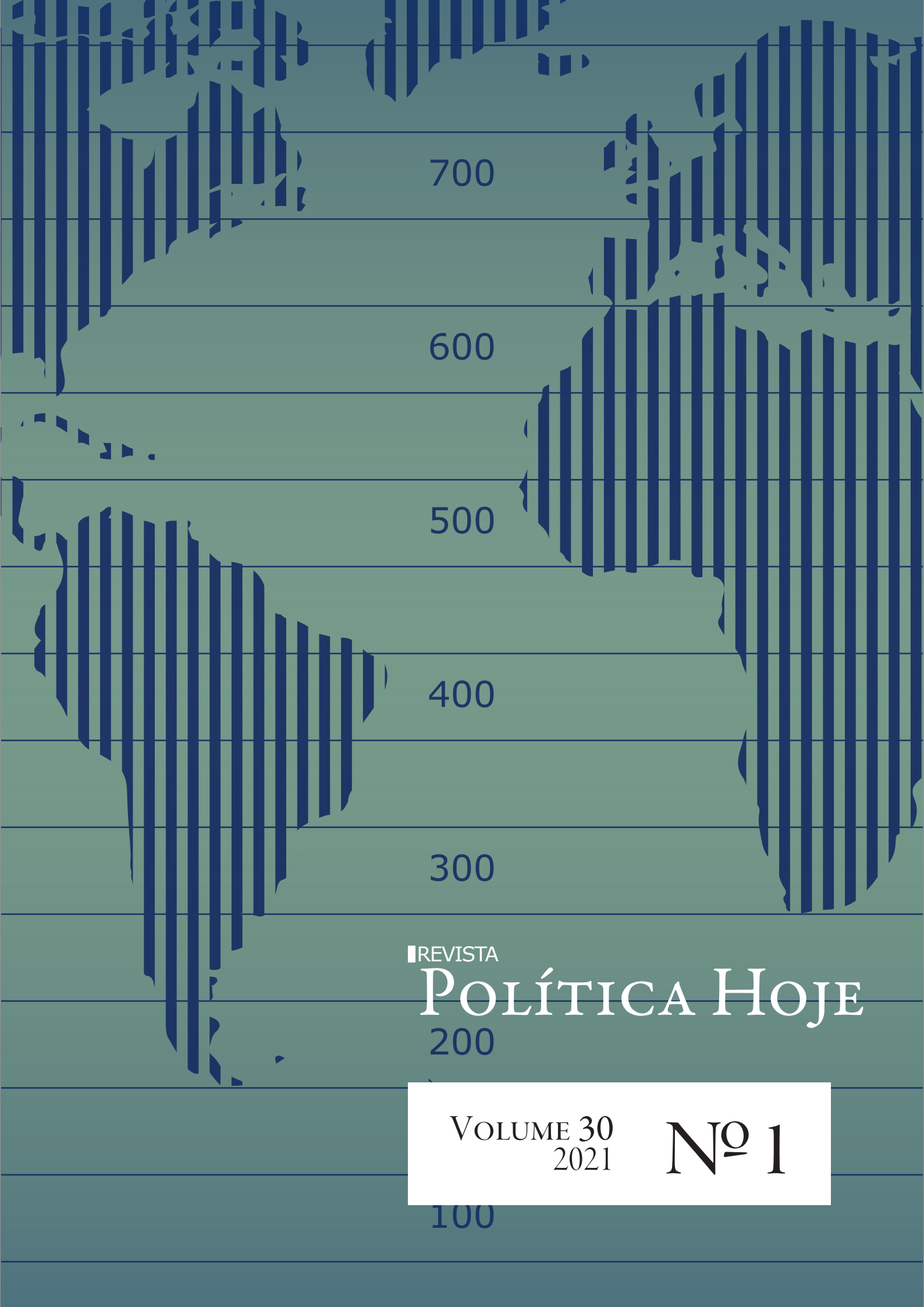


REVISTA
Política Hoje

Volume 30
2021

Nº 1

REVISTA

Política Hoje

Equipe Editorial:

Professores Organizadores
Rafael Mesquita (UFPE)
Virgínia Rocha (UFPE)
Palloma Marciano (UFPE)
Dalson Britto Figueiredo Filho (UFPE)

Editor Chefe
Ernani Carvalho (UFPE)

Editora Executiva
Ulle Ráfaga Campos e Figueiredo (UFPE)

Capa e Diagramação
Laís Marques Arcelino de Macedo

Conselho Editorial:

UNB, Brasil	André Borges
UFRGS, Brasil	André Marenco
IPESPE, Brasil	Antônio Lavareda
UFMG, Brasil	Carlos Ranulfo
UFBA, Brasil	Celina Souza
FGV, Brasil	Cláudio Couto
IBGE, Brasil	Eduardo Leoni
USP, Brasil	Eduardo Marques
UTDT, Argentina	Enrique Peruzzoti
UFPE, Brasil	Flávio da Cunha Rezende
FGV, Brasil	George Avelino
CIDE, México	Julio Ríos Figueroa
UDESA, Argentina	Lucas Gonzalez
UNB, Brasil	Lúcio Rennó
UDESA, Argentina	Marcelo Leiras
UFPE, Brasil	Marcelo Medeiros
UFPE, Brasil	Marcus André Melo
USP, Brasil	Marta Arretche
USP, Brasil	Matthew Taylor
UFPE, Brasil	Mauro Soares
UERJ, Brasil	Miriam Saraiva
UFRGS, Brasil	Paulo Peres
USP, Brasil	Rafael Duarte Villa
UEL, Brasil	Raquel Kritsch
UFPE, Brasil	Ricardo Borges Gama Neto
UFSCAR, Brasil	Simone Diniz
UERJ, Brasil	Thamy Pogrebinschi
UNESP, Brasil	Tullo Vigevani
USP, Brasil	Wagner Pralon Mancuso

POLÍTICA HOJE – Revista Semestral do Programa de Pós-Graduação em Ciência Política, da Universidade Federal de Pernambuco. Recife, Vol. 30, n.1, 2021.

13. Ciência Política-Periódicos

REVISTA
POLÍTICA HOJE

Universidade Federal de Pernambuco
Centro de Filosofia e Ciências Humanas – 14º andar. CEP: 50670-901
Tel/Fax: (81) 2126-8283 / 2126-8922
E-mail: politica.hoje@ufpe.br
periodicos.ufpe.br/revistas/politicahoje
ISSN 0104-7094

Índice

05 **Apresentação**
Virginia Rocha - Palloma Marciano

DOSSIÊ

11 **Methodological Transformations in Contemporary Political Science**
Flávio da Cunha Rezende

40 **Um mapa para a transparência e replicabilidade na ciência social empírica: o Protocolo TIER**
Amanda Domingos - Ian Rebouças Batista

64 **A map for transparency and replicability in empirical social science: the TIER Protocol**
Amanda Domingos - Ian Rebouças Batista

87 **Uma Introdução à Bibliometria e Cientometria: o caso do presidencialismo latino-americano**
Amanda Sangalli - Samira Kauchakje

124 **An Introduction to Bibliometrics and Scientometrics: the case of Latin American presidentialism**
Amanda Sangalli - Samira Kauchakje

161 **A abordagem da grounded theory como alternativa metodológica qualitativa para pesquisas em Ciência Política**
José Alberto de Siqueira Brandão

180 **Grounded theory approach as a qualitative methodological alternative for Political Science research**
José Alberto de Siqueira Brandão

197 **DA TEORIA À ANÁLISE: Uma introdução ao uso de entrevistas individuais semiestruturadas na ciência política.**
Virgínia Rocha

DOSSIÊ

226 **From theory to analysis: An introduction to using semi-structured individual interviews in political science**
Virgínia Rocha

252 **Ideologia e Partidos no Brasil: reflexão e prática a partir dos programas de governo**
Nara Salles

296 **Ideology and Parties in Brazil: considerations and methods using party**
Nara Salles

339 **Mapeando comportamentos com estimação de pontos ideais**
Rodrigo Martins

368 **Mapping behavior with ideal point estimation**
Rodrigo Martins

395 **Tempo, tempo, tempo: regressão de Cox na ciência política**
Rodrigo Lins

415 **All kinds of time: Cox regression in political science**
Rodrigo Lins

434 **Uma introdução à regressão com dados de painel**
Rafael Mesquita - Antônio Alves Tôrres Fernandes - Dalson Britto Figueiredo Filho

472 **An introduction to panel data regression**
Rafael Mesquita - Antônio Alves Tôrres Fernandes - Dalson Britto Figueiredo Filho

Apresentação

Uma abordagem didática aos métodos de pesquisa em Ciência Política e Relações Internacionais: Em busca de uma ciência mais rigorosa e acessível

Em 2005, Soares pautou o debate sobre os desafios enfrentados pela ciência política brasileira para o uso e ensino de métodos quantitativos e qualitativos na área. O autor ressaltou uma rejeição pelos pesquisadores em relação aos métodos quantitativos, ao passo que as técnicas qualitativas, apesar de preferidas, não eram aplicadas com rigor. Estudos posteriores indicam que houve avanços, mas ainda existem desafios a serem superados. Neiva (2015) encontrou que apenas 5,2% dos trabalhos publicados em revistas das ciências sociais classificadas com Qualis B3 ou mais apresentam técnica estatística avançada. Por outro lado, Barberia, Godoy e Barboza (2014) argumentam que houve progresso no ensino de métodos no Brasil, haja vista o aumento de oferta de cursos de metodologia. Essa melhoria, no entanto, ocorreu de forma concentrada em algumas universidades e técnicas de pesquisa.

Numa perspectiva mais recente, Figueiredo, Fernandes, Borba e Aguiar (2020) demonstram que houve incremento no uso de métodos em artigos publicados em 6 periódicos relevantes na ciência política no Brasil. Houve crescimento na quantidade de trabalhos que usam métodos quantitativos ou qualitativos, mas esses últimos possuem predominância do uso de entrevistas.

No que tange às Relações Internacionais, Medeiros, Barnabé, Albuquerque e Lima (2016) analisam artigos publicados entre 2006 e 2014, acessados pela plataforma ScieELO, e identificam que apenas 12,94% da amostra analisada fazia menção ao método utilizado, o que os pesquisadores chamam de "inconsciência metodológica". Dentre esses trabalhos, a maioria deles usava métodos qualitativos, seguidos de análise quantitativa, métodos mistos e modelos formais. Já Albuquerque, Lima e Brito (2020) apresentam um mapeamento da formação de internacionalistas na América do Sul. Seus resultados preliminares, que vão na mesma linha dos achados de Barberia et al. (2014), revelam que 83% dos cursos de pós-graduação ofereciam ao menos 1 ou 2 disciplinas de metodologia. Ainda encontram que os conteúdos das ementas são similares aos aplicados na Ciência Política, ambos tendo obras em comum como base das disciplinas. Neste sentido, há prevalência de artigos e livros internacionais em relação a autoras/es nacionais.

De maneira geral, é notável que houve avanço na formação de estudantes de Ciência Política e das Relações Internacionais em metodologia de pesquisa, bem como no uso de métodos na produção científica nessa grande área. Embora tais mudanças sejam relevantes, ainda parece haver um longo caminho até a superação do calcanhar meto-

dológico. Os avanços indicados pela literatura não fazem, em geral, um recorte sobre gênero, raça ou região, por exemplo. Mas a discussão sobre o uso e ensino de métodos na ciência política e nas relações internacionais não pode ser descolada dos debates sobre desigualdades.

Candido, Feres Junior e Campos (2019) observam a composição da docência dos programas de Pós-Graduação em Ciência Política e seus achados indicam desigualdade de gênero e profunda desigualdade racial. Essa desigualdade é maior do que aquelas existentes nas áreas de Sociologia e Antropologia. Barberia et al. (2014) encontram que programas mais antigos, e considerados, portanto, mais institucionalizados, são aqueles que possuem maior oferta de disciplinas de métodos. Além disso, mesmo os programas com melhor desempenho na análise dos/as autores/as, possuem um número substancialmente menor ao dos cursos oferecidos em programas internacionais (BARBERIA et al., 2014). Um levantamento da CAPES de 2019, como explicam Colombo et al. (2020), revela que 60% dos Programas de Pós-Graduação em Ciência Política brasileiros estão concentrados na região Sudeste.

Coelho, Elias e Santos (2020) avaliam a formação e docência na ciência política e nas relações internacionais e encontram dados que corroboram estudo anterior (Coelho, Santos e Elias, 2019) que demonstra manutenção do gender gap, mesmo diante da expansão das pós-graduações da área nos últimos anos. Coelho et al. (2019) verificam desigualdade constante entre homens e mulheres em termos de quantidade de publicações, bem como a maior propensão de homens serem os primeiros autores dos trabalhos. Em termos de estratégias de internacionalização, Rangel (2020) pontua que homens e mulheres possuem trajetórias internacionais similares de maneira geral - no que se refere às publicações e à formação, por exemplo - mas mulheres passam, em média, menos tempo nas atividades realizadas fora do país. A autora observa essa discrepância a partir de demandas e tarefas de cuidado exercidas pelas mulheres, que pode afetar o tempo disponível que as pesquisadoras possuem para se dedicar às experiências internacionais.

Nesse cenário, iniciativas que democratizem o acesso ao conhecimento e que promovam a divulgação de materiais sobre metodologia de forma didática podem contribuir para que sigamos avançando no caminho para uma ciência mais rigorosa, transparente e pluralizada. Esforços importantes têm sido feitos, como o trabalho essencial da área temática sobre "Ensino e Pesquisa em Ciência Política e Relações Internacionais", da ABCP; o Prêmio Anpocs de Divulgação Científica em Ciências Sociais; além de grupos de pesquisa dos Programas de Pós-Graduação e de atividades de discentes relacionados à metodologia, como a iniciativa "Métodos em Pauta", por exemplo, composta por pós-graduandos e egressos do PPGCP-UFPE.

Nesta direção, o Dossiê intitulado "Uma abordagem didática aos métodos de pesquisa em Ciência Política e Relações Internacionais: Em busca de uma ciência mais rigorosa e acessível" tem como objetivo apresentar alguns dos principais métodos de pesquisa em Ciência Política e Relações Internacionais com um objetivo expressamente didático. Os artigos deste dossiê apresentam diferentes técnicas, dentre qualitativas e quantitativas, definindo as abordagens, discutindo suas limitações e apresentando as etapas necessárias para aplicar a técnica de forma adequada. A edição é bilíngue e os artigos possuem uma versão de apresentação dos trabalhos (em powerpoint ou PDF), a fim de facilitar o acesso e o aproveitamento pedagógico dos materiais. Também são disponibilizados os scripts dos artigos, não só para efeitos de transparência e replicabilidade, mas também para facilitar o aprendizado dos/as estudantes.

Os artigos que compõem este Dossiê foram organizados da seguinte maneira. Como ponto de partida, o artigo intitulado "Methodological Transformations in Contemporary Political Science", de Rezende (2021), procura entender e identificar o conjunto de instituições que têm influenciado a produção de conhecimento científico na Ciência Política nas últimas duas décadas. O autor argumenta que há uma transformação metodológica na área, acarretada pelas mudanças nos valores, crenças e parâmetros sobre inferência causal. O artigo oferece, portanto, uma visão minuciosa da evolução do debate inferencial na Ciência Política e traz uma compreensão aprofundada sobre os padrões atuais de cientificidade que norteiam os debates metodológicos de ponta na área.

Ainda no que se refere às práticas de pesquisa numa ótica mais ampla, Domingos e Batista (2021), no artigo "Um mapa para a transparência e replicabilidade na ciência social empírica: o Protocolo TIER", oferecem o que podemos chamar de "o caminho das pedras" para obter mais transparência nas pesquisas. Os autores apresentam uma maneira de utilizar o Protocolo TIER (Teaching Integrity in Empirical Research) para documentar todo o processo de pesquisa e os materiais produzidos por ela. Ao longo do texto, é possível identificar as três dimensões do padrão de replicação: substantiva, pedagógica e transparente. Uma parte do trabalho é fornecer elementos de exercício para compreensão prática do protocolo. Esse trabalho é o ponto inicial para quem procura desenvolver pesquisas de forma crível e replicável.

Já Sangalli e Kauchakje (2021), no estudo chamado "Introdução à bibliometria e cientometria: exemplo prático de aplicação ao tema presidencialismo latino-americano", procuram explicar a definição, o objetivo e a aplicabilidade da bibliometria e da cientometria na ciência política. As autoras apresentam tais métodos como procedimentos para medir o desenvolvimento da ciência no que tange a produção científica, as trajetórias das áreas de pesquisa e até mesmo como está distribuída a produção em diferentes centros. A partir de um exemplo de coleta de dados das coleções bibliográficas Web of Science e SciELO Citation Index sobre o presidencialismo latino-americano, as autoras explicam o passo-a-passo para utilização dos métodos citados, discutindo também as vantagens e os limites para a sua aplicação na Ciência Política.

Dado que o objetivo deste dossiê é apresentar métodos relevantes nas áreas de ciência política e relações internacionais, tanto para a análise qualitativa, quanto para a análise quantitativa, apresentamos estudos que trazem explicações práticas e didáticas nesses dois campos. No que tange a análise qualitativa, Brandão (2021), em trabalho chamado "A abordagem da Grounded Theory como alternativa metodológica qualitativa para pesquisas em ciência política", apresenta um estudo bibliográfico sobre a Grounded Thoery, também chamada Teoria Fundamentada, para elucidar sua definição e desenvolvimento enquanto possível abordagem metodológica. Em um segundo momento, o autor explana as principais etapas para aplicação da teoria, a partir de uma posição construtivista. Tendo em vista a atual predominância da posição positivista na Ciência Política (MOSLEY, 2013), o trabalho traz um aporte teórico e prático que pode melhor orientar estudantes e pesquisadoras/es interessadas/os na técnica.

O que são entrevistas semiestruturadas e como elas podem ser usadas na ciência política? Essa é a pergunta que Rocha (2021) procura responder em seu artigo, intitulado "Da teoria à análise: Uma introdução ao uso de entrevistas individuais semiestruturadas na ciência política". A autora versa sobre as entrevistas em profundidade semiestruturadas, discutindo seu conceito e apresentando exemplos práticos do seu uso na Ciência Política como uma

técnica que possibilita o acesso a dados como sentimentos, emoções e crenças dos indivíduos. Para isso, descreve as principais etapas para a execução dessa técnica qualitativa de coleta de dados - aporte teórico, elaboração do roteiro, seleção dos/as entrevistados/as, aprovação do comitê de ética (quando necessário), aplicação das entrevistas e compilação e análise dos dados. Tal discussão é feita à luz dos desafios de validade e confiabilidade dos dados coletados e dos limites da transparência na utilização da técnica.

Além disso, quatro estudos abordam técnicas quantitativas como classificação ideológica dos partidos, pontos ideais, regressão de Cox e regressão em painel. No primeiro caso, chamado "Ideologia e Partidos no Brasil: reflexão e prática a partir dos programas de governo", Salles (2021) questiona se é possível classificar ideologicamente os partidos através dos programas de governo apresentados por seus candidatos. A autora discute como as propostas de governo podem fornecer informações e atalhos para classificação ideológica de atores e partidos políticos. Segundo a pesquisadora, essa análise consegue extrair quais são os temas salientes para esses grupos e partir deles posicioná-los entre direita e esquerda. Iniciativas como o "Manifesto Project/MARPOR" e os avanços mais recentes das técnicas de análise automatizada de texto dão a base metodológica para a comparação entre programas de forma estruturada. A análise automatizada de texto wordfish é utilizada no trabalho para a classificação de 889 programas de governo nas três esferas (federal, estadual e municipal). A partir disso, a autora busca identificar as posições dos partidos no espectro programático e se há constância nos três níveis de disputa.

Por outro lado, a estimação de pontos ideais é uma das metodologias mais utilizadas para entender as preferências de partidos políticos e parlamentares. No estudo "Mapeando comportamentos com estimação de pontos ideais", Martins (2021) apresenta uma forma simples e intuitiva para explicar a técnica e sua aplicação em pesquisa. Segundo, Martins o método é um derivado do modelo espacial do voto construído por Downs (1957) que tinha como objetivo demonstrar posições políticas em um espaço unidimensional. Então, a estimação de pontos ideais simplifica e permite a comparação da distância das preferências dos indivíduos analisados para fornecer um mapa de posicionamentos a respeito de uma votação dividida de projeto lei no parlamento, por exemplo. Assim, as votações anteriores podem indicar quais serão as prováveis preferências dos legisladores em futuras votações. O texto apresenta também as variações e modelos populares utilizados pelos principais trabalhos da área como NOMINATE e seus derivados, de Pool e Rosenthal, e os modelos bayesianos. É possível ao final da leitura, aplicar um modelo não paramétrico para os dados de comportamento da Câmara dos Deputados no início do governo Bolsonaro (2019).

No que tange às diversas técnicas de regressão existentes, Lins (2021) e Mesquita, Fernandes e Britto (2021), apresentam as regressões de Cox e de painel, respectivamente. No primeiro caso, no artigo "Tempo, tempo, tempo: regressão de Cox na ciência política", Lins (2021) apresenta uma visão detalhada do modelo de regressão de Cox – técnica comumente usada em desenhos de análise de sobrevivência. Além de uma abordagem teórica que expõe a definição e uso desse modelo de regressão na ciência política, o autor traz um glossário que introduz conceitos específicos da técnica em questão. Por fim, Lins traz uma aplicação prática do modelo de regressão de Cox no software R, explicando cada etapa do processo.

Já no estudo "Uma introdução à regressão com dados de painel", Mesquita, Fernandes e Britto (2021) trazem uma introdução à regressão de painel. Em primeiro lugar, apresentam um balanço do baixo uso da técnica na área

de Ciência Política e Relações Internacionais brasileira através da análise de conteúdo de revistas acessadas na Scientific Electronic Library Online Platform (Scielo). Em seguida, os autores apontam como a combinação de informações espaciais e temporais traz mais confiabilidade nos resultados de determinados desenhos de pesquisa e indicam as principais diretrizes do método, desde os conceitos ao modelo matemático. Segundo eles, sua aplicação traz benefícios em quatro aspectos: em termos de mensuração, controle de confounders, aprimoramento de inferências causais e tamanho da amostra. Por fim, trazem uma aplicação da técnica com dados sobre a relação do crescimento econômico e a eleição de presidentes, finalizando com uma seção de Perguntas e Respostas com os principais pontos do tema.

É evidente que um único dossiê não é capaz - e nem pretende ser - de esgotar a ampla gama de métodos de pesquisa aplicados na ciência política e nas relações internacionais. Algumas ausências importantes são uma discussão mais ampla sobre inferência causal e as vantagens e os desafios da utilização de desenhos experimentais na nossa área. Outra importante questão reside nas discussões sobre ética em pesquisa, incluindo tanto pesquisas experimentais, quanto observacionais. Embora esses aspectos sejam observados nos artigos sobre transformações metodológicas na ciência política (REZENDE, 2021) e sobre transparência e replicabilidade (DOMINGOS; REBOUÇAS, 2021), há uma larga quantidade de aspectos importantes desse - e de outros - debates que poderiam ser explorados em dossiês futuros, que possuam também uma abordagem prática e didática. Para além das limitações, esperamos que este dossiê possa contribuir para o aumento na oferta de material de apoio de qualidade para docentes e estudantes da área e para incentivar que mais publicações com este objetivo continuem a ser produzidas.

Pessoalmente para nós, Palloma Marciano e Virginia Rocha, construir esse projeto junto a Dalson, Rafael, Ulle e as/os colaboradoras/es, foi uma forma de retribuir o investimento feito em nossa formação. O início da nossa graduação (somos contemporâneas na UFPE) foi marcado pela insipiência de material com linguagem introdutória e traduzido para as ciências sociais nas disciplinas de metodologia. Hoje, apresentar este Dossiê é garantir mais ferramentas para quem forma e quem se forma cientista. Desejamos uma boa leitura!

Virginia Rocha
Palloma Marciano

REFERÊNCIAS BIBLIOGRÁFICAS

BARBERIA, G. L.; GODOY, S. R. de; BARBOZA, D. P. Novas Perspectivas sobre o 'Calcanhar Metodológico': O Ensino de Métodos de Pesquisa em Ciência Política no Brasil. Teoria & Sociedade, 2014.

CANDIDO, M. R.; FERES JÚNIOR, J.; CAMPOS, L. A. Desigualdades na elite da Ciência Política brasileira. Civitas. v. 19, n. 3, p. 564-582, 2019.

COELHO, A.; ELIAS, I.; SANTOS, V. A Participação das Mulheres na Produção Acadêmica da Área de Relações Internacionais no Brasil. Mural Internacional. v. 10, 2019.

COELHO, A.; ELIAS, I.; SANTOS, V. As Mulheres na Pós-Graduação nas áreas de Ciência Política e Relações Internacionais no Brasil: um olhar sobre a Formação e a Docência. 12° Encontro da Associação Brasileira de Ciência Política. João Pessoa, 19 a 23 de outubro de 2020.

COLOMBO, L. A.; SANTANA, L. da C. F.; NASCIMENTO, E. O. do. Um Mapeamento dos Cursos de Graduação e de Pós-Graduação em Ciência Política no Brasil: Avanços e Perspectivas. 12° Encontro da Associação Brasileira de Ciência Política. João Pessoa, 19 a 23 de outubro de 2020.

FIGUEIREDO, D.; FERNANDES, A.; BORBA, L.; AGUIAR, T. H. Metodologias de Pesquisa em Ciência Política: Uma Breve Introdução. BIB. 2020.

MEDEIROS, M. de A.; BARNABÉ, I.; ALBUQUERQUE, R.; LIMA, R. What does the field of International Relations look like in South America? Revista Brasileira de Política Internacional. v. 59, n. 01, 2016.

MOSLEY, Layna. (2013), Interview Research in Political Science. Cornell Univ. Press: Ithaca.

NEIVA, P. Revisitando o calcanhar de Aquiles metodológico das ciências sociais no Brasil. Sociologia, Problemas e Práticas. 2015.

OLIVEIRA, L.; NICOLAU, J. Métodos e Metodologias da Ciência Política no Brasil: Uma Análise dos Currículos de Pós-Graduação. 8° Encontro da Associação Brasileira de Ciência Política. Gramado, 1 a 4 de Agosto de 2012.

SOARES, G. A. D. O Calcanhar Metodológico da Ciência Política no Brasil. Sociologia, Problemas e Práticas. n. 48, 2005, p. 27-52.

Methodological Transformations in Contemporary Political Science[1]

Flávio da Cunha Rezende[2] - Universidade Federal de Pernambuco

Abstract

This article seeks to identify, map, and understand a set of institutions – understood as values, beliefs, and parameters – that structure the scientific knowledge in political science over the past 20 years. Its basic purpose is to map the "fundamental" values that had produced a new paradigm for academic production in the contemporary political science. The central argument of the paper is that Political Science is moving throughout a methodological transformation in which causal inference is pursued by means of several research design types, condition I named as inferential pluralism.

1 This paper was originally written for the journal Política Hoje. The author thanks the CNPQ for financing studies on research design and causal inference in political science over the past five years. These resources were key in enabling the continuation of empirical research in the contemporary production of high-impact journals in political science.

2 Ph.D Cornell University, United States. Political Scientist. Associate Professor IV in UFPE's Political Science Department. CNPQ Productivity Researcher. Coordinator of the Group in Epistemology and Comparative Method in Political Science – CNPq. Email: rezzende2005@gmail.com

1. Introduction

A detailed examination of high-impact journals in contemporary Political Science reveals a high number of transformations in the "scientificity" of political science over the past two decades. Although political science still studies classical themes such as collective action, electoral behavior, democratic institutions, or the paradox of voting, the substantive nature of the knowledge produced today is considerably different than it was a quarter of a century ago. Political science is going through a massive redefinition of its methodological bases, its technology, and its conception of what "science" is, in such a way that, from an objective standpoint, producing knowledge in political science today is a substantially different task.

As it is widely known by institutional theories, "changes" usually occur through the introduction of new values, beliefs, or parameters that are shared by the agents, organizations, and institutions that produce the action. In the case of the scientificity transformations in political science, which is of particular interest to us, it is plausible to state that there is a set of new shared institutions[3] in the scientific community that have substantially altered the way of producing, organizing, and diffusing scientific knowledge from at least the 90s. The current production of valid knowledge requires of political scientists the adherence to a set of values that give meaning to the new standard of scientificity observed in the discipline.

Understanding Methodology, in its broadest sense, as the systematic study of the rules that give meaning and validate the production of knowledge we are, ultimately, concerned with comprehending the precise form "Methodological Transformations" have occurred in the Political Science scientific community. These transformations, produced by the introduction of a new set of beliefs and values considered necessary and sufficient to give "scientificity" to political science, deeply affect the rigor, nature, and meaning and logic of the knowledge produced in the past decades.

Comprehending these institutions in the particular case of contemporary political science is the basic motivation of this article. It is focused in presenting in a broad and systematic way the specific set of "beliefs, values, and parameters" that, in an integrated fashion, allowed the structuring of a new conceptualization of scientificity. The main question this paper aims to answer is precisely: "what are the values and beliefs that shape the production of knowledge in political science, and why do these introduce a new concept of science to the study of politics?".

To answer this question, this paper looks to identify, map, and understand more broadly the institutions – understood as values, beliefs, and parameters – that structure scientific knowledge of politics currently, backed by an ample examination of works published in scientific journals over the past 20 years. For the sake of parsimony, we will call contemporary political science "post-kkv political science[4]". A considerable part of these beliefs and values emerge from the large methodological production that flows from the crucial work Designing Social Inquiry (King, Keohane, and Verba, 1994), which has reconfigured methodological debates within political science since.

3 In contemporary institutional theories there is a broad conceptualization of what are "institutions". I align here with Ostrom and Hess (2011. p.42), who define institutions "as formal and informal rules that are understood and used by a community. Institutions, as we use the term here, are not automatically what is written in formal rules. They are the rules that establish the working "do's and don'ts" for the individuals in the situation". For this paper in particular, agents are the researchers and the knowledge-producing institutions; and the behavior to be understood is the production of knowledge in political science.

4 The use of "post-KKV" was influenced due to Mahoney (2010). He elaborates an analysis of the transformations in the reflexivity of the so-called New Qualitative Methodology field for political science, produced by the book Designing Social Inquiry. Several key works in contemporary methodology have used this work as a watershed in precipitating substantial transformations in the way of perceiving political science, especially when it comes to the issue of being guided by inferences and the unification of the quantitative and qualitative logics.

The basic goal of the paper is, therefore, to map the "fundamental" values that offer a new scientificity for contemporary political science in an integrated way. This mapping will be conducted from the identification of the main constitutive elements that "give meaning" to the production of knowledge in post-KKV political science. I will identify these values from dimensions considered essential by the scientific community.

The central argument of the paper is that Political Science is going through a profound transformation in its scientificity over the past two decades and this is what differentiates it from other social sciences fields. There is an introduction of new institutions in political science that given meaning to the new standards of scientificity. These institutions are understood from seven interconnected dimensions. The first dimension is the Growing Concern with Causal Inference, where a growing concern with causation in Political Science is observed, especially with the production of causal inferences. More broadly, it is observed that several logics of how to produce inferences coexist, developing what Rezende (2015) conceptualizes as *inferential pluralism*.

The second dimension considered is Orientation by Hypothetico-Deductive Models, which finds support in the extensive implementation of the so-called Empirical Implications of Theoretical Models (EITM). These gain ample space in the formation of new generations in political science. Political science is substantially guided by the adoption of hypothetico-deductive models (H-D), with an emphasis in hypothesis testing, massive use of formal models, and a deep interaction with econometrics and statistical analysis.

The third institution considered here is the connection between Research Design and Inferential Quality. The central importance of research designs as elements considered fundamental to create inferential quality is pivotal condition for KKV political science. This relationship has enabled more and more the coexistence between several types of research design, which become more rigorous to elevate the chances of producing inferential quality.

The fourth fundamental institution resides in the Integration of Quantative and Qualitative Methods as desirable scientific research. Political science has broadened its possibilities to allow plausible alternatives that boost the advantages of studies that adequately mobilize the logic of the "two cultures" (Goertz and Mahoney, 2012). The integration of methods becomes an ideal more valued and sought after in contemporary political science.

The fifth dimension compared here attempts to understand Analytical Eclecticism, which has been one of the fundamental presuppositions of scientificity. This is when scientists start to soften the traditional "paradigmatic barriers" between theories to increase explanatory power and deal in a more satisfactory way with the usual gap between theories and useful applications. Contemporary political science has been progressively "pushed" towards handling more complex problems without losing theoretical rigor. Analytical eclecticism proposes clear alternatives in this direction, and it has been one of the important constitutive elements in the beliefs regarding the new scientificity in political science.

The sixth dimension considered is the Growing Demand for Causal Identification. Much influenced by the interaction with applied econometrics, contemporary political science, especially one more empirically-minded and concerned with inference, has increased its attention towards the need to understand more closely identification strategies. The demand for more precise inferences, understood as parameter estimation, begins at a clearer search for identification.

Lastly, another institution that is considered fundamental is the Primacy of Methods over Political Analysis. Over the last few years, political science has gone through an adherence to the known argument by Rubin (2008) that "design trumps analysis", in which methods have prevalence over analysis. In economics, this statement structured a complete "credibility revolution", in which research designs and inference production methods became much more rigorous. This belief has had similar impact (although with less intensity) in political science, enacting a situation that can be known as the primacy of methods (and research designs) over political analysis. Considerable empirical evidence of this process has been the progressive strengthening and institutionalization of the "political methodology" disciplinary field over the past 20 years.

The paper is structured thusly. In the first section, the conditions that made contemporary political science be fundamentally different at present are more broadly discussed. I will focus on four key factors: the notion of cosmopolitan political science proposed by Norris (1997); the publication of the work by KKV in 1994, which redefines beliefs concerning scientificity in political science; the considerations by Shapiro (2004) on the issue of the disconnect between political science and the problems of reality (usually referred to as "flight from reality"); and finally, the extensive technological and informational changes that condition the production of knowledge in political science. Next, each of the beliefs considered fundamental in this paper[5] are analyzed "in isolation" in each subsection. Lastly, final considerations are offered and their basic implications for the "practical" production of knowledge in political science. The goal is to broaden the awareness of researchers in understanding the larger "context" in which each work done in political science is inserted in, being profoundly affected regarding its scientificity.

2. CONTEMPORARY POLITICAL SCIENCE AND ITS CONDITIONING FACTORS

In the 1990s, the conditions were gathered to enable an ample set of transformations that occurred in the scientificity standards within Political Science. These changes are, in large part, furthered by some endogenous factors that are associated, on one hand, to the methodological developments that happened within the discipline; and, on the other, as a result of important exogenous factors that were produced from broader changes in the economy and society[6].

The key endogenous factors identified in this paper are, essentially, three[7]. First, the notion of Cosmopolitan Political Science, proposed by Norris (1997), which begins to represent the "shared" nature of the beliefs and values that guide the production in political science, especially when it comes to research agendas, theories used, and essentially the methods and techniques mobilized by the scientific community. The second is directly associated with the methodological impact of the work Designing Social Inquiry by King, Keohane, and Verba (1994), which introduced new scientificity parameters and, especially, brought to the surface the need to focus on the issue of causal inference in political

5 For the sake of parsimony, it is important to realize that each one of these beliefs would require a more extensive treatment than what is offered here. The goal is not to conduct an "exhaustive review" of the methodological literature, for the sake of objectivity, to give the basic elements that allow to minimally comprehend the relevance of each for the transformation of scientificity in political science. It would not be an exaggeration to state that, in light of the enormous literature available, each section would need an individual article.

6 As it is widely known in theories of change (Mahoney and Thelen, 2010; Rezende, 2012), these are best explained by gradual approaches in which the endogenous and exogenous causes are considered.

7 These factors were selected to ensure a good reduction of the problem and to identify the main conditions that brought about a substantial modification of the values concerning political science's scientificity. These factors unleashed a series of other intentional and unintentional consequences on the organization and scientific production in the discipline and cannot be explored here due to space.

science. Lastly, I consider as fundamental the important reflection proposed by Shapiro (2004), which puts forward the issue of the gap between theory and reality in political science and points to the imperative need of using political science to deal with relevant problems attached to empirical reality.

Regarding exogenous factors to be considered in the analysis, I will mention the important dimension of scientific production in contemporary political science being conditioned by the transformations brought about by processes of diffusion of democracy, globalization, and especially, the associated technological changes. The most visible effect of these processes on political science's scientificity is structured from the ample availability of new technologies, data analysis software, and databases (information). In a radically distinct way from any other period in the history of the discipline, political science has plenty of resources to produce scientific knowledge.

Analyzing production in important European and North American journals between the 1970s and 90s, Norris (1997) gathers evidence to argue that Cosmopolitan Political Science is a reality. Indeed, political scientists, in several contexts and realities, have focused on ever more similar research problems, methods, and techniques. Consequently, the intensification of globalization and the increased interaction of researchers from diverse traditions, deepened over the last quarter of a century an intense process of "isomorphic convergence" of research agendas, concern with methodological rigor, and theoretical sharing, as well as the proliferation of analytical strategies for the development of inferential compared analysis.

To Norris, the condition of Cosmopolitan Political Science introduces an important set of transformations that decisively impact values concerning scientificity. The professionalization of practices and values from sharing regional and national institutions enables, on one hand, that the standards of production be more and more "globalized". On the other hand, an intensification of the sharing processes of the knowledge produced regarding research problems, methods, and theories becoming more common is observed.

Concerning scientific production itself, given the notion of sharing, cosmopolitan political science is guided by five key processes. Methodological convergence becomes a growing reality. Political scientists become more "engaged" by methodologies that are shared more. The convergence of research agendas among different realities. The increase of internalization processes of scientific publications. And a considerable expansion of shared publications involving authors from different realities.

The second endogenous factor considered in this paper is the important contribution by KKV (1994) in Designing Social Inquiry. This pivotal methodological work launches the framework for introducing new scientificity standards. Based on the presupposition of science as a shared activity[8], the authors highlight a conception of scientificity based on four fundamental values: a) the centrality of causal inference; b) the primacy of methods; c) more transparency in scientific production; and, d) the uncertainty of the knowledge produced.

The crucial contribution by the authors is due to focusing on causal inferences and its connection to research designs. Inferential quality in political science is desirable, possible, and linked to research design. In an original way, they argue that quantitative and qualitative research traditions are similarly guided by the same rules of inference.

8 KKV (1994) consider that scientific activity as a shared among values by the scientific community. They state that "science at its best is a social enterprise" (p.9). The authors believe that the distinctive feature of a scientific contribution is a combination of two elements: adherence to the scientific principles and using "public" methods and procedures to produce valid causal inferences.

Consequently, post-KKV political science, as an empirical science, was unprecedentedly guided by the pair research design-causal inference. Research designs may grant quality causal inferences, increasing reliability, validity, transparency, and honesty in the analyses. To do so, they must be aligned with the rules of inference.

This methodological orientation profoundly changed the scientific status, rigor, and research practices within the discipline. Post-KKV political science introduced a new universe of problems and methodological issues concerning to matters of causation, inference, estimation, modelling, and more broadly, the quality of research designs. The institutional consolidation of the field known as Political Methodology, which will be explored further down, is a direct consequence of this professional concern.

Another argument considered an important reference in transforming the discipline's scientificity was the one made by Shapiro (2002), backed by a strong critique of the "pathologies" of Rational Choice models that became hegemonic in the discipline in the 90s. Shapiro believes that Political Science must be useful and solve the so-called "flight from reality" problem, which expresses the usual gap between theory (and methods) and real world problems.

The main issue faced by a science that is strongly rigorous, formal, and guided by causal inference, is its "natural propensity" to distance itself from problems that affect the real world and actually matter (from a theoretical and empirical standpoint). Shapiro believes it necessary that Contemporary Political Science be concerned with a stronger adherence to being relevant to real world problems.

This presupposition's main overall implication is that models and theories developed within the discipline must be continuously confronted with relevant and empirically treatable problems. There is nothing wrong in establishing the scientificity of the discipline on causal inferences but, it is necessary that political science be capable of creating useful knowledge.

Having presented the endogenous factors that contribute towards the transformation of scientificity in Political Science, I will examine the exogenous factors. I highlight the important fact that production of knowledge in the discipline today is strongly conditioned by technological transformations promoted by combined processes in democracy, globalization, and changes in the sense of an information society, which become prevalent in the 1990s.

Scientificity in Political Science is radically altered by the unprecedented affordance of technological resources and the growing availability of information (and access) in the form of databases shared by researchers. Political scientists currently count on a massive array of tools to collect, organize, and analyze data, a fact that makes knowledge production ever more "shared" among different realities and able to accomplish analytical tasks that were unthinkable 20 years ago. The breadth of analytical possibilities becomes vastly larger at present.

Arguably, we are experiencing a political science in the "big data" age, in which a massive amount of data is produced and accessed at an astonishing speed, by scientific knowledge "producers" and "consumers". Monroe (2013) states that there are five fundamental characteristics that guide big data production in Political Science: volume, variety, velocity, vinculation, and validity. The scientificity standard exhibited by the discipline enables the increase of interactions between computational science, political science, and the so-called data science.

Monroe et alli (2015) believe that the interaction of Political Science with these fields has introduced substantial innovations for collection, manipulation,

management of data, information extraction, as well as an ample set of inferential techniques such as "statistical or machine learning". The authors argue that the use of "big data" is perfectly compatible with the production of better causal inferences or with the methods mobilized by political science or, more broadly, by the social sciences.

Producing scientific knowledge in a context in which a considerable quantity and diversity of data is produced on the interaction of individuals and groups, on a large variety of issues, has introduced a series of new interaction possibilities with formal models and inferential strategies that enable the development of substantial innovations. They offer an interesting discussion on how these innovations have contributed towards political science producing better experimental designs; making better comparisons between populations of interest; and creating more relevant observations on social and political behavior that were difficult to be ascertained by more traditional approaches.

The availability of data has been accompanied by an expansion in the offer of data analysis technologies at the individual or organizational levels. An important visible manifestation of this process is the ample availability of data analysis software which may be completely accessed by individual researchers, organizations, or teaching and research institutions, such as R, Stata, SPSS, SAS, NVIVO, QDAMAX Plus. These tools create a new frontier of possibilities in the production of more sophisticated analyses, whether in the quantitative research tradition, qualitative, or mixed-methods.

The amount of resources created by these packages contributes to a considerable increase in the dialogue between political science and fields such as mathematics, statistics, and econometrics. The qualification to adequately handle these tools has been an important addition in the training of new generations of political scientists. A careful examination of the main high-impact journals enables the observation of a proliferation of new techniques, methods, and data analysis possibilities that are mobilized by the current production which is immensely different from the what was done before the 1990s.

Teaching and training focused on methods using these packages has become an important "disciplinary" factor in the institutionalization of political science. Rare are the political scientists presently who do not spend a considerable amount of time in qualifying for the use of these tools. For example, currently, platforms such as R have been an important drive for political scientists to quickly get on board with a larger investment in programming activities to better handle the ample resources of this tool. The possibilities enabled by R for producing graphs, data analysis, and tasks such as online extraction of information, for example, has systematically attracted the community's attention in substantially investing in these activities.

Beyond the technological resources themselves, there is a new affordance of fast and reliable sharing of observational and experimental datasets among several research networks. This factor substantially reinforces the belief that valid knowledge in Political Science is more and more guided by the assumption of integrating theories and models to empirical data via hypothesis testing. This context opens the borders for dialogue with statistical analysis and econometrics to ensure valid inferences. The access to shared comparative data increases the possibilities for operationalizing concepts developed by theories.

Lastly, it is also clear that these elements allow, in a completely unprecedent way, access by researchers to the journals produced by the discipline. This feature ends up "increasing the degree of sharing" ideas, values, and

beliefs of what comes to be scientificity in political science and enables a growing institutionalization of such "parameters" that guide contemporary production. In the following section, I will focus on seven dimensions considered essential to understanding scientificity in the particular case of post-KKV political science.

3. DIMENSIONS OF THE METHODOLOGICAL TRANSFORMATIONS

This section will analyze "in isolation" each of the seven dimensions presented in the introduction as the main institutions that shape the new scientificity conceptualizations in the present study of politics.

3.1. Growing Concern with Causal Inference

Contemporary political science has as one of its main principles being guided by the production of valid causal inferences. Causal inferential analysis represents one of its fundamental purposes[9] and, therefore, this disciplinary field has sought analytical strategies capable of producing inferences through experimental and observational data. Gerring (2005), with a pragmatic approach on causation, argues that the social sciences must produce empirical knowledge that is reliable, objective, falsifiable, replicable, and inferential. The matter of causation becomes decisive[10].

The connection between scientificity and inference is directly associated with the pivotal argument by King, Keohane, and Verba (1994). As mentioned in the previous section, the authors consider the inferential issue as one of the four fundamental elements of scientificity in political science (and social science, more broadly). The authors propose a necessary connection between research designs and inferential quality, in which the creation of knowledge that adheres more to the "rules of inference" strengthens the quality of empirical research. This work transformed the scientific community's beliefs and behavior regarding the epistemic status of the inferential issue in political science.

In the recent production for the case of political science, the coexistence of multiple paths to reach causation empirically can be identified, going well beyond the estimation of parameters in linear models, as it is typical of applied econometrics. Observing random samples in the publications of the 20 Top Journals in Political Science, Rezende (2015) identifies that political science is strongly characterized by an "inferential pluralism", in which several logics and multiple avenues to generate valid inferences coexist. The basic assumption of inferential pluralism to generate valid causal inferences is that political scientists situated in different "disciplinary quadrants" must have at their disposal multiple patterns of causation. And, unlike in economics, none of these models has a legitimate monopoly on inferential validity.

Due to the strong interest in causal inference to confer scientificity, political science would then move towards two large categories: experimental and observational research, to access causal effects as well as to understand the cause of those effects[11]. As Gerber

9 Empirical data presented by Box-Steffensemeir, Brady, and Collier (2008) reveal that between 1995 and 1999 approximately 33% of articles published in the American Political Science Review (ASPR) and 19% of all journals published by JSTOR, use the expression "causal inference". They note that when the expression. "cause" is considered, this percentage grows to 67% for the ASPR and to 60% for all JSTOR journals. These data tend to grow if we consider the period following the one analyzed by the authors.

10 Quality empirical research has two basic characteristics: to generate causal inferences from collected data and to have this process follow the "rules of inference" (King and Epstein, 2011. p.14).

11 As Przeworski (2009) states, a considerable amount of problems treated by comparative politics involves understanding the impact of variable X on variable, in which usually dependent variables are institutions, policies, or certain events, conditions where experiments are not possible.

(2004) states, the main concrete difference between these two types resides in the possibility of randomization for the data generation.

There are four fundamental logics to understand causation in social science, ones which directly apply to political science: covariational, experimental, and counterfactual models, and causal mechanisms. He argues that there are higher quality inferences and these have four characteristics for which he states some basic principles: (a) "no covariation, no causation" – covariation between causes and effects; (b) "no counterfactuals, no causation" – in which there is no given causal effect when the cause does not operate in a counterfactual reality; (c) "no manipulation, no causation" – in which causal effects may be generated from experimental conditions that are effectively manipulated by the researcher; and d) "no mechanism, no causation" – in which the connection between causes and effects demand the presence of identifiable causal mechanisms (Brady, 2008. p.218).

The basis for these causation models is the logic of potential outcomes, which is known as the Neyman-Rubin-Holland model, originally conceived for experimental designs. Later, as Angrist and Prischke (2009) show, this was expanded for observational research.

The potential outcomes model seeks to "find a solution" to the central problem of causal inference, categorically stated as: it is impossible to simultaneously observe a same unit of analysis being exposed to treatment and control conditions. To generate causal inferences and access effects, counterfactual scenarios (Morgan e Winship, 2007) must be built, that effectively allow comparing the cases. When randomization is possible, counterfactual groups are "naturally" created. The problem becomes more acute when one is handling observational data.

When randomization is not possible, as with observational studies, other techniques such as matching, propensity score analysis, and others are used. The key problem in observational research is how to control "confounders" that tend to substantially affect inferential results when not adequately controlled. Imai et al. (2011) argue that when we are dealing with experimental studies, the mediation effect may be found through the estimator of the difference between means. The authors believe that the main advantage of the potential outcomes model is related to the attention it can give to the problem of causal heterogeneity, usually inadequately grasped by linear regression models, given that it is often "placed" in the error term. Premises of exogeneity are indeed insufficient to understand and identify the causal mechanisms involved in a causal proposition.

Sekhon (2009) believes there are three basic paths to making causal inferences in the political science quantitative tradition: experimental, inferences focused on models, and ones guided by research design. The first goes back to the experimental tradition that was the first formative value in North American political science since Merriam[12], in the beginning of the 20th century. The most frequent path is the one in which inferences are created from models that involve causal propositions, guided by theory, and tested with multivariate regressions. The more modern tradition focuses on research design. Nonetheless, research designs almost always seek to build alternatives to experiments to overcome imperfect experimental conditions for generation of causal inferences, as I will discuss further on.

Presently, without exaggeration, it can be argued that without an adequate research design to access causal inferences, it is unlikely that econometric or statistical

12 To understand the role of the experimental tradition in North American Political Science in its formative years, see Merriam's (1921) pivotal paper.

analysis can create knowledge that goes beyond covariation relationships between variables. The presence of artificial control conditions, manipulability, and crucial attention to research designs (and the conditions under which they must be applied), shows itself to be vital in producing valid causal inferences. This belief has been altering political science considerably, towards building research designs – experimental, quasi-experimental, or comparative – that truly produce more inferential knowledge. This feature has increased the degree of rigor and methodological sophistication in recent production.

One of the characteristic effects of this logic resides in the growing relevance of the "inferential problem" in the empirical political science of several countries. As econometrics, political science seeks to attain the causal-inferential model for production of knowledge as an ideal. Possibly, political science would be more systematically nearing the "ideal model of empirical research" proposed by Angrist and Pischke (2009) in econometric analysis. For these authors, modern empirical science must, therefore, be supported by empirically treatable causal questions and must, necessarily, contain in its analyses the following foundational elements: (a) a causal proposition; (b) "an ideal experiment"; (c) identification strategies; and, lastly; (d) causal inference.

The growing importance of inferences has become a reality in the construction of research designs that come considerably close to the "experimental ideal". Indeed, in political science, as I argue later on, this concern is translated in the proliferation of innovations of more rigorous research designs. In political science, a broad ecology of possibilities coexists, involving experiments, quasi-experiments, comparative studies, and case studies (e.g., Gerring, 2005; Rohlfing, 2012) to generate valid causal inferences. This distinctive configuration enables the idea of inferential pluralism.

3.2. Orientation by Hypothetico-Deductive Models

The production in Contemporary Political Science is strongly organized by compliance with hypothetico-deductive models (H-D). This belief represents a clear alignment with the creation of falsifiable theories, orientation by formal models, and essentially, the intensive use of hypothesis testing. The founding concept is that theories developed must allow for empirical confirmation. This value attributed to H-D models brings about the need to make political science closer to economy, reaching beyond being a merely discursive science.

H-D models are constructed from simple logic. Theories and formal models are built in ways that allow empirical falsification; from these, implications in the form of hypotheses are made, ones which are tested with empirical data. The corroboration of theories using statistical methods and data analysis techniques represents the fundamental element. Therefore, the Popperian principle of scientificity attached to falsifiability still finds substantial room.

This belief regarding hypothetical models became institutionalized in Political Science in the 2000s, with the formulation of the Empirical Implications of Theoretical Models – EITM, a program developed by the Political Science Program of the National Science Foundation. This program has had substantial impact in guiding scientific research since then, with strong influence over the standards of disciplinary scientificity of many countries. This project's impact on disciplinary production, professional formation, and reorganization of the concept of scientificity is of utmost importance to understand the nature of production in high-impact research.

Based on the ideas of the Cowles Commission[13] in the 30s, the main goal of the EITM project is to "reduce the deficit" between formal models and empirical analysis which were pervasive in political science, as well pointed out by Morton (1999). The adequate combination between these two logics allows the introduction of comparative advantages to generate valid scientific knowledge. On the one hand, the project believes that the intensive use of formal models enables the contemplation of deductive modelling strategies from theorems and demonstrations of proofs or computational simulations. On the other, the stronger inclination towards empirical analysis for hypothesis testing demands the intensive use of statistical tools for data analysis.

The use of formal models supposedly introduces a larger demand for clarification regarding key premises and concepts mobilized; introduces a stronger logical consistency which increases disciplinary rigor; and allows, at the same time, more space for identification strategies that can truly understand how certain mechanisms produced certain results in the reality analyzed. For its part, statistical analysis enables a higher capacity for generalization of analysis and aids in taking more robust decisions in the elimination of rival hypotheses based on a multivariate analysis. In addition, statistical analysis allows the identification of causes and effects, treatment of reverse causation, and estimation of the causal effects in a given model.

As Granato et alli (2004) state, the EITM is run on the institutionalization of three crucial elements. The first is the creation of an incentive strategy so that Political Science can base itself on the notion of cumulative production of knowledge. The second element, strongly supported in the presuppositions of the Cowles Commission, seeks to promote scientificity from a larger interaction with disciplines such as economics, statistics, and mathematics, to foster the necessary connection between formal models and statistical analysis. The project's third fundamental objective is to go beyond the proposals of the Cowles Commission and create clear strategies for making causal inferences with emphasis in identification.

Granato and Scioli (2004) show that, within the context of EITM, research in political science is guided by five founding values: (1) orientation by theory, which must be created from case studies, field research, or "research problems"; (2) formal models that can identify the relevant causal connections; (3) deductive analytical strategies from falsifiable/testable hypotheses; (4) considerable measuring effort and attention to research designs; collection, systematization, and data analysis strategies.

The fundamental belief that sustains the project is the notion of connection between theory and data. This feature allows for more explanatory power, more consistent empirical models, and development of relevant deductions in knowledge. The connection between formal models and data enables scientists to construct theories that can indeed be tested, i.e., falsifiable, with data, clearly rejecting the construction of theories that cannot be operationalized or converted into testable hypotheses, as it is still the case in other areas of the social sciences. The test requirement allows the knowledge produced to be more capable of generating valid inferences.

Consequently, the EITM models are not restricted to using quantitative methods for tests. Qualitative analyses may be considered fundamental in validating concepts generated by theory, or even in understanding relevant causal mechanisms. Quantitative and qualitative methods should be considered complementary and supporting of each other. Mixed-methods research should be considered preferable.

13 The Cowles Commission was created in the 1930s by six economists to promote and foster the development of formal methods – mathematical and statistical – of analysis for application in economics and social sciences.

The impact of this important project is decisive in consolidating a new conception of scientificity through continued formation. Aldrich, Alt, and Lupia (2008) reveal that soon after the launch of the program in 2001 by the NSF, the dissemination of summer courses was promoted in important universities such as Berkeley, Duke, Harvard, Michigan, UCLA, and Washington University, involving hundreds of participants. These training sessions were expanded to a large set of Universities around the world. They strongly introduce the notion that research designs centered on formal models and empirical analysis can indeed give credibility to the research.

3.3. Research Designs

As mentioned in the previous section, the "disciplinary effect" created by KKV expanded the methodological reflection surrounding the inferential possibilities and its connections with research designs. The connection between research design and inferential quality represents one of the most important values that affect the high-end production in political science.

Political science has been progressively guiding itself on the classical recommendation proposed by Rubin that "design trumps analysis" and has substantially altered its scientificity, coming remarkably close to initiatives created by applied econometrics. Indeed, political science leaning more and more towards hypothetico-deductive models and causal inference-making, must attempt to understand how research design – experimental or observational – may in fact generate valid inferences. Broadly speaking, research guided by inferences is currently between two logics: experimental and observational.

Morton and Williams (2010) argue that the growing demand for experiments – in its diverse variants – in political science is derived from two basic reasons. First, the continued "failure" by traditional quantitative methods in offering more satisfactory responses to the demand for causal inference. Green and Gerber (2009) believe to be also important the "exhaustion" of observational research designs, leading to a higher demand for experimental models.

One of the most intriguing issues is related to the criteria that motivate the decision by researchers to choose experiments. One of the most satisfactory answers is formulated by Gerber (2004) in the Illusion of Observational Learning Theorem. They suggest that, even with large samples, grave inferential errors may compromise research designs when potential biases threaten the efficiency and consistency of estimators. Facing an elevated bias variance, the data (and the results) generated by observational research have limits to indeed advance inferential knowledge. To overcome such problems, the experiments are considered crucial to expand causal knowledge.

One of political science's most distinctive features in the last decade is the return to "experimental reason", that marked the birth of the scientific study of politics in the 1930s, in the United States. The return to experimentalism comes about with a serious consideration of the use of experimental research design when the conditions are favorable. In its many variants – field experiments, natural experiments, laboratory experiments, survey experiments –, experimental methodologies have acquired consistency, sophistication, and a set of "comparative advantages" to produce causal inferences in contrast with observational studies.

Broadly speaking, the more intensive use of experimental designs in political science emerges as one of the most important methodological trends, and is an important response to four large problems: a) an expansion of the classical models of estimation, through linear regres-

sion; b) an increasing methodological sophistication to control biases and reduce effects of other causal factors; c) an expansion of inferential quality, increasing rigor regarding problems of internal and external validity; and lastly, d) allowing more robust forms of integrating quantitative and qualitative methods.

The superiority of experimental methods in generating causal inferences has produced strong incentives for the concern with the introduction of randomization and manipulation strategies to be taken seriously, to overcome the issue of confounders in the models. This problem represents a serious limitation of analyses that count on observational data alone.

When randomization is imperfect, and data is generated by observational research, quasi-experimental designs become a preferable option in the creation of research designs. Campbell, Cook, and Shadish (2002) suggest that when facing "imperfect" controls and randomization, research designs must be considered as quasi-experimental, a typical situation in which observational studies enable generation of inferences from controls "similar" to ideal experimental conditions. Collier et alli (2004) argue that these designs face obstacles and threats like experiments regarding the generation of causal inferences, which ensures that they are considered "as if they were", to some degree, experiments.

Imperfect experimental possibilities have been a favored field in the rise of considerable innovations in research designs that mobilize observational studies. The problem of how to generate causal inferences when facing imperfect experimental conditions, with strong endogeneity problems (Przeworski, 2007), and the continued problem of omitted variables, as well as other threats to internal and external validity, makes quasi-experiments one of the privileged innovation fields in terms of design for causal effects estimation. For this reason, considerable part of the production mobilizes research designs in the form of regression discontinuity, instrumental variables, difference-in-differences, hierarchical models, and synthetic controls (Abadie et al., 2012).

This field has also been responsible for producing critical innovations in large areas of empirical political science. Therefore, political science would be aligning itself more and more with Angrist and Prischke's (2009) proposition that one of the main decisions in making a research design is associated with the important question of finding "an ideal experiment", so that a specific causal proposition guided by theory can be studied.

The critical aspect of observational research design, typical in social sciences and, in particular, political science, lies in the fact that, even when these conditions are present, there is a strong possibility that confounders or "other non-observable factors" in the difference among control and treatment groups can affect the results. Attention to this problem makes other quasi-experimental research designs be complemented by other ways of thinking and manipulating causation, as with counterfactuals, mechanisms, and inferential strategies with few cases. The key challenge lies in finding alternatives to combine research designs to reduce threats to internal and external validity.

Beyond the quasi-experimental research designs mentioned, it is observed in the particular case of political science, the introduction of three important types of research design that are having more presence in analysis: case studies, process-tracing approaches, and Qualitative Comparative Analysis (QCA). These forms are considered valid to generate inferences, in political science.

Case studies (Gerring, 2004, 2005; Rohlfing, 2012; Rezende, 2011) have been one of the privileged fields in political science, both in the experimental perspec-

tive (usually strong in internal validity) and traditional small-n research. These are considered alternatives to "controlled experiments" when one wants to understand causation with special attention to conditions and causal mechanisms, or deal with problems related with endogeneity.

Processes of formalization and increase in rigor of comparative methodologies continue to point to case studies as legitimate and potentially heuristic sources to create theoretical knowledge as well as causation. Cases do not just produce descriptive knowledge, but also advance it, in a world where phenomena are loaded with endogeneity, causal complexity, and dependency on specific conditions. Cases may be relevant tools to understand the diversity of causal patterns, direction of causalities and, more importantly, to observe the degree of non-spuriousness in each research design. The connection between cases and theories is ever more frequent and recommended as valid analytical strategies.

Approaches such as process-tracing (Beach and Pedersen, 2013) represent valid alternatives when researchers' attentions are focused on diving in the connection between agents, institutions, and contexts to produce satisfactory causal explanations from mechanisms. Attention to causal processes has given substantial relevance to producing inferences centered on causal mechanisms. George and Bennet (2005) argue that process-tracing strategies seek to identify intervening causal processes that connect independent variables to a given dependent variable. These models are ideal when one wants to truly dive into the "large chain" of causal mechanisms that are operating in a given reality, which is usually done through cases studies.

Lastly, political science has also observed the introduction of the so-called set-theoretical methods (Weller and Barnes, 2014). Based on a Boolean logic originally developed by Ragin (1989) to think about causation, these methods mobilize several types of strategies and techniques to formulate concepts, create typologies, and produce causal interpretation[14], such as fuzzy sets QCA[15], crisp QCA, Multi-Value QCA, and Temporal QCA.

Configurational methods introduce a new logic to generate causal inferences from small-n research designs, in which it is possible to understand how conditions – necessary, sufficient, or a combination of these – may create causal analyses from so-called "truth tables". These research designs are especially used when researchers face causal complexity patterns marked by equifinality (multiple causation), joint causation, and causal asymmetry.

3.4. Integration of Quantative and Qualitative Methods

One of the key values for political science goes through the effective possibilities of connection between quantitative and qualitative strategies of analysis to produce causal inferences. Political science has progressively sought to introduce the connection between methods since the original argument proposed by King, Keohane, and Verba (1994) on methodological unification.

The authors argue that there is a common inferential logic between the two research universes. The argument lies in the fact that there are no substantial differences between quantitative and qualitative methods, which must be guided, oriented, and structured by a similar logic for the generation of causal inferences. This basic belief has profoundly altered the scientific community's perceptions on the issue of integrating these two universes.

14 APSA (2014) gives a perspective on the limits of configurational methods for political science.

15 For an updated critique on the limits of fuzzy methods, see Krogslund et al. (2015).

The impact of this argument was crucial. It brought about a significant reconfiguration of the methodological debates within the discipline, especially regarding the use of qualitative methods for producing causal inference and the possible ways of integration what historically coexisted with strong divisions when it came to its logics and purposes. The unifying logic has created a gradual process of eroding traditional beliefs, with the identification of a new moment of profound institutional transformation in the discipline, in which mixed-methods research goes from desirable to ideal.

One of the important impacts developed by this work is the creation of methodological studies, which open new possibilities and frontiers in research design, offering fertile answers to the "problem of methodological unification" in producing causal inferences in qualitative research. Even authors who do not directly agree with this position (Brady, Collier and Seawright, 2004, Mahoney, 2010, Mahoney and Goertz, 2006; 2012) have strongly contributed to bring in "qualitative methods" for reformulation under the parameters of inferential quality.

The problem of integrating quantitative and qualitative research designs represents, therefore, one of the main areas of concern in the recent political science debates and an expressive number of works have been creating opportunities for a larger methodological awareness of this issue. More than in any other field in the social sciences, there is in political science still a strong commitment with causality, causation, and causal inference and reflecting on the bases of the viability of integration takes on a clear importance in contemporary debates. The production of more robust theories, explanations, models, and tests depends on the possibility of promoting more fruitful discussions on how to integrate the quantitative and qualitative paradigms.

The central issue in these reflections lies in the specific conditions that allow thinking of causal inference (and explanation) from a more "integrated" perspective, as ideal. Far from thinking that these discussions immiscible or even incompatible, they are guided by more clearly elucidating the frontiers of integration and differentiation between methods (Rezende, 2011). The fundamental question is understanding the conditions that make possible integration or differentiation between approaches in specific research designs. As the pivotal study by Ahmed and Sil (2012) states, the problem of how to integrate methods in the "objective" reality of practice research takes center stage.

Why is the integration of methods desirable as a scientificity ideal? Fearon and Laitin (2008) argue that the belief regarding integrated research designs has become highly popular in political science and in large part this is due to the recognition of the limits of the validity of traditional quantitative research to produce inferential knowledge. The authors believe that integrated strategies, when well-done, increase the effective chances of combining the power of quantitative analysis with case studies and small-n to increase the chances of better inferences.

On the one hand, the quantitative dimension would be more fruitful to enable the identification of causal patterns and empirical regularities, supported by statistical analysis and, on the other, the qualitative dimension – usually in the shape of case studies and small-n – would be focused on creating knowledge about causal mechanisms that generate the phenomena of interest in certain contexts. The complementarity between the methods would therefore be a desirable "balance".

One of the central arguments for the issue of integration via complementarity is proposed by Tarrow (1995). The complementarity thesis between the methods proposed by him argues that the quality of the integration

depends on how researchers understand the logic, role and the division of labor in these designs. Qualitative researched focusing on identifying the non-systematic dimension of reality and its variation logic among contexts. Conversely, the systematic dimension of reality, usually expressed by causal regularities, must be apprehended by quantitative methods.

The second advantage for integration lies in the argument that it generates better quality of inference and greater error reduction, thus increasing analytical efficiency and explanatory power of comparative studies. The belief regarding the viability of integrated conceptions is organized around the assumption that "shared scientific principles" among quantitative and qualitative traditions such as falsifiability, parsimony, reduction, logical coherence, and validity must be present in research designs to ensure more chances of promoting greater inferential quality.

The growing institutionalization of this belief as one of the most beneficial to operationalize the diffusion of benefits from methodological pluralism (and analytical eclecticism) has been essential in contemporary political science, making room for similar conceptualizations. Clear examples of this integration have substantially influence the standards of contemporary political science's scientificity.

One of them is in the clear tripartite integrative conception to "discipline political science as science", as suggested by Laitin (2002). He is supported by a conception of scientificity that involves the combined use of formal models, statistical analysis, and qualitative analysis focused on increasing inferential quality. The notion of integration is also clear in the formulation "nested analysis" proposed by Lieberman (2005), where the simultaneous use of large and small-n strategies to identify cases, to generate historically situated explanations, as well as hypotheses that can be tested in a larger population, appears as a key factor in the quality of causal inferences.

Known the comparative advantages regarding integration, it is important to highlight that the confection of integrative strategies in comparative research has sensitive limitations.

Although desirable, integration is not always doable, and we need to understand regarding what issues (or topics) these limitations are based on. Understanding these limits, therefore, enables the researcher to understand with greater depth the key issues to create integrated research designs. There are tensions among the qualitative and quantitative approaches, which are decisive in structuring these limits. The first falls on the possibility of integrating ontologies and methodologies; the second lies in the complex matter of multiple conceptions about causal inference, causation, or even causality; the third focuses on the different perceptions of what are external and internal validity; and, lastly, the issue of causal complexity.

Hall (2003) believes that there is a growing disconnect between the "ontologies" of statistical analysis models and traditional comparative methods (qualitative and small-n research), becoming a fundamental dilemma for political science and its inferential possibilities. He argues that comparative models have been trying to incorporate crucial analytical elements such as joint causation, causal complexity, agency, and equifinality, which are incompatible with the fundamental assumptions of traditional statistical analysis.

Therefore, as Ahmed and Sil (2012) argue, the issue of integration is associated with the effective possibilities of ontological alignment and the conceptions of causality of two or more methods. When ontological conditions do not allow integration due to incompatible conceptions of causality (or causation), integrated

research design may, unlike what is supposed, generate errors and inferential problems. This fact is incredibly important given that not always the analytical efficiency of integration would be higher in relation to non-integrated designs. The most important dimension connected to choosing methods lies not in the research question, as some authors mechanically assume, but in the ontological conditions that give meaning to a potential integration of methods. This point aligns with the essential demarcation question proposed in the article. Method does not prevail over ontology, but the other way around.

Another relevant point lies in the perceptions that quantitative and qualitative researchers have on internal and external validity. Research designs usually involve two fundamental questions regarding the "problem of inference": the issue of inferential quality (issues relative to internal validity or credibility); and the issue of the capacity for inferential transfer (external validity or transmissibility). "Possible agreements" are not always present between these dimensions and the research must be aware of when and how these issues must be integrated.

Mahoney (2008) suggests an important integrative look when facing this dilemma, by proposing that the validity of a unified theory of causality which enables that the "language" of variables and causal mechanisms be common to quantitative and qualitative universes. He argues that case studies and multivariate regression analyses have different logics for different purposes: on the one hand, case studies seek to understand why certain results occurred in specific conditions; on the other, regression analyses are made to estimate the mean effect of causes. The analytical power of case studies lies in their ability to identify and analyze causal processes, and in these cases, one can almost always observe intense endogeneity and the presence of causal effects with complex interdependence among the variables.

Ideally, any attempt at an integrated explanation must mobilize elements that inform on the integration of concepts. For quantitative researchers, the ideal position would the one in which concepts are "perfectly" operationalized in measurable variables and hypothesis testing for statistical analysis. This "reduced" version of reality assumes possible that concepts be simple, clear, and objective, allowing precision to be achieved. Qualitative researchers, conversely, usually believe that concepts (and research questions) mobilized by political science are complex, multidimensional, and polysemic, making complex, when not impossible, an adequate generalization for empirical analysis in ideal terms.

3.5. Analytical Eclecticism

Analytical eclecticism represents another important shared value for scientificity in contemporary political science. Originally proposed by Sil and Katzenstein (2012), analytical eclecticism is a plausible alternative to deal with the "flight from reality" problem, raised by Shapiro (2005). This is characterized by the tension between theoretical debates and the demands for practical knowledge in the empirical world, as mentioned previously.

Political science is constantly pressured to offer relevant causal explanations in the face of a multidimensional, dynamic, and complex reality, and well-known for a high degree of endogeneity. These factors, on the one hand, make political explanation demand models and theories with high complexity, since they usually involve causes and causal mechanisms that are situated within institutions, culture, and history, and still demand the presence of rational agents. On the other hand, because it is a science that deals more closely with real life problems that directly affect the lives of large groups, there is a considerable demand on the part of

several knowledge producing bodies to generate useful knowledge that can indeed contribute with intervention via public policies.

To handle all these problems, the strategy suggested by analytical eclecticism is about overcoming traditional research conceptions that tend to be organized by isolated research paradigms or traditions, and starts to combine inputs from several models and theories generated from within the discipline. This belief has been substantially altering the way scientific knowledge is made in political science, especially in its more applied fields, such as international relations or public policy.

The fundamental assumption of analytical eclecticism lies in the belief that, in a context of abundance of theory, models, or paradigms competing for better scientific explanations for problems of high complexity in the real world, it is important to introduce "flexibilization" efforts among research paradigms when producing better causal inferences and relevant knowledge in political science. This perspective requires new ways of thinking about premises, fundamental concepts, theories, disciplinary organization and, ultimately, the connection with actual relevant problems.

The main implication of this belief is that scientific research in political science bases itself on the assumption that no theoretical paradigm solely possesses the legitimate monopoly of analytical superiority. Thus, it is necessary to seriously consider the connection or reduction of barriers between concepts, evidences, and causal analysis created with paradigms traditionally considered "incompatible". This assumption, going well beyond the simple matter of "integration", involves the effort of exploring hidden connections between theories and models in order to produce new approaches and innovations for analysis of relevant phenomena applied to the real world.

The main comparative advantage of analytical eclecticism lies in the increase of possibilities to create conceptual and theoretical innovations and as well as causal explanations produced. Rejecting the position that explanations must be confined to specific boundaries within a given paradigm (which the authors call paradigm-bounded research), eclecticism enables the development of more powerful analyses, creatively and consistently meshing together contributions from several research traditions that are usually mobilized "in isolation".

The eclectic position requires political scientists to progressively abandon their "narratives and preferred models" and start to pursue more satisfactory answers, broadening the scope of relevant causal factors and causal mechanisms involved in the construction of explanations of a particular phenomenon. The broadening of paradigmatic frontiers enables the creatively handling of new analytical possibilities.

Analytical eclecticism must not be mistaken with mixed-methods research in political science. To more adequately understand complex phenomena that involve several dimensions of analysis and require a great dose of contextuality, explanations in Political Science must be more attentive to a gamut of contributions made by several paradigms. Thus, a larger awareness regarding the limits, potentialities, and implications regarding the methodological choices when facing several approaches.

Analytical eclecticism is differentiated from traditional research led by paradigms on three dimensions related to: how research problems are built; the causal strategies for analysis; and the degree of adherence of theories and models to the context. The first dimension is related to the complexity and multidimensionality of the relevant phenomena for political analysis. Research problems often are built by theories that are adjusted

to paradigms. Eclecticism demands the transposition of paradigmatic barriers, as in, enabling the explanation of complex problems. Lastly, eclecticism assumes as necessary to construct theories and analysis that can effectively contribute towards innovations to treat relevant empirical problems.

Analytical eclecticism's fundamental concern is the pressing need to develop greater interaction between theoretical knowledge, analytical models, and the empirical evidence. Useful knowledge must be created from middle range theories that can unravel relevant causal mechanisms in political phenomena. The innovative proposal brought about by analytical eclecticism is about the need to explore new avenues between research traditions, enabling a more intense communication of concepts, mechanisms, theories, and explanations in the face of specific research problems. These assumptions tend to substantially increase the probability of solving the chronic gap between theories and evidence in political science.

3.6. Growing Demand for Identification

As analyzed previously, Contemporary Political Science has focused on generating valid causal inferences. To do so, it uses of empirical research in research designs that formulate several strategies of identification that enable the understanding of how causes lead to the effects in the empirical reality observed. Scientificity then depends on how the problem of identification in the social sciences can be solved (Manski, 1991).

The identification problem lies in establishing if the causal propositions assumed by the theoretical models effectively occur (and are confirmed) in empirical reality, i.e., if the causes indeed are causes and the effects are actually generated by the causes supposed, given observational data. How can we suppose exogeneity of causes when using observational data? That is an incredibly complex problem and substantial analytical energy has been devoted to it in applied econometrics, with considerable impacts on comparative politics.

With experimental designs, this problem is "easily" solved due to randomization and experimental manipulation of causes and effects. However, with observational studies, given the problem of omitted variables (confounders), the demand for empirical strategies of identification make inference issue more complex.

Angrist and Prischke (2009) believe that empirical research seeks to answers questions of causal inference from "ideal experiments", in which researchers aim to better identify to solve problems of endogeneity, selection bias, reverse causation, and causal mechanisms involved in the explanation. It is precisely through strategies of identification that such inferences from observational data are operationalized, and researchers truly seek plausible empirical alternatives to reach the desired causal effects.

The connection between good identification and inferential quality has been decisive in observational research in economics, instigating the credibility revolution, which has had similar impacts on political science. In political science, there has been an increase in demand to handle the identification that accompanies the rigor in the creation of research designs.

Considered as the "critical point" in observational research, identification performs a key role in the quality of research and inferential knowledge produced. With adequate identification strategies, researchers can empirically deal with the usual inferential problems associated with endogeneity generated by non-observable factors or causal complexity caused

by institutions, history, or more broadly, culture, as it is typical in political science.

Attention to possibilities of identification have allowed observational research to produce valid causal inferences. Scientificity depends, in large part, in making theories consistent with observational empirical data. Considering a comparative politics perspective, Przeworski (2009) states that when political scientists cannot control causes, we are completely dependent on history and thus, the chances of good identification are really reduced.

In historical analysis and political science, where the degree of endogeneity is extremely high, this problem becomes acute. For him, comparative political science lives in a paradox: "The better we specify our models, the more endogenous loops we consider, the more difficult it becomes to identify their causal structure." (Przeworski, 2007; p.168). The main challenge when facing endogeneity is in the fact that it is difficult to empirically distinguish the causal effects from the effects generated by the particular conditions in which the studied phenomenon is occurring.

Understanding the problems regarding identification is also discussing the limits regarding internal and external validity in observational research. The demand for identification is one the most important methodological trends and it is an important answer to three problems highlighted by Dunning (2012): a) overcoming typical problems in multivariate regression analyses; b) needing more robust research designs, with the inclusion of field and natural experiments; c) an effective possibility to increase inferential quality through constructing good identification strategies in observational research.

This hypothesis is also supported by the relevant argument proposed by Brady, Collier and Seawright (2006) in Rethinking Social Inquiry, in which the growing demand for a creative dialogue between quantitative and qualitative strategies is needed to enhance the capacity of causal inference in theory production (and hypothesis testing). Thus, experiments arise as a plausible alternative that allows (via other means) the combination of quantitative data analysis and analysis of causal processes.

The demand for identification from quasi-experimental designs may be understood as requirements considered fundamental so that knowledge on the several logics of causation in political science can be generated.

Thus, many authors working on the frontier of empirical research assume that identification strategies are the fundamental dimension in constructing research designs. These strategies are broadly known as the alternatives research designs build to estimate causal effects for specific empirical problems. The central issue lies in understanding how contemporary studies have been mobilizing innovations in identification strategies in research designs centered around quasi-experiments and how these contribute to inferential quality.

Consequently, refining identification strategies would have an important role in increasing "experimental reason" in political science, where the combination of experimental, quantitative, and statistical analysis elements can lead the production of valid causal inferences. Thinking about the identification problem from quasi-experiments means creating the conditions for a more robust inferential causal knowledge, which is needed for the advancement of political science as a discipline.

3.7. Primacy of Methods over Analysis

The last influential belief I consider here has to do with the primacy of methods. This can be clearly identified

in the transformations occurred in recent production, with the advancement and sophistication of methods and techniques applied, as well as institutional efforts to introduce the disciplinary field of Political Methodology. These characteristics have made political science consistent in the assumption that methods represent the crucial elements that create the conditions for diffusion and institutionalization of values regarding a new scientificity. In fact, it is possible to state that political science approaches a "Sartorian ideal" of increased awareness about methods.

The influx of new methodologies in political analysis can be attributed to, on the one hand, the increased interaction with statistical analysis, especially multivariate analysis techniques for hypothesis testing. On the other, a simultaneous expansion of the reflexivity regarding comparative methods is observed, which has substantially incited the real confrontation between theories and empirical data. The expansion of rigor and sophistication of comparative methodologies is clearly associated with an exponential increase of quantitative and qualitative techniques available to political scientists.

Certainly, the orientation by hypothetico-deductive models centered on hypothesis testing via the diffusion of new values introduced by the EITM/NSF program since 2001[16], established the conditions for increasing the degree of rigor and sophistication, both in quantitative and qualitative research. In the quantitative tradition, the intensive use of multivariate techniques in its several variants, or even more sophisticated techniques such as hierarchical models, temporal series, or computational simulations, is clearly observed. The degree of integration with econometrics made so that models and testable theories hit an extremely high degree of formalization and analytical rigor in the current production.

There is today ample discussion in the qualitative research tradition in political science on the limits and potentials of these methods, as well as a large offering of possibilities for inference generation, especially when configurational models or ones based on causal mechanisms are used. Approaches such as process-tracing, case studies, or even the several QCA variants, have introduced a substantial influx of possibilities for developing new methodologies to generate valid causal inferences.

Another important symptom of the expansion of methods as a crucial element of scientificity in the context of the EITM project is the growth in the status of journals that publish papers mobilizing formal models, theory, and empirical data from research designs using a high level of methodological rigor. Data in Figure 1 compare the positions of the top five journals in SCImago Journal & Country Rank (www.scimagojr.com).

16 The diffusion of values regarding Methods in Political Science have been well supported in professional formation by the Inter-University Consortium of Political and Social Research (ICPSR) Summer Courses, created in 1963, at the University of Michigan and the Essex Summer School in Social Science Data

Analysis and Collection, founded in 1967. New initiatives, such as the creation of the Society for Political Methodology da American Political Association, created new opportunities for the massive diffusion of values related to the importance of method and political methodology. In Brazil, the International Political Science Association's (IPSA) summer schools are annually offered by the University of São Paulo USP.

Figure 1 – Political Science Journals

Ranking SCImago Journal & Country Rank 2000; 2005; 2010

ANO 2000	ANO 2005	ANO 2014
World Politics	American Political Science Review	American Political Science Review
American Journal of Political Science	American Journal of Political Science	Political Analysis
American Political Science Review	Comparative Political Studies	American Journal of Political Science
Journal of Conflict Resolution	World Politics	Journal of Conflict Resolution
Comparative Political Studies	Jounal of Theoretical Politics	Annual Review of Political Science
Journal of Politics	Journal of Conflict Resolution	Journal of Politics

Source: the author, with data from www.scimagojr.com.

The data show that journals such as the American Political Science Review, American Journal of Political Science, Journal of Politics, and the Journal of Conflict Resolution continue to maintain their high impact over the years when it comes to diffusion of methods and techniques mobilized by the discipline's mainstream. The novelty emerging from the data and that reflects the recent primacy of methods is the appearance of the journal Political Analysis, as the second main journal in 2014.

This journal has strengthened since its creation in 2006 and it has been the main source of publication for high quality papers that mobilize rigorous and current methodologies in political science. This journal has always promoted innovations in political analysis that massively includes formal analysis and statistical analysis. On the other hand, an analysis of the content of the papers also reveals its relevance as an important diffusion source for sophisticated techniques to integrate quantitative and qualitative methods.

The increased relevance of the Political Analysis journal for scientificity in political science can be "objectively" observed in data from Figure 2. These show that there has been an exponential increase in its impact factor, going from 0,917 to 4,655. This illustrates its power of diffusion of new forms of analysis in the discipline. In addition, it goes from 20th in the ranking (out of 85 journals) in the year it was created, to 1st in 2014.

Figure 2 – Political Analysis: Impact Factor and Ranking (Political Science Journals) 2006-2014

Year Impact Factor Ranking (Political Science Journals)

Ano	Fator de Impacto	Posição Ranking (Periódicos Ciência Política)
2006	0,917	20 (85)
2007	2,535	1 (93)
2008	4,78	1 (99)
2009	3,756	1 (112)
2010	1,864	9 (139)
2011	2,191	5 (148)
2012	2,231	5 (157)
2013	2,879	3 (156)
2014	4,655	1 (161)

Source: http://www.oxfordjournals.org/our_journals/polana/about.html

The importance of journals in the formation and "disciplining" of production is decisive. As it is known, political science follows the trend of other disciplines and is heavily impacted by reading "publications in scientific journals". Undoubtedly that this impact, especially from main journals, ends up creating an isomorphic process of values around what are methods and finally shape the parameters of scientificity in the context of a cosmopolitan political science.

In addition, the primacy of methods can also be observed by the massive production in books on applied methodology specifically for political science. Publishers such as Cambridge University Press, Oxford University Press, and other high-caliber companies have contributed immensely to the diffusion of high-quality works in the field of methodology, empirical research, quantitative and/or qualitative, that significantly expand the discussion and awareness regarding methods in political science[17]. Specific series focused on research were created, enabling the rapid dissemination of beliefs around the demand for methodological renewal in the discipline. It is important to highlight that Oxford University Press published the Handbook on Political Methodology in 2007, making available a collection of works produced by the most qualified authorities on the subject with broad impact on the professional formation in political science.

4. CONCLUSIONS

This paper sought to present a set of key transformations that guided high impact factor academic production in political science over the past 20 years. The changes carried out were "beliefs and values" that quickly became institutionalized in the scientific community and today set an ideal for producing any scientific work in the field. It is possible to state that

17 For obvious reasons of economizing space, examples of these pivotal works are not included in the text. In the references, however, works considered fundamental for the diffusion of the "primacy of methods" are mentioned.

the nature, quality, and sophistication of contemporary production is radically distinct from what was thought of as "mainstream" before the revolution instigated within the discipline in the 1990s.

As discussed, these beliefs and values gain meaning and are diffused by structural transformations generated by endogenous and exogenous pressures within political science. These factors make the "scientificity indicators" completely specific in relation to other social sciences such as sociology, anthropology, law, or even in the humanities. Understanding these values more closely, we can comprehend what "is characteristic in political science". This comprehension cannot be considered obvious or trivial. The meaning of disciplinary scientificity that has been built in contemporary political science makes it an "outlier" in the social sciences that are still focused on methodologies centered around interpretation.

Currently, political science has been guiding itself with values that privilege empirical analysis centered around explanations, the intensive use of research designs, and the growing demand for causation. Rezende (2015) believes that these dimensions are the main constitutive axes that "discipline" academic production.

The new values that guide disciplinary production have been progressively constructed, diffused, and institutionalized, beginning with the institutions that formally organize the discipline at the national and international levels, as well as in the university centers at the undergraduate and postgraduate levels in political science around the world. The speed in which these values are consolidated within the discipline and that decisively affect production when it comes to quality, sophistication, and rigor, make contemporary production radically different than what it was a decade ago. High impact journals have progressively given incentives so that papers are guided by high sophistication, intensive use of formal models, and statistical analysis, both in the quantitative tradition or even in mixed-methods research.

The substantial effort by political science is to clearly promote a stronger "adherence" to rules of inference in a world of big data, technological abundance, and methodological diversity. Thus, it is possible to consider that political science is approaching a "credibility revolution" that has been occurring in economics since the 90s, in which research designs and associated methods take primacy over traditional data analysis. To move from correlations to understanding causation, it is necessary to introduce efforts to create identification, with substantially more rigorous research designs. Political science's connection with applied econometrics has been a natural path to developing this science as an empirical science.

Therefore, empirical political science would be moving towards more interfacing with econometrics, applied statistics, computational sciences, and even mathematics. Orientation by formal models has demanded a bigger understanding of the role of models in the social sciences, and its connection with hypothesis testing. In the case of political science, which is not true for most of the social sciences, the demands for testable, objective, and replicable knowledge takes focus.

Another important factor to consider as foundational reference in scientificity is the return to the experimental tradition (or quasi-experimental), which has been widely used in the analysis of significant phenomena such as corruption, clientelism, performance of institutions, collective action, and a whole gamut of other events that continue to be relevant. Political science today has a large scope of possibilities to produce valid inferential knowledge, whether with observational or experimental data, despite their limitations. Innovations developed by research designs

via experiments or quasi-experiments have allowed advances in a more robust understanding of causal effects or the causes of effects. The impact of these methods has substantially expanded the frontiers of analysis in comparative political science.

Another substantive innovation has been the progressive rupture of traditional cleavages between quantitative and qualitative methods that characterized the discipline before the 1990s. The traditional limits for data analysis techniques have also been progressively eased and, currently, higher standards of scientificity demand that researchers mobilize their creative efforts to produced valid inferential analysis with quantitative and qualitative strategies. The incentive for more integrated productions, despite their known limitations in practical research, has been noteworthy. The issue of integration between quantitative and qualitative methods since KKV has been profoundly formulated, and new paths to develop connections between these "cultures" (Mahoney and Goertz, 2012) have been enhanced in teaching and research institutions in political science.

Lastly, it is worth highlighting the important dimension related to the primacy of methods, which, as analyzed, introduces higher standards of reflexivity by scientists regarding new methodologies and research designs that can, indeed, be more adequate in treating specific research problems. Even in fields less quantitatively oriented, the concern with causal inferences has moved a substantial number of researchers in the comparative historical tradition to adhere to the new methodological "canons" available in the discipline. Political science would thus be moving towards what Sartori, in the 70s, conceptualized as "methodological awareness", needed to promote relevant knowledge when one is attempting to mesh empirical political science and science guided by formal models.

BIBLIOGRAPHICAL REFERENCES

Abadie, Alberto, Alexis Diamond, and Jens Hainmuller (2012). Comparative Politics and the Synthetic Control Method. Working Paper. Mimeo.

Ahmed, Amel and Rudra Sil (2012). When Multi-Method Research Subverts Methodological Pluralism—or, Why We Still Need Single-Method Research. Perspectives on Politics, 10 (4), pp.935-53.

Aldrich, John H., James E. Alt, and Arthur Lupia (2008). "The EITM Approach: Origins and Interpretations". In Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier (eds.) The Oxford Handbook of Political Methodology. Chapter 37. pp.828-843.

Angrist, Joshua D. and Jörn-Steffen Pischkie (2009). Mostly Harmless Econometrics: An Empiricist Companion. New Jersey. Princeton University Press.

APSA (2014). Symposium: The Set-Theoretic Comparative Method: Critical Assessment and the Search for Alternatives. Qualitative & Multi-Method Research, 12(1). Organized Section for Qualitative and Multi-Method Research.

Beach, Derek and Rasmus Brun Pedersen (2013). Process-Tracing Methods: Foundations and Guidelines. Ann Arbor. University of Michigan Press.

Blyth, Mark (2006). Great Punctuations: Prediction, Randomness, and the Evolution of Comparative Political Science. American Political Science Review, 100(4). pp.493-498.

Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier (2008). "Political Science Methodology". In Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier (eds.) The Oxford Handbook of Political Methodology. Chapter 1. pp.3-31
Brady, Henry E. and David Collier (2004). Rethinking Social Inquiry: Diverse Tools, Shared Standards. New York. Rowman & Littlefield Publishers, Inc.

Brady, Henry E., Collier and Jason Seawright (2006). "Toward a Pluralistic Vision of Methodology" Political Analysis, 14(3). pp14(3):353–368.

Braumoeller, Bear F. and Gary Goertz (2000). "The Methodology of Necessary Conditions". American Journal of Political Science, 44(3), pp.844-858.

Clarke, Kevin A. and David M. Primo. (2012). A Model Discipline: Political Science and the Logic of Representations. New York. Oxford University Press.

Collier, David (2008). Symposium: Case Selection, Case Studies, and Causal Inference. Introduction. Qualitative & Multi-Method Research. Fall. pp.2-4

Fearon, James (1991). Counterfactuals and Hypothesis Testing in Political Science. World Politics, 43(2), pp. 169-195

George, Alexander L. and Andrew Bennett (2005). Case Studies and Theory Development. Cambridge. , Massachussets. The MIT Press.

Gerber, Alan, Donald P. Green, and Edward H. Kaplan (2004). "The Illusion of Learning from Observational Research". In Shapiro, Ian, Rogers M. Smith, and Tarek E. Masoud (eds.) (2004). Problems and Methods in the Study of Politics. New York. Cambridge University Press. pp.251-273.

Gerring, John (2004). What Is a Case Study and What Is It Good for? American Political Science Review , 98 (2) pp. 341-354.

Gerring, John (2005). Causation A Unified Framework for the Social Sciences. Journal of Theoretical Politics, 17(2). pp. 163-198.

Goertz, Gary and Harvey Starr (2002). Necessary Conditions: Theory, Methodology, and Applications. Lanham. Rowman & Littlefield Publishers, Inc.

Granato , Jim , and Frank Scioli (2004) "Puzzles, Proverbs, and Omega Matrices: The Scientific and Social Significance of Empirical Implications of Theoretical Models (EITM) ". Perspectives on Politics, 2 (2), pp. 313 –23.

Granato, Jim, Melody Lo And M. C. Sunny Wong (2005). The Empirical Implications of Theoretical Models (EITM): A Framework for Methodological Unification. Working Paper presented at Southern Political Science Association. Mimeo.

Hess, Charlotte and Elinor Ostrom (2011). "A Framework for Analyzing the Knowledge Commons". In Hess, Charlotte and Elinor Ostrom (eds.) Understanding Knowledge as Commons: From Theory to Practice. Cambridge, Massachussets. The MIT Press. Pp.41-81.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. (2011) "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." American Political Science Review 105 (4), pp. 765–89.

King, Gary, Robert Keohane and Sidney Verba (1995). "The Importance of Research Design in Political Science" American Political Science Review, 89(2), pp. 454-456.

King, Gary and Lee Epstein (2001). The Rules of Inference. University of Chicago Law Review, XXX (1). pp.1-93.

King, Gary, Robert O. Keohane and Sidney Verba (1994). Designing Social Inquiry: Scientific Inference in Qualitative Research. New Jersey. Princeton University Press

Krogslund, Chris, Donghyum D. Choi and Mathias Poertner (2015). Fuzzy Sets on Shaky Grounds: Parameter Sensitivity and Confirmation Bias in QCA. Political Analysis, 23, pp.23-41.

Laitin, David (2002). "Comparative Politics: The State of the Subdiscipline" in Ira Katznelson and Helen V. Milner, eds. Political Science: The State of the Discipline New York: Norton. pp.630-660.

LIEBERMAN, Evan. (2005). Nested Analysis as a Mixed-Method Strategy for Comparative Research. *American Political Science Review, 99*(3), 435-452.

Mahoney, James (2007). Qualitative Methodology and Comparative Politics. Comparative Political Studies, 40, 122-144.

Mahoney, James (2010). After KKV: The New Methodology of Qualitative Research. World Politics, 62(1). pp.120-147.

Mahoney, James and Gerry Goertz (2012). A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research in the Social Sciences. New Haven. Princeton University Press.

Mahoney, James, Erik Kimball and Kendra L. Koivu (2008). The Logic of Historical Explanation. Comparative Political Studies, 42(1). pp.114 -146.

Mahoney, James. (2000) 'Strategies of Causal Inference in Small-N Analysis', Sociological Methods and Research 28(4): 387–424.

Mahoney, James. (2003). "Strategies of Causal Assessment in Comparative Historical Analysis", in J. Mahoney and D. Rueschemeyer (eds) Comparative Historical Analysis in the Social Sciences, pp. 337–72. New York: Cambridge University Press.

Manski, Charles F. (1999). Identification Problems in the Social Sciences. Cambridge. Harvard University Press.

Marsh, D. and Savigny, H. (2004) 'Political Science as a Broad Church: The Search for a Pluralist Discipline', Politics 24(3): 155–168.

Merriam, Charles E. (1921). The Present State of the Study of Politics. American Political Science Review, 15(2), pp.173-185.

Monroe , Burt L . (2013). "The Five Vs of Big Data Political Science: Introduction to the Virtual Issue on Big Data in Political Science ". Political Analysis , Virtual Issue, 5.

Monroe , Burt L., Jennifer Pan, Margaret E. Roberts, Maya Sen, and Betsy Sinclair (2015). No! Formal Theory, Causal Inference, and Big Data Are Not Contradictory Trends in Political Science. American Political Science Association, PS. January, pp. 71-74.

Morgan, Stephen L. and Christopher Winship (2007). Counterfactuals and Causal Inference: Methods and Principles for Social Research. New York. Cambridge University Press.
Moses, Jonathon, Benôit, Rihoux, , and Bernhard Kittel (2005). Mapping Political Methodology: reflections on an European Perspective. European Political Science, 4. pp.55-68.

National Science Foundation (2002). Empirical Implications of Theoretical Models Report. 2002. Political Science Program, Directorate of Social, Behavioral, and Economic Sciences. Arlington, Va.

Norris, Pipa. (1997) 'Towards a More Cosmopolitan Political Science?' European Journal of Political Research, 31(1–2): 17–34.

Przeworski, Adam (2007). "Is The Science of Comparative Politics Possible?". In Boix, Carles and Susan Stokes (orgs.). Oxford Handbook of Comparative Politics. New York. Cambridge University Press. pp.147-171.

Ragin, Charles C. (1987). The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies. Berkeley. University of California Press.

Ragin, Charles C. (2000). Fuzzy-Set Social Science. Chicago. The University of Chicago Press.

Rezende, Flávio da Cunha (2010). Analytical Challenges for Neoinstitutional Theories of Change in Comparative Political Science. Brazilian Political Science Review, vol 3(2). pp.98-126.

Rezende, Flávio da Cunha (2011). "A Nova Metodologia Qualitativa" e as Condições Essenciais de Demarcação entre Desenhos de Pesquisa na Ciência Política Comparada. Revista
Política Hoje, vol 20(1), pp.218-252.

Rezende, Flávio da Cunha (2011). Razões emergentes para a validade dos estudos de caso na ciência política comparada. Revista Brasileira de Ciência Política, 6. pp. 297-337.

Rezende, Flávio da Cunha (2015). Modelos de Causação e o Pluralismo Inferencial em Ciência Política. Working Paper #1. Núcleo de Epistemologia e Método Comparado. Recife. Universidade Federal de Pernambuco. Departamento de Ciência Política. Mimeo.

Rihoux, Benoit and Charles C. Ragin (2009). Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Applied Social Resarch Methods Series, vol 51. Thousand Oaks, California. SAGE Publications.

Rohlfing, Ingo (2012). Case Studies and Causal Inference: an integrative framework. New York. Palgrave MacMillan. ECPR Research Methods Series.
Sekhon, Jasjeet S (2009). "Opiates for the Matches: Matching Methods for Causal Inference." Annual Review of Political Science 12(1), pp.487–508.
Schneider, Carsten Q. and Claudius Wageman (2012). Set-Theoretical Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis. New York. Cambridge University Press.

Shapiro, Ian (2002). Problems, Methods, and Theories in the Study of Politics, or What´s Wrong with Political Science and What to do About it. Political Theory, vol 30(4). pp. 588- 611.

Weller, Nicholas and Jeb Barnes (2014). Finding Pathways: Mixed-Method Research for Studying Causal Mechanisms. New York. Cambridge University Press.

Um mapa para a transparência e replicabilidade na ciência social empírica: o Protocolo TIER

Amanda Domingos - Federal University of Pernambuco
Ian Rebouças Batista- Federal University of Pernambuco

Resumo

Transparência e replicabilidade são o padrão ouro da pesquisa científica. Mas, alcançar esse padrão pode ser um caminho longo e trabalhoso. Nesse artigo, apresentamos uma maneira de aumentar a transparência de sua pesquisa através da documentação. Argumentamos que o Protocolo TIER é uma importante ferramenta de transparência e de auxílio na organização de materiais para replicação. Trata-se de uma metodologia de documentação que compila os principais materiais produzidos numa pesquisa empírica social. Além disso, destacamos a adequação do protocolo nas três dimensões do padrão de replicação (substantiva, pedagógica e transparente), uma vez que contribui para o avanço da ciência, introduz jovens graduandas à análise de dados de uma forma prática e permite a transparência quanto aos procedimentos empregados na pesquisa. Por fim, discutimos atividades relacionadas ao Projeto TIER que contribuem com os objetivos do protocolo. Especificamente, apresentamos um exercício de introdução à análise de dados a ser realizado no R, onde a documentação do protocolo é requerida e sua aplicação, portanto, ilustrada.

Palavras-chave: Transparência; Replicabilidade; Documentação; Protocolo TIER

Abstract

Transparency and replicability are the gold standard in empirical research. But reaching that standard can be a long and tortuous journey. In this article, we present a way to increase the transparency of research through documentation. We argue that the TIER Protocol is an essential tool for transparency and assistance in organizing replication materials. It is a documentation methodology that compiles the main materials produced in empirical social research. Also, we highlight the adequacy of the protocol in the three dimensions of the replication standard (substantive, pedagogical, and transparent) – since it contributes to the advancement of science, introduces young undergraduate students to data analysis practically, and allows transparency of research procedures. Finally, we discuss activities related to the TIER Project that contribute to the objectives of the protocol. Specifically, we present a data analysis introduction exercise for R, where the protocol documentation is required and its application, therefore, exemplified.

Keywords: Replicability; Documentation; TIER Protocol

1. Introdução

Baixar o banco de dados, tratar o banco, elaborar *scripts*, programar funções, executar comandos, analisar os resultados. Se essas são etapas corriqueiras de sua pesquisa acadêmica, existe um protocolo que irá te ajudar a organizar a sua pesquisa e a divulgar, de maneira transparente e replicável, seus resultados.

Princípios e práticas de transparência e replicabilidade na pesquisa científica têm mobilizado crescente atenção da comunidade acadêmica nas ciências sociais, ganhando destaque em publicações nacionais (PARANHOS, 2012; PARANHOS et al, 2014; FIGUEIREDO et al, 2019; AVELINO e DESPOSATO, 2018) e internacionais (CHRISTENSEN, 2016; MUNAFÓ et al, 2017; CHRISTENSEN e MIGUEL, 2018; STOCKEMER, KOEHLER E LENZ, 2018; DE LA GUARDIA e STUDY, 2019), sendo ainda incentivados através de iniciativas inter-universitárias como *Berkeley Initiative for Transparency in the Social Sciences* (BITSS) e *Teaching Integrity in Empirical Research* (TIER), que promovem cursos, *workshops* e eventos voltados à troca de experiências e ideias em prol da transparência.

Identificada como padrão-ouro (KING, 1995), a transparência quanto aos procedimentos adotados em uma pesquisa é aqui entendida como a disponibilização dos materiais utilizados e a indicação das etapas percorridas até se chegar aos resultados. Para além dos benefícios que isso traz ao adequado funcionamento da ciência, permitindo acesso público a suas escolhas metodológicas e possibilitando que seu trabalho seja replicado e avaliado mais rigorosamente, a transparência permite: (i) teste de hipóteses mais preciso, (ii) a colaboração com outros pesquisadores com os mesmos interesses e (iii) aumento na credibilidade da autora. Argumentamos que a pesquisadora transparente, que documenta seu trabalho à medida que é feito, tem maior facilidade na elaboração de sua pesquisa. Uma documentação maior traz benefícios em termos de reputação científica, uma vez que contribui para boas práticas da ciência e gera benefícios práticos, como o aumento de citações (FIGUEIREDO et al, 2019).

A partir disso, alguns instrumentos têm sido empreendidos para difundir e implementar os princípios e as práticas da transparência e replicabilidade nas Ciências Sociais, sendo as três práticas principais: planos de divulgação, pré-registro e abertura de dados e materiais (MIGUEL et al, 2014). O objeto desse estudo se enquadro no terceiro tipo. Nesse artigo, apresentamos e sugerimos a adoção do protocolo TIER, uma metodologia para documentação adequada dos arquivos utilizados no desenvolvimento da pesquisa (bancos de dados, *scripts, outputs*) e posterior disponibilização desses materiais para replicação. Acreditamos que a adoção de uma regra específica quanto à organização dos arquivos de pesquisa aumenta a transparência e favorece a replicabilidade do estudo.

O artigo se divide como segue. Na próxima seção, apresentamos o padrão-ouro da ciência em termos de transparência dos procedimentos e replicabilidade dos resultados, indicando porque a leitora deve almejar esses valores e práticas. Na sequência, apresentamos o protocolo TIER, destacando os seus aspectos fundamentais. Em seguida, enfatizamos os benefícios e incentivos a se adotar uma estratégia de documentação e disponibilização para tornar a pesquisa transparente. Por fim, elencamos algumas atividades desenvolvidas pelo Projeto TIER que corroboram e buscam difundir as boas práticas de transparência e replicabilidade nas ciências sociais e apresentamos um exercício introdutório ao protocolo. As considerações finais encerram este trabalho.

2. Por que um Protocolo para transparência?

Dentre as características de uma boa pesquisa científica estão a inovação conceitual, o rigor metodológico

e a riqueza do conteúdo substantivo (PRZEWORSKI e SALOMON, 1998). Nas Ciências Sociais, no entanto, outros dois elementos foram adicionados a essas características nas últimas décadas: a transparência e a replicabilidade, tidos como o padrão-ouro da pesquisa empírica (KING, 1995; PARANHOS et al, 2014; JANZ, 2016). Os benefícios de uma pesquisa transparente são diversos, tanto para a disciplina como um todo quanto para a pesquisadora, em particular. A adoção de práticas transparentes permite que a roda da ciência gire de forma mais rápida, ao permitir o teste de hipótese mais preciso, facilitar a colaboração com outros pesquisadores com os mesmos interesses, aumentar a credibilidade da pesquisa, além de retirar a costumeira atenção da significância estatística dos resultados e aumentar a relevância da pesquisa *per se* (DE LA GUARDIA e STUDY, 2019).

No que se refere a benefícios para autores e periódicos, Figueiredo et al (2019) apresentam sete razões para adotar procedimentos transparentes em pesquisas empíricas nas ciências sociais: (I) prevenir que as pesquisas contenham erros honestos ou, até mesmo, fraudes em suas análises empíricas; (II) tornar a escrita de artigos científicos mais fácil, uma vez que permite o teste de outras hipóteses com a mesma base de dados; (III) procedimentos transparentes permitem que pareceristas tenham acesso a dados e *scripts* e deem recomendações mais precisas sobre como melhorar o trabalho; (IV) ao garantir a sistematização de dados brutos e analisados, permite que o trabalho cientifico do autor continue; (V) ajudam a construir reputação científica, ao demonstrar a boa-fé por apresentar material de replicação; (VI) ajudam a aprender análise de dados, ao permitir que os alunos tenham contato com dados e procedimentos estatísticos utilizados na vida real e (VII) aumenta o impacto do trabalho acadêmico, uma vez que evidências empíricas demonstram que trabalhos que adotam procedimentos e práticas transparentes são duas vezes mais citados.

Por outro lado, os custos da não adoção de procedimentos de transparência e replicabilidade também são variados. Pesquisas que não mantém procedimentos públicos podem contribuir para a diminuição da confiança em seus resultados, e, além disso, podem ser uma barreira à publicação em periódicos de alto fator de impacto que solicitam matérias de replicação – o que diminui o alcance da pesquisa (MUNAFO et al, 2017; DE LA GUARDIA e STUDY, 2019). Diante de tudo isso, uma pessoa poderia argumentar que não adota princípios e práticas de transparência por ser muito trabalhoso. Aqui, vale a pena reforçar que o padrão-ouro de replicação não solicita que um indivíduo reproduza os resultados apresentados em livros ou artigos científicos, apenas recomenda que sejam disponibilizadas tanta informação quanto for suficiente para que, caso algum pesquisador deseje, consiga reproduzi-los (KING, 1995).

Apesar disso, a prática de transparência não são a regra. Cerca de 67% dos artigos publicados na *American Political Science Review* entre 2013 e 2014 não são replicáveis (KEY, 2016). No Brasil, apenas 5% dos artigos publicados entre 2012 e 2016 em periódicos da área de Ciências Sociais eram passíveis de replicação completa (AVELINO e DESPOSATO, 2018)[1,2]. O padrão ainda é a obscuridade, não a transparência. Nesse sentido, não é incomum encontrar pesquisadores que não tornam públicos os procedimentos e práticas que os levaram às conclusões de seus produtos científicos. Questionamos a cara leitora: quantas vezes, na produção de sua própria pesquisa empírica, se esforçou para lembrar os caminhos que a levaram ao resultado que encontrou? Mais do que isso, o que você tem feito para ser transparente? Se pensou um tanto antes de responder essa última pergunta, esse artigo tem uma contribuição a te fazer. Nas próximas seções, vamos sugerir um protocolo

1 O relatório é fruto de do projeto "*Research Transparency in Brazilian Political and Social Science: A First Look*" dos autores junto à *Berkeley Initiative for Transparency in the Social Sciences* (BITSS). Para maiores informações, visite: < https://www.bitss.org/projects/research-transparency-in-brazilian-political-and-social-science-a-first-look/>. Acesso em: 20 mar. 2010.

2 O relatório final está publicamente disponível em: <https://osf.io/f279z/>. Acesso em 20 mar. 2020.

de organização para facilitar a implementação de práticas transparentes e discutir um pouco sobre o processo de documentação de uma pesquisa empírica.

3. O protocolo TIER

O protocolo TIER é uma metodologia de documentação dos materiais de replicação de uma pesquisa empírica. De modo simplificado, o protocolo traz uma indicação geral sobre a organização dos arquivos utilizados em um estudo, indica a construção de pastas e quais informações devem ser disponibilizadas em cada uma delas. O protocolo foi desenvolvido originalmente para lidar com um problema de sala de aula. Os trabalhos que eram entregues pelos alunos de Introdução à Estatística, disciplina ministrada por Richard Ball na graduação de economia na Haverford College (EUA), apresentavam problema similar àqueles existentes nos artigos científicos publicados nos periódicos acadêmicos de grandes revistas: eles não eram transparentes. Ao entregarem os trabalhos finais da disciplina com resultados errados, era praticamente impossível para o professor entender em que momento da atividade de análise dos dados foram realizados procedimentos equivocados, o que impedia a adequada correção do professor e, consequentemente, o processo de aprendizado do aluno.

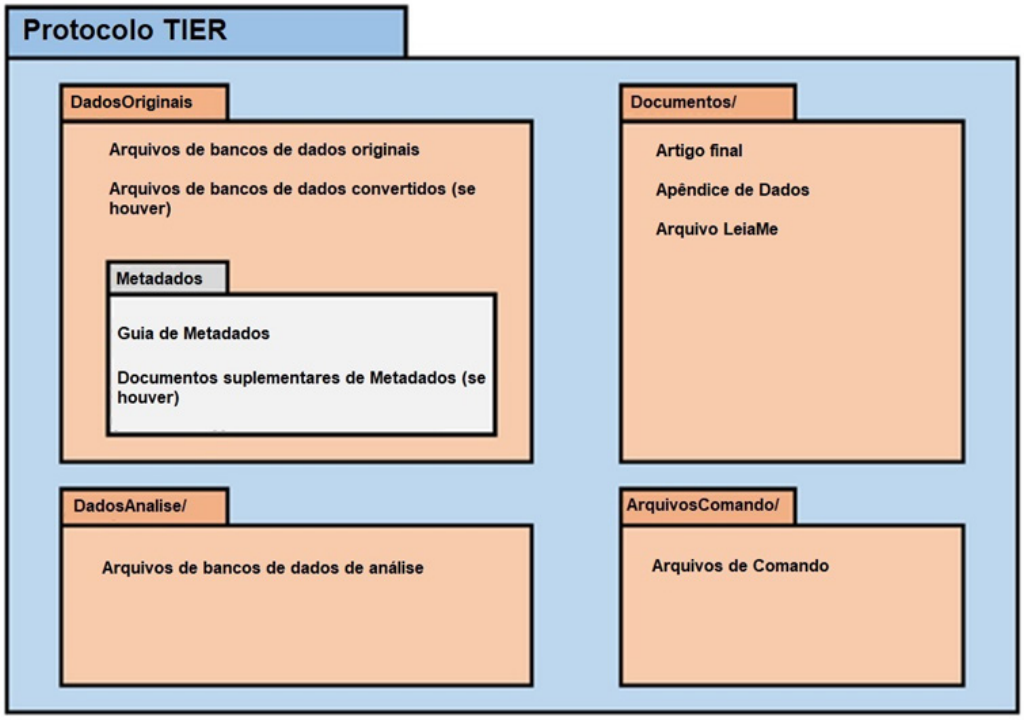
Para garantir que existisse um padrão para a entrega das atividades, tornando possível que o professor corrigisse de forma eficiente, tendo acesso aos *scripts* e dados de forma inteligível e organizada, Ball elaborou, junto com o bibliotecário do College, Norm Medeiros, o protocolo TIER (BALL e MEDEIROS, 2012). O protocolo consiste em uma organização e hierarquia de pastas específica que separa as informações sobre determinada pesquisa (dados brutos, dados analisados, *scripts*, etc.) e facilita a replicação dos achados. Apesar dessa origem pedagógica, o Protocolo TIER vai além das portas da sala de aula. Diversos pesquisadores na área de Ciências Sociais podem utilizar o protocolo como uma maneira de aumentar o nível de transparência de suas pesquisas. Nesse sentido, argumentamos que a adoção do Protocolo TIER acaba por colaborar nas três dimensões do padrão de replicação: substantiva, pedagógica e transparência (PARANHOS et al, 2013).

O protocolo TIER possui uma orientação geral, tal como se encontra ilustrado na Figura 1, e possui versões que aprimoram a organização e a operacionalização da replicação de seu trabalho, como a versão 4.0 apresentada na Figura 2. Discutiremos as duas versões na sequência, mas antes alguns esclarecimentos procedimentais dessa metodologia organizacional que o TIER propõe ajudam a entender sua proposta.

As diferentes pastas que irão compor o material de replicação de seu trabalho devem ser criadas no início da pesquisa. Elas acompanharão a pesquisadora ao longo de todo o processo: tratamento do banco de dados original, processamento das análises, elaboração do artigo e dos apêndices. A pesquisadora deve ir salvando e trabalhando com os arquivos nas pastas adequadas e referenciando, nos seus *scripts* ou arquivos de comando, os diretórios de trabalho de acordo com a hierarquia de pastas, o que torna fundamental que essa organização inicial seja mantida ao longo do trabalho. Por exemplo, no *script* onde você irá tratar o banco de dados original e irá gerar o banco da análise, você deve carregar o banco de dados original a partir da pasta *DadosOriginais* e deve salvar o banco tratado na pasta *DadosAnalise* (Figura 1). A ideia é que usando um *software* como o R, seu *script* contenha as linhas necessárias para abrir o banco original, que estará na pasta adequada, e salvar o banco de análise na pasta correta[3].

3 Como uma medida de replicabilidade, o ideal é que você adote um diretório de trabalho relativo. Para maiores informações, ver: https://ssc.wisc.edu/sscc/pubs/R_intro/book/1-10-paths-and-working-directories.html. Acesso em 07 de outubro de 2020.

Figura 1: Visão geral da documentação sugerida pelo Protocolo TIER



Fonte: Project TIER. Tradução dos autores4

Mas que organização de pastas você deve adotar, uma vez que as pastas propostas na Figura 1 e na Figura 2 são distintas? O objetivo do protocolo é fornecer uma orientação para organização e disponibilização de todo o material digital utilizado na elaboração de uma pesquisa empírica. É importante que essa orientação seja coerente e que seja, se não intuitiva, justificada e apresentada. Sendo assim, a exata configuração das pastas e dos arquivos fica a gosto do freguês. A Figura 1 é o exemplo mais genérico, a partir do qual se elaboraram quatro outras versões, e a Figura 2 é a formatação sugerida mais recente pelo Projeto (4.0). Contudo, se o argumento que o protocolo faz, sobre organização da pesquisa e sobre transparência, convencer a pesquisadora/leitora, você é encorajada a organizar o seu material para disponibilização a partir de um protocolo que funcione para você, ou seja, na sua própria hierarquia de pastas que achar conveniente.

Para que os protocolos individuais funcionem a qualquer pesquisador interessado em replicar o seu trabalho e conferir os resultados, são necessárias algumas orientações básicas que compõe a mensagem do TIER. O primeiro deles é a elaboração de arquivos LeiaMe (geralmente em .txt, para facilitar a leitura independente do sistema operacional utilizado) que orientarão a atividade de replicação desde o início. Uma das premissas de uma replicação transparente é que arquivos desse tipo são os primeiros a serem lidos, pois além de apresentar e explicar a configuração de pastas adotada, que a replicadora estará se deparando ao ler, são esses documentos de texto que dizem que arquivos abrir primeiro e qual será a sequência do trabalho de replicação. É interessante, portanto, que ele apareça logo na primeira pasta, como consta na Figura 2. A elaboração de um arquivo LeiaMe claro e objetivo acaba demonstrando a preocupação da pesquisadora com a replicação futura de seus achados. Assim, deve fornecer um pano-

4 O *template* utilizado é o *template* original utilizado pelo Projeto TIER. Disponível em: https://www.projecttier.org/tier-protocol/specifications/#overview-of-the-documentation. Acesso em 23 de março de 2020

rama completo da maneira com que as pastas estão organizas e das etapas de replicação.

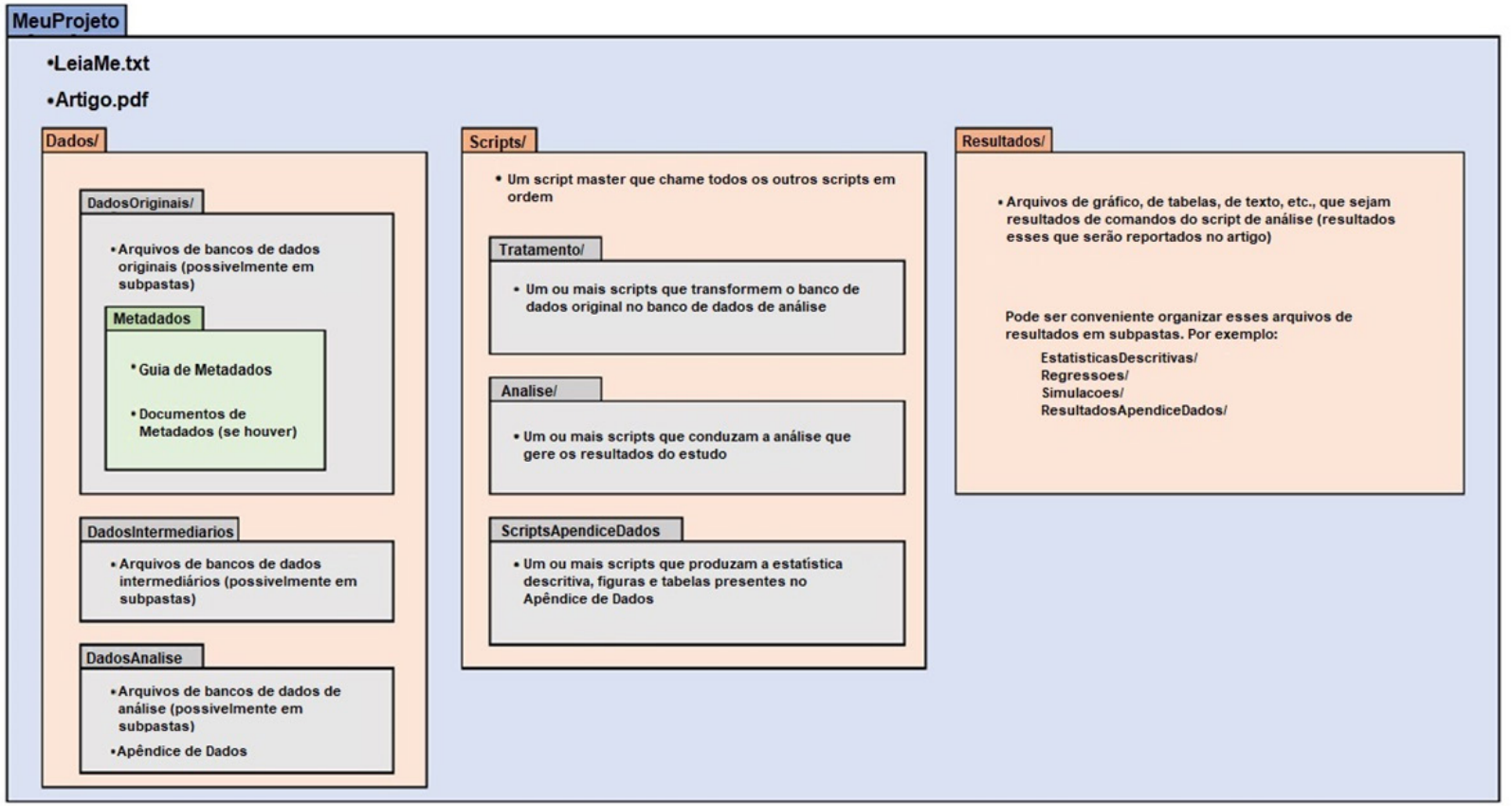
Por exemplo, para um trabalho que utiliza o protocolo inspirado no apresentado na Figura 2, o arquivo LeiaMe pode conter um texto semelhante a esse: "Inicialmente, vá até a subpasta *Tratamento*, localizada na pasta *Scripts,* e abra o *script* Tratamento.R. Nele você irá tratar o Banco.dta, que está na subpasta *DadosOriginais,* dentro da pasta *Dados*, e irá gerar o BancoNovo.dta, que será salvo na subpasta *DadosAnalise*. Em seguida, também dentro da pasta *Scripts*, vá até a subpasta *Analise*, abra o *script* Analise.R. Nele você irá gerar os Gráficos 1, 2 e 3, que serão salvos na pasta *Resultados*". Os arquivos LeiaMe podem ainda conter informações sobre o *software* de análise de dados utilizado, os endereços da *web* onde os bancos originais foram baixados ou ainda informações para contato com a pesquisadora.

Semelhante ao LeiaMe, mas com outra função, o protocolo TIER considera ainda que é fundamental que a pesquisadora elabore um Apêndice de Dados (ou *codebook*). É nesse documento que serão descritas com detalhes as variáveis que constam no banco de dados de análise, incluindo o tipo dessa variável, como ela é mensurada e como ela foi construída (se já estava pronta no banco de dados original ou se foi elaborada pela pesquisadora no processo de tratamento dos dados). Pode-se ainda incluir estatísticas descritivas para cada uma das variáveis. É nesse documento também que maiores especificidades sobre as versões do *software* utilizado e dos pacotes estatísticos necessários são fornecidas.

Outra orientação básica que permeia o protocolo TIER em qualquer de suas versões, e que deve estar presente também num possível protocolo pessoal inspirado no TIER, é a separação dos arquivos por tipos. Tanto na Figura 1 quanto na Figura 2 existem pastas próprias para bancos de dados, *scripts* e documentos. Analisemos de modo mais aprofundado essas figuras agora.

Na Figura 1, a orientação geral do protocolo, tem-se quatro pastas ao todo: *DadosOriginais*; *DadosAnalise*; *AquivosComando*; *Documentos*. Dentro da pasta *DadosOriginais*, encontramos os arquivos de bancos de dados originais, baixados diretamente da fonte que for, os bancos de dados convertidos, caso a pesquisadora tenha que converter o formato do banco de dados original baixado (note que mesmo nesse caso é importante manter o banco também no formato original), e uma subpasta chamada de *Metadados*, onde se encontram um Guia dos Metadados (um Apêndice de Dados para o banco original, incluindo a citação bibliográfica indicada, o identificador DOI, a data em que o banco foi baixado e o endereço de web) e documentos suplementares desse banco original (se houver). Na pasta *DadosAnálise*, tem-se o(s) banco(s) de dado(s) tratado(s), com o qual a pesquisadora fará as análises do seu artigo. Na pasta de *Documentos*, encontram-se o artigo final, o Apêndice de Dados e o arquivo LeiaMe. Por fim, na pasta *ArquivosComando*, estarão todos os *scripts* utilizados na pesquisa, desde o tratamento do banco de dados original até as análises realizadas.

Figura 2: Protocolo TIER na sua versão 4.0 (2020)



Fonte: Project TIER. Tradução dos autores

A Figura 2, por sua vez, baseia-se em uma elaboração preliminar da versão 4.0 do Protocolo, apresentada em novembro de 2019. Nela, na pasta principal, tem-se logo de cara dois arquivos, o artigo final e o arquivo LeiaMe, que irá orientar o procedimento daquele que deseja replicar o presente trabalho, e três pastas: *Dados*, *Scripts* e *Resultados*. Na pasta *Dados*, encontram-se três subpastas: *DadosOriginais* (com a pasta de *Metadados* dentro dela); *DadosIntermediarios*, onde estarão bancos de dados por ventura convertidos em outro formato; e *DadosAnalise*, onde também estará o Apêndice de Dados. Na pasta *Scripts*, encontra-se um *script* "*master*", que combinará todos os *scripts* do trabalho, mas também três subpastas: *Tratamento*, *Analise* e *ScriptApendiceDados*, cada pasta contendo o *script* que realiza cada uma das etapas referentes. Por fim, na pasta *Resultados*, estarão todos os arquivos que foram gerados ao longo da análise a partir dos *scripts*, como tabelas, gráficos e figuras, que podem, por sua vez, estar subdivididas também em subpastas.

Como se percebe, as duas versões do protocolo compartilham do mesmo ideal de documentação e transparência da pesquisa realizada, ainda que adotando organizações e hierarquias de pastas diferentes. Compreendendo quais os objetivos do protocolo e a lógica da separação dos arquivos presentes em ambas as versões, é possível então que a pesquisadora interessada numa pesquisa transparente e replicável possa, se não adotar exatamente alguma das versões expostas, adotar uma versão própria do protocolo TIER. Uma versão que faça sentido para a pesquisadora e beneficie o seu fluxo de trabalho, ao mesmo tempo que seja clara e acessível a qualquer interessado em replicar os resultados de sua pesquisa.

Após elaborado todo o material seguindo o Protocolo TIER, a disponibilização deve ocorrer da maneira mais acessível possível. Como a transparência no uso dos dados é um assunto em ascensão, temos a nosso dispor uma gama de repositórios digitais que fornecem a estrutura adequada para essa disponibilização. O *Open Science Framework*[5] (OSF), o *Dataverse*[6] e o *Github*[7] são três exemplos de repositórios adequados para o compartilhamento de todo o material de seu trabalho, inclusive com a possibilidade de divisão nas próprias pastas do protocolo. Nesse artigo, consideramos que o trabalho da pesquisadora tenha uma abordagem quantitativa, dada a natureza do protocolo que beneficia trabalhos deste tipo. Mas, também existem repositórios digitais para trabalhos com uma abordagem qualitativa[8], a saber: *Qualitative Data Repository*[9], UK data archive[10], National Anthropological Archives[11].

4. Documentação, protocolo TIER e dimensões do padrão de replicação

Em 1995, ano de lançamento do dossiê "*verification/replication*"[12], a realidade da produção científica nas ciências sociais era diferente. O dossiê propôs aos pesquisadores da área tentar tornar disponíveis maiores detalhes sobre as escolhas empíricas realizadas em suas pesquisas em notas de rodapé e apêndices metodológicos, mas esses esforços esbarravam nos limites impostos pelos periódicos e editoras de livro para divulgar tais detalhes. Por outro lado, a disponibilização de informações e armazenamento eram um problema: as facilida-

5 Ver: https://osf.io . Acesso em 20 mar. 2020.

6 Ver: https://dataverse.org/ . Acesso em 20 mar. 2020.

7 Ver: https://github.com/ . Acesso em 20 mar. 2020.

8 Para maiores informações sobre repositórios que aceitam dados qualitativos, ver: https://www.sesync.org/list-of-qualitative-data-repositories. Acesso em 24. Mar. 2020.

9 Ver: https://qdr.syr.edu/. Acesso em 24 mar. 2020.

10 Ver: https://www.ukdataservice.ac.uk/manage-data. Acesso em 24 mar. 2020.

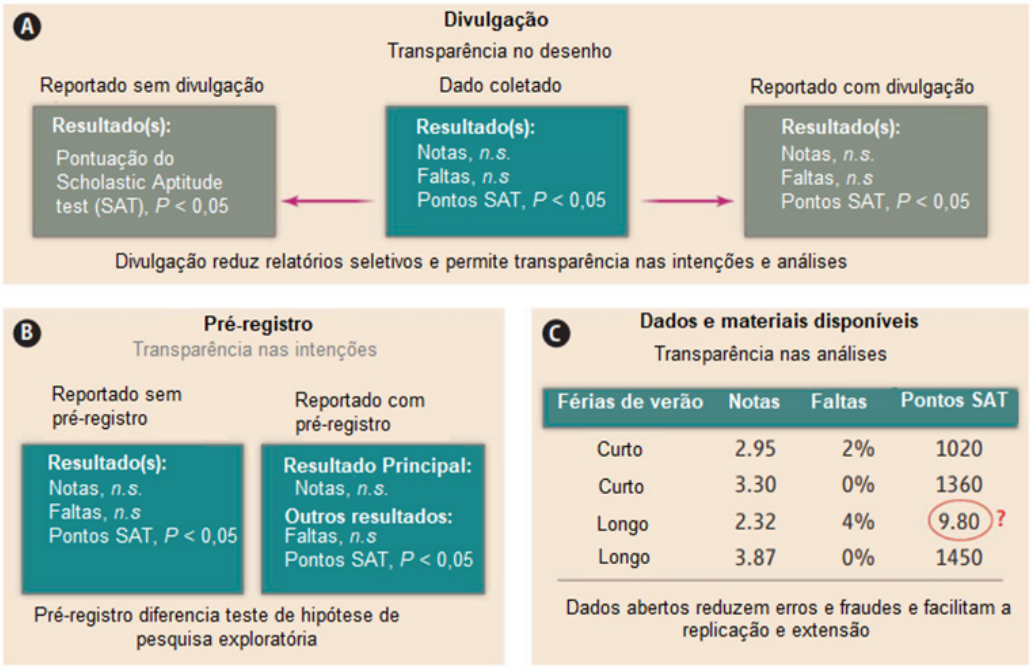
11 Ver: http://anthropology.si.edu/naa/search.html. Acesso em 24. mar. 2020.

12 O dossiê está disponível em: < https://www.cambridge.org/core/journals/ps-political-science-and-politics/issue/DAEAAA8412CD56791EAFCC75082A4C15>. Acesso em 24 mar. 2020.

des de *drives* e nuvens não existiam; o arquivo digital relacionado a publicações do *Inter-University Consortium for Political and Social Research* (ICPSR)[13], por exemplo, apenas seria lançado algum tempo depois do dossiê (ALVAREZ et al, 2018).

Nos dias atuais, com o surgimento de plataformas e repositórios *online* adequados, a disponibilização de informações tem se tornado cada dia mais fácil. Nesse sentido, diversos esforços têm sido empreendidos para difundir e implementar os princípios e as práticas da transparência e replicabilidade nas Ciências Sociais. Tais empenhos têm se concentrado em três práticas principais: planos de divulgação, pré-registro e abertura de dados e materiais (MIGUEL et al, 2014), conforme ilustrado na Figura 3.

Figura 3: Três mecanismos para aumentar a transparência nos relatórios científicos (Demonstração com a pergunta de pesquisa "as férias de verão mais curtas melhoram os resultados educacionais?", onde "n.s" significa p-valor > 0,05.)



Fonte: Miguel et al (2014), p.30. Tradução dos autores.

Os planos de divulgação surgem para advogar pela produção de relatórios sistemáticos que buscam garantir que os pesquisadores documentem e divulguem a forma com que procederam quanto a coleta e análise dos dados – essa é uma estratégia muito utilizada na área da medicina e implementada em alguns periódicos acadêmicos na Ciência Política e Psicologia[14] (MIGUEL et al, 2014). Por sua vez, o pré-registro é um documento público onde o pesquisador especifica variáveis, procedimentos de processamentos de dados e técnicas de análise (OLKEN, 2015). A adoção do pré-registro visa diminuir o viés de publicação[15], prevenir a apresentação de resultados seletivos da pesquisa e reforçar a credibilidade desta (CASEY et al, 2012). Por fim, a disponibilização de dados e materiais permitem que outros pesquisadores reproduzam os resultados da pesquisa, testem hipóteses remanescentes com os dados e identifiquem resultados não declarados ou, até mesmo, fraudes (MIGUEL et al, 2014; MUFANO et al, 2017; CHRISTENSEN e MIGUEL, 2018).

13 Disponível em: <www.icpsr.umich.edu/icpsrweb/deposit/pra/index.jsp>. Acesso em 24 mar. 2020.

14 De acordo com o levantamento dos autores, as revistas que adotaram políticas em direção a esse tipo de procedimento são: *Journal of Experimental Political Science*, *Management Science* e *Psychological Science*. Ver: Miguel et al (2014), p. 30.

15

É na terceira prática que o Protocolo TIER se enquadra. Mais especificamente, na documentação que precede a publicitação de dados e materiais de replicação. A adoção de políticas de replicabilidade por periódicos de alto fator de impacto forçam que as pesquisadoras documentem, de forma sistematizada, os arquivos de suas pesquisas. Apesar da demanda, a qualidade dos materiais disponibilizados não é alta. Stockemer, Koehler e Lenz (2018) analisaram todos os artigos publicados em 2015 nos três principais periódicos da área de comportamento político[16] e identificaram que em 25% deles os dados e os códigos estavam tão mal organizados que tornou a replicação impossível.

Materiais de replicação bem organizados aprimoram a extensão de pesquisas relacionadas e possibilitam que as análises sejam mais acessíveis (ALVAREZ et al, 2018).

No nível mais básico de preparação para se tornar um pesquisador transparente, é recomendado que se adote algum tipo de protocolo de organização desses dados, com o objetivo último de ter dados que sejam passíveis de reutilização e humanamente compreensíveis (FIGUEIREDO et al, 2019; OSINSKI, 2019).

Alvarez, Key e Nuñez (2018) apresentam algumas recomendações para aumentar a organização, clareza e utilização dos dados. Nele, os autores dão dicas sobre como arquivar os materiais relacionados à pesquisa de forma útil e segura, como produzir arquivos LeiaMe que sirvam como ajuda para replicações futuras, indicações sobre como devem ser nomeados os arquivos nos materiais de replicação, além de indicar como devem ser reportados *softwares* (utilizados ou criados pelo autor), *scripts* e *outputs* gerados na pesquisa. O Quadro 1 sumariza as sugestões dos autores.

16 Os autores analisaram todos os artigos presentes nos volumes de 2015 dos seguintes periódicos acadêmicos: *Electoral Studies*, *Party Politics* e *Journal of Elections, Public Opinion & Parties*.

Quadro 1 - Recomendações para criação de materiais de replicação (Alvarez et al, 2018)

Processos	Recomendações
Arquivamento	Os materiais de replicação devem ser armazenados em arquivos permanentes que estejam disponíveis por longos períodos. Atualmente, os mais confiáveis são os que têm garantias institucionais, sejam de consórcios ou universidades; Autores devem salvar seus arquivos de modo que possam ser utilizados no futuro (ex: arquivo separado por vírgula), ao invés de formatos software-specifc.
Arquivos LeiaMe	O arquivo deve ser claro e suficientemente detalhado, mas não em demasia. Alguns itens deveriam ser incluídos, a saber: (a) referência do artigo ou publicação associada; (b) uma breve explicação dos arquivos e tipos de arquivos incluídos, ex: dados originais, dados processados, scripts, etc; (c) uma indicação da ordem em que os scripts devem ser executados; (d) uma lista de softwares, pacotes de softwares e sistema operacional utilizado para produzir os resultados do artigo; (e) se extensões incomuns forem utilizadas, o autor deve indicar como proceder com esses arquivos.
Nomes dos arquivos	Os nomes dos arquivos devem ser fáceis de compreender e prover informação sobre seu conteúdo, isso é especialmente importante para matérias de replicação que incluem vários scripts e arquivos de dados. Caso existam arquivos que devem ser utilizados seguindo determinada ordem, eles devem ser nomeados de modo que incluía a ordem (ex: "1_DadosProcessados"; "2_Estimação", etc.).
Softwares e pacotes criados pelo usuário	Em casos onde o autor tenha criado um pacote estatístico, é necessário que estes estejam (idealmente) arquivados em um local estável, como o CRAN para o R. Se possível, o autor deve incluir uma cópia do software ou pacote nos materiais de replicação.
Compatibilidade com sistemas operacionais	Idealmente, é necessário que os autores garantam que seus materiais de replicação funcionem nos sistemas operacionais mais comuns. Caso não seja possível, devem indicar em que sistema eles funcionam.
Scripts	(a) devem conter uma breve discrição do que fazem, quais as dependências, quais pacotes são necessários e quais são os outputs; (b) deve-se evitar linhas sem especificação do que se faz, pois tal indicação ajuda a identificar partes especificas do gerenciamento e processamento de dados e a identificar erros ou possíveis problemas; (c) deve-se planejar bem e evitar recodificação de dados no meio da análise, ao menos quando seja extremamente necessário.
Resultados	Os materiais de replicação devem conter tabelas, gráficos e figuras da mesma forma que aparecem no artigo.
Resultados excluídos	Geralmente, materiais de replicação apresentam scripts ou dados que não foram apresentados no artigo ou que estavam vinculados a um apêndice online. Nesse caso, o autor deve diferenciá-los dos que estão no artigo.
Resultados intermediários	Algumas vezes os artigos apresentam resultados intermediários que levam a algo essencial na pesquisa. Nesse caso, devem estar presentes nos materiais de replicação.
Sequências aleatórias	Para estudos de simulação ou estratégias de estimativa que usem randomização, os autores devem sempre incluir a sequência aleatória de números usada para produzir os resultados no artigo.
Diretórios e pastas	Os caminhos de diretórios devem ser facilmente identificáveis no script, facilitando a alteração pelos usuários. É recomendado que se coloque o diretório apenas no início, modificá-los ao longo do script pode gerar confusão.
Computing time	Os autores devem indicar (no código e no arquivo LeiaMe) se algum script específico demora muito tempo para calcular. Isso é importante para que o usuário não se depare com uma estimativa longa e acredite que esteja passando por algum problema.
Parallel Computing	Cada vez mais os autores estão utilizando parallel computing nos seus scripts a fim de acelerar o processo computacional. Infelizmente, isso não funciona em todos os computadores. Logo, devem indicar em que momento do script a estratégia foi utilizada.
Avisos e erros	Não é incomum encontrar scripts que produzam erros ou avisos – que geralmente estão relacionados ao uso inadequado de pacotes ou compatibilidade com o sistema operacional. Nesse caso, os autores devem indicar as razões pelas quais eles não são um motivo de preocupação.

Fonte: Alvarez et al (2018), p. 425-426. Tradução dos autores

As sugestões de Alvarez et al. (2018) são, definitivamente, um caminho para a produção de uma pesquisa mais transparente e replicável. Apesar disso, a quantidade de passos necessários e a falta de um direcionamento mais específico sobre como implementá-los acaba tornando-as um caminho trabalhoso. É nesse sentido que defendemos a adoção do protocolo TIER como uma ferramenta para a organização de materiais de replicação.

Acreditamos que a adoção e ampla divulgação do Protocolo TIER oferece um benefício claro à comunidade acadêmica: ao oferecer uma maneira de padronizar os materiais de replicação, acaba fazendo com que pesquisadores e diferentes áreas saibam, de forma intuitiva, o que devem encontrar em cada uma das pastas. Em outras palavras, a organização dos materiais de replicação se tornará mais orgânica para os pesquisadores e mais intuitivas para os replicadores. Nesse sentido, o protocolo TIER surge como um caminho menos tortuoso à adoção do padrão de replicabilidade.

Tais benefícios estão amplamente fincados nas três dimensões da replicabilidade, a saber: substantiva, pedagógica e transparente. Na dimensão substantiva, o padrão de replicação possibilita que os trabalhos possam ser falseados e, com isso, que a ciência avance mais rápido (PARANHOS et al, 2012; PARANHOS et al, 2013). De modo similar, a adoção do Protocolo TIER permitiu que os trabalhos dos alunos de Ball fossem avaliados de forma mais cautelosa e que os alunos recebessem comentários mais direcionados sobre os problemas enfrentados na produção do artigo. Por sua vez, para os pesquisadores que resolvem adotar o protocolo, esta estratégia de organização permite que seus trabalhos recebam melhores avaliações pelos seus pares e pelos avaliadores de periódicos especializados.

No que diz respeito à dimensão pedagógica, o padrão de replicação permite que os alunos sejam iniciados no mundo da pesquisa e da análise de dados ao terem contato com dados e problemas do mundo real (KING, 1995). Mais recentemente, os alunos de Ball resolvem exercícios *soup-to-nuts* (mais sobre eles abaixo), que levam os alunos a implementar o protocolo e exercer todo o processo de pesquisa empírica com dados estatísticos, desde a coleta dos dados à escrita do relatório com as análises estatísticas. Nesse mesmo sentido, o *hub* do Projeto TIER no Brasil, baseado em Recife, tem iniciado a produção de *workshops* para alunos de graduação em Ciência Política e Relações Internacionais com o objetivo de difundir a utilização do protocolo e iniciá-los na pesquisa empírica.

Por fim, na dimensão da transparência, o padrão de replicação permite que os procedimentos e dados sejam publicamente disponíveis para todos os pesquisadores que se interessem sobre o tema (KING, 1995; PARANHOS et al, 2013; FIGUEIREDO, 2019). O protocolo permite que a disponibilização das informações necessárias para a replicação seja humanamente inteligível e disponibilizada de modo com que qualquer pesquisador possa realizar a replicação dos resultados do estudo sem necessitar de ajuda do autor. Para os alunos do Ball, possibilitou que os resultados fossem replicados sem a presença dos alunos. Para a ciência, é a maneira das pesquisas, discussões e hipóteses conflitantes avançarem rumo à produção do conhecimento.

Nesse sentido, acreditamos que a adoção do protocolo pode ser uma ferramenta útil na caminhada para os autores das Ciências Sociais que desejam alcançar maiores níveis de transparência em suas pesquisas, auxiliando não só quem está conhecendo o valor da transparência agora, mas também fornecendo um mapa claro sobre como proceder para aqueles que já tinham o interesse, mas não sabem como. Na próxima seção, apresentamos um exercício *soup-to-nuts* de replicação que serve como uma introdução para aplicação do protocolo.

5. Exercício de Aplicação do Protocolo

Nesta seção, apresentaremos o primeiro exercício *soup-to-nuts* desenvolvido para exemplificação da aplicabilidade do protocolo (BALL, 2018). Os exercícios *soup-to-nuts* são originalmente desenvolvidos com o intuito de serem aplicados em turmas de Introdução à Estatística ou Introdução à Análise de Dados[17]. Nessas atividades, o intuito é levar os alunos a realizarem um processo completo de pesquisa empírica, desde o download de um banco de dados, até o tratamento, análise e entrega dos resultados. A descrição das etapas é detalhada, e originalmente é prevista para ser introduzida em sala de aula e concluída pelos alunos em casa. Nessa atividade, os alunos devem construir a sua hierarquia de pastas logo no início do projeto, trabalhando, portanto, com o Protocolo TIER ao longo de toda a jornada, uma vez que serão cobrados de realizar a entrega dos resultados com a documentação completa e organizada. Além de facilitar o trabalho do professor na correção dos exercícios, como já comentamos da experiência do Richard Ball, o protocolo ajuda aos alunos em seus fluxos de trabalho, configurando assim um exemplo poderoso das potencialidades dessa metodologia organizacional. Os exercícios de *soup-to-nuts* são, portanto, verdadeiros tutoriais para os alunos trabalharem com análise de dados de maneira transparente e replicável.

Neste artigo, utilizaremos uma versão encurtada do exercício "Consumo de Álcool em Universidades Americanas" como exemplificação do protocolo[18]. Recomendamos que dedique algum tempo para sua elaboração, de maneira que compreenda a maneira pela qual o protocolo contribui ao fluxo de trabalho e facilita a replicação de uma pesquisa. A versão completa e em português deste exercício você encontra na página do OSF[19], e a versão original você encontra no site do TIER[20]. O exercício resolvido, com modelo de relatório e de *scripts*, também estão disponíveis nos endereços indicados. A versão apresentada abaixo é para ser resolvida no *software* R.

Atividade de Replicação (Protocolo TIER): Consumo de Álcool em Universidades Americanas

Nesta atividade, você irá explorar se o consumo de bebidas alcoólicas é menor entre estudantes que vivem em dormitórios onde a prática é proibida em relação àqueles que moram em dormitórios onde não existe essa regra. Os dados utilizados são fruto de uma pesquisa conduzida em 2001 em universidades americanas, da *Harvard School of Public Health College Alcohol Study, 2001*. Essa atividade pedirá que você use os dados dessa pesquisa para criar alguns gráficos de barra e que responda a uma série de questões sobre a interpretação desses gráficos no que se refere aos padrões de consumo de álcool entre estudantes, comparando os que podem beber em seus dormitórios e os que não deveriam. A entrega deverá ser não somente o relatório com os gráficos e as respostas, mas todo o material utilizado para processamento e análise realizados nesta atividade, seguindo a documentação indicada.

I. PREPARATIVOS

IA. Crie uma hierarquia de pastas para guardar e organizar seu trabalho.

Crie uma pasta chamada de **ExercicioAlcool/.**

17 Ver: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility . Acesso em 20 de março de 2020.

18 Ball (2018) é o autor da versão original deste exercício, e Jenna Krall é creditada enquanto colaboradora. A tradução para o português e para a linguagem R foi realizada pelos membros do TIER hub Recife. Esta versão encurtada é de responsabilidade dos autores deste artigo, com autorização do autor da versão original.

19 Versão completa do exercício em Português: https://osf.io/mc2e9/?view_only=32e9fa640c83422ba7d6e9ab91e65743. Acesso em 06 de outubro de 2020.

20 Versão original completa: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility . Acesso em 10 de outubro de 2020.

Dentro de **ExercicioAlcool/**, crie as seguintes pastas e subpastas:

- **Dados/** (uma subpasta de **ExercicioAlcool/**)
  - **DadosOriginais/** (uma subpasta de **Dados/**)
  - **DadosAnalise** (uma subpasta **Dados/**)
- ***Scripts/*** (uma subpasta de **ExercicioAlcool/**)
- ***Graficos/*** (uma subpasta de **ExercicioAlcool/**)

Salve essa hierarquia de pastas no local que escolheu guardar o trabalho para esse exercício.

I.B Instale os pacotes

Dois pacotes serão úteis nessa atividade. Instale antes de iniciar o seu trabalho.

- *haven* (para ler o banco de dados no formato *.dta*)
- *tidyverse* (para facilitar os procedimentos de processamento e criação dos gráficos)

II. BAIXE E EXAMINE O BANCO DE DADOS E O *CODEBOOK* QUE VOCÊ UTILIZARÁ NESSE EXERCÍCIO.

II.A. Encontre e explore o site onde os dados para esse exercício estão disponíveis.

Os dados da pesquisa está arquivado no *Inter-University Consortium for Political and Social Research* (ICPSR).

Vá até o site do ICPSR (www.icpsr.umich.edu) e procure por *Harvard School of Public Health College Alcohol Study, 2001*.

Quando você tiver encontrado a página desse estudo, leia as informações fornecidas.

II.B. Baixe o banco de dados e o *codebook*.

- Antes de baixar o banco de dados, você precisará criar uma conta no site do ICPSR.

Uma vez feito isso, faça o login e vá para a página do estudo: https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4291

Clique no botão de *download*. Você verá um menu de formatos para escolher, incluindo SAS, SPSS, Stata ou ASCII. Nesse menu, selecione Stata (com o pacote *haven* o R irá ler o formato *.dta*).

O arquivo do banco será baixado numa pasta zipada, *ICPSR_04291-V2.zip*.

Quando você extrair os arquivos, verá que são vários documentos, incluindo uma subpasta chamada **DS0001**. Para esse exercício, você usará somente dois desses arquivos que vieram na pasta zipada e ambos estão na subpasta **DS0001**:

- *04291-0001-Data.dta*. Esse arquivo contém os dados da pesquisa em formato *.dta*. Esse arquivo será referido como "o banco de dados original" nesse exercício.
- *04291-0001-Codebook.pdf*. Esse é o *codebook* para a pesquisa utilizada. Será referido como "metadados do banco de dados original".

Você deve salvar cópias desses dois arquivos na pasta **DadosOriginais/**.

II. C. Examine o *codebook* e o banco de dados.

Examine o *codebook* e o banco de dados original para se familiarizar com as informações que eles contêm. Para esse exercício, você precisará utilizar cinco variáveis desse estudo. No banco de dados original, essas variáveis têm os nomes *A6, B8, B9, C13* e *G11*. Para cada uma dessas variáveis, analisando o *codebook*, responda as seguintes questões (as respostas a essas perguntas não precisarão estar no relatório final, mas são fundamentais para que você compreenda a natureza dos dados com que lidará):

- Como a variável é definida? O que ela representa?
- Qual a classe da variável (contínua, ordinal, categórica, discreta?)
- É qualitativa ou quantitativa?
    - Se for quantitativa, encontre a média, o desvio padrão, os três quartis, o valor mínimo e o valor máximo.
    - Se for qualitativa categórica, quantas categorias existem, o que elas representam e qual proporção das observações por categorias? Para cada variável categórica, decida se as categorias estão ordenadas ou não.

III. ESCREVENDO OS *SCRIPTS*

Convenções para o diretório de trabalho

Em todos os *scripts*, adote as seguintes convenções para o diretório de trabalho:

- Quando seus *scripts* forem executados, o diretório de trabalho do R deve ser definido para a pasta ***Scripts/***.
- Não escreva comandos em seus *scripts* que mudem o diretório de trabalho. Por exemplo, evite utilizar a função *setwd()*.
- Quando você precisar abrir ou salvar um arquivo em outra pasta que não ***Scripts/***, especifique a localização da pasta em questão utilizando um *relative directory path*.

Para esse exercício, você escreverá três *scripts*:

III.A. *tratamento.R*

- O propósito desse *script* é modificar o banco original (*04291-001-Data.dta*) para prepará-lo para a análise que você vai realizar.
    - As modificações incluem a criação de uma série de novas variáveis e a exclusão de algumas observações do banco.
- Salve o novo banco em um arquivo chamado *analise.RData*.
    - Esse arquivo será chamado de banco de análise.

Criando o *script*

- Abra o RStudio.
- Abra um novo *script* de R.
- No início do novo *script*, escreva um comentário indicando que o diretório de trabalho do R deve ser a pasta **Scripts/**. Por exemplo:

  # Quando esse *script* for executado, o diretório de trabalho deve ser a pasta **Scripts/**.
- Salve esse novo *script* na sua pasta **Scripts/**, com o nome *tratamento.R*.
    - Para evitar perder trabalho, salve o *scripts* constantemente enquanto o escreve.
- Carrege os pacotes *haven* e *tidyverse*
- Abra (importe) o banco *04291-0001-Data.dta*.
    - Use a função *read_dta()* do pacote *haven*.
    - Como seu diretório de trabalho é a pasta **Scripts/**, você precisa usar um *relative directory path* para especificar que o arquivo *04291-0001-Data.dta* deve ser importado de **DadosOriginais/**.
- Para essa atividade, nós queremos considerar apenas estudantes que vivem no campus ou em fraternidades/sororidades. Assim, exclua todos os casos os quais a residência do aluno seja "*Off*

campus house/apt" ou "*other*" ("fora de casa/apto no campus" ou "outro"). Também exclua todos os casos onde a variável representando a residência do aluno seja uma NA.

- Crie uma variável chamada *drunk*, que seja:
    - Igual a 0 se o estudante bebeu o suficiente para ficar bêbado menos do que três vezes nos últimos trinta dias.
    - Igual a 1 se o estudante bebeu o suficiente para ficar bêbado três ou mais vezes nos últimos trinta dias.
    - Igual a NA se a variável indicando quantas vezes o estudante bebeu o suficiente para ficar bêbado nos últimos trinta dias for NA.
- Crie uma variável chamada *hsdrunk*, que seja:
    - Igual a 0 se o estudante bebeu cinco ou mais bebidas seguidas em duas ou menos ocasiões no último ano do High School
    - Igual a 1 se o estudante bebeu cinco ou mais bebidas seguidas em três ou mais ocasiões no último ano do High School
    - Igual a NA se a variável indicando quantas vezes o estudante bebeu cinco ou mais bebidas seguidas no último ano do High School for uma NA
- Crie uma variável chamada *free*, que seja:
    - Igual a 0 se o estudante não vive num dormitório onde é proibido consumir bebida alcoólica (*alcohol-free housing*)
    - Igual a 1 se o estudante vive num dormitório onde é proibido consumir bebida alcoólica (*alcohol-free housing*)
    - Igual a NA se a variável indicando se o estudante vive ou não num dormitório onde é proibido consumir bebida alcoólica for NA.
- Crie uma variável chamada *volfree*, que seja:
    - Igual a 1 se *free*=1 (o estudante vive num dormitório onde é proibido ingerir bebida alcoólica) e o aluno requisitou morar num dormitório desse tipo
    - Igual a 0 se *free=1* (o estudante vive num dormitório onde é proibido ingerir bebida alcoólica) e o aluno foi designado a viver num dormitório desse tipo
    - Igual a NA em todos os outros casos
- Crie uma variável chamada *housing*, que seja:
    - Igual a 1 se *free*=0 (o estudante não vive num dormitório onde é proibido ingerir bebida alcoólica)
    - Igual a 2 se *free*=1 (o estudante vive num dormitório onde é proibido ingerir bebida alcoólica) e o aluno foi designado a viver num dormitório desse tipo
    - Igual a 3 se *free* = 1 (o estudante vive num dormitório onde é proibido ingerir bebida alcoólica) e todas as habitações do campus possuem a regra de proibição de consumo de bebida alcoólica.
    - Igual a 4 se *free* = 1 (o estudante vive num dormitório onde é proibido ingerir bebida alcoólica) e o aluno requisitou morar num dormitório desse tipo
    - Igual a NA em todos os outros casos
- Exclua todas as variáveis que não sejam as seis que você acabou de criar.
- Exclua todos os casos onde o valor de *hsdrunk, drunk* ou *housing* seja NA.
- Salve o novo banco de dados na pasta **DadosAnalise/,** nomeando o arquivo de *Analise.RData*.
    - Você salva o banco com a função *save()*.
    - Aqui também você precisará indicar um diretório de trabalho relativo para especificar a localização da pasta

DadosAnalise/.

Lembre-se de salvar tratamento.R na sua pasta ***Scripts/***.

III.B. *ApendiceDados.R*

O *script* ApendiceDados.R irá conter comandos que geram as informações que vão aparecer no Apêndice de Dados (mais informações abaixo).

Criando o *script*

- Abra o RStudio.
- Abra um novo *script* de R.
- Abra o arquivo do banco de análise, analise.RData (que, após ter escrito e rodado tratamento.R, deve estar salvo na pasta **DadosAnalise/**).
- Crie os seguintes resultados para cada uma das cinco variáveis no banco de análise:
    - O número de NAs e o número total de observações (incluindo NAs)
    - A distribuição da frequência e a distribuição da frequência percentual
    - Um gráfico de barra que apresente a distribuição da frequência percentual.

Lembre-se de escrever comentários explicando quais linhas produzem qual informação e gráfico. Cada linha que produzir informação ou gráfico que componha o Apêndice de Dados deve ser precedido por um comentário que descreva o que se espera que esse comando faça. Por exemplo, acima de uma linha que gere uma tabela mostrando a distribuição da frequência da variável *housing*, você escreve um comentário do tipo "A seguinte linha gera uma tabela mostrando a frequência da distribuição de *housing*".

Lembre-se de salvar ApendiceDados.R em sua pasta ***Scripts/***.

III.C. *analise.R*

O analise.R gera os gráficos de barra que constituem a principal análise que você fará para essa atividade, utilizando os dados do seu banco de dados de análise (analise.RData).

Criando o *script*

- Abra o RStudio.
- Abra um novo *script* de R.
- Abra o analise.RData (que, depois de você ter escrito e rodado tratamento.R, deve estar salvo na sua pasta **DadosAnalise/**).
- Crie um gráfico de barras, com duas barras, onde:
    - Uma barra represente estudantes vivendo em dormitórios onde é proibido consumir bebida alcoólica
    - Uma barra represente estudantes que não vivem em dormitórios onde é proibido consumir bebida alcoólica
    - A altura de cada barra é igual à proporção de estudantes (do grupo de estudantes que a barra representa) que beberam o suficiente para ficar bêbado três ou mais vezes nos últimos trinta dias.
    - Salve esse gráfico na pasta **Graficos/** (subpasta de **ExercicioAlcool/**), com o nome Figura1.jpg. (Use um diretório de trabalho relativo para especificar a localização da subpasta **Graficos/**)
- Crie um gráfico de barra, com duas barras, onde:
    - Uma barra represente estudantes que

vivem em dormitórios onde é proibido consumir bebida alcoólica porque eles requisitaram isso
    - Uma barra represente estudantes que vivem em dormitórios onde é proibido consumir bebida alcoólica porque eles foram designados
    - A altura de cada barra deve ser igual à proporção de estudantes (no grupo representado pelas barras) que beberam o suficiente para ficar bêbados três ou mais vezes nos últimos trinta dias
    - Salve esse gráfico na pasta **Graficos/** (subpasta da pasta **ExercicioAlcool/**), com o nome Figura2.jpg.
- Crie um gráfico de barrar, com quatro barras, onde:
    - Cada barra represente estudantes em uma das quatro categorias definidas na variável *housing*
    - A altura de cada barra é igual à proporção de estudantes (no grupo representado pelas barras) que beberam o suficiente para ficar bêbados três ou mais vezes nos últimos trinta dias
    - Salve esse gráfico na pasta **Graficos/** (subpasta da pasta **ExercicioAlcool/**), com o nome Figura3.jpg.
- Crie dois gráficos de barra (lado-a-lado), cada um possuindo duas barras, onde:
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em três ou mais ocasiões no último ano do *High School*
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em menos de três ocasiões em seu último ano de *High School*
    - E em cada um desses gráficos,
        - Uma barra representa estudantes que vivem em dormitórios onde o consumo de álcool é proibido
        - Uma barra representa estudantes que não vivem em dormitórios onde há essa proibição
        - A altura de cada barra é igual à proporção de estudantes (no grupo representado pelas barras) que beberam o suficiente para ficar bêbados três ou mais vezes nos últimos trinta dias
        - Salve esse gráfico na pasta **Graficos/** (subpasta da pasta **ExercicioAlcool/**), com o nome Figura4.jpg.
- Crie dois gráficos de barra (lado-a-lado), cada um possuindo duas barras, onde:
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em três ou mais ocasiões no último ano do *High School*
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em menos de três ocasiões em seu último ano de *High School*
    - E em cada um desses gráficos,
        - Uma barra represente estudantes que vivem em dormitórios onde é proibido o consumo de bebidas alcoólicas porque requisitaram
        - Uma barra represente estudantes que vivem em dormitórios onde é proibido o consumo de bebida alcoólica por que eles foram designa-

dos
    - A altura de cada barra é igual à proporção de estudantes (no grupo representado pelas barras) que beberam o suficiente para ficar bêbados três ou mais vezes nos últimos trinta dias
    - Salve esse gráfico na pasta **Graficos/** (subpasta da pasta **ExercicioAlcool/**), com o nome Figura5.jpg.
- Crie dois gráficos de barra (lado-a-lado), cada um possuindo quatro barras, onde:
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em três ou mais ocasiões no último ano do *High School*
    - Um dos gráficos represente apenas estudantes que tiveram cinco ou mais *drinks* em sequência em menos de três ocasiões em seu último ano do *High School*
    - E em cada um desses gráficos,
        - Cada barra represente estudantes em uma das quatro categorias definidas na variável *housing*
        - A altura de cada barra é igual à proporção de estudantes (no grupo representado pelas barras) que beberam o suficiente para ficar bêbados três ou mais vezes nos últimos trinta dias
        - Salve esse gráfico na pasta **Graficos/** (subpasta da pasta **ExercicioAlcool/**), com o nome Figura6.jpg.

Escreva comentários indicando quais linhas produziram que gráficos. Por exemplo, antes da linha que gera a Figura 4, escreva um comentário como "A seguinte linha de comando gera a Figura 4".

Lembre-se de salvar analise.R em sua pasta de ***Scripts&Relatorio/***.

IV. O RELATÓRIO

O relatório que você deve entregar para essa atividade consistirá na apresentação das Figuras que foram geradas e de respostas a uma série de perguntas que pedem que você interprete os gráficos de barra que você criou. Salve em *.pdf.*

Apresente as Figuras e escreva uma legenda abaixo de cada gráfico que indique a numeração e um título informativo. Por exemplo, uma legenda apropriada para a Figura 1 seria: "Figura 1: Taxa de Consumo Excessivo de Álcool em Dormitórios onde é proibido consumir álcool vs. Dormitórios onde não há essa proibição".

Questões

1) Descreva a amostra no arquivo BancoAnalise.dta que você criou. Qual a unidade de análise? Quantas observações existem? Descreva o método utilizado para criar a amostra e quais critérios foram considerados para incluir ou excluir grupos de indivíduos.

Note que essa pergunta é sobre seu banco limpo e processado, ao invés do banco original que você baixou do ICPSR. De qualquer forma, para responder, você terá que recorrer ao *codebook* do banco original.

2) Explique em suas palavras o que a Figura 1 apresenta. Era o que você esperava? Se não, como a Figura 1 é diferente do que você esperava?

3) Explique em suas palavras o que a Figura 2 apresenta. O que você vê na Figura 2 repercute de alguma forma o

que você viu na Figura 1? Explique.

4) Quantos indivíduos são representados na Figura 1? (Ou seja, quantas observações foram utilizadas para gerar a figura?). E quantos indivíduos estão representados na Figura 2? Se o número de indivíduos representados na Figura 1 e Figura 2 não é o mesmo, explique: alguns indivíduos apresentados na Figura 1 não estão representados na Figura 2? Se sim, que indivíduos foram esse? E alguns indivíduos apresentados na Figura 2 não foram representados na Figura 1? Se sim, quais?

5) Considerando a Figura 3, quais partes contêm informação que já foram apresentadas nas Figuras 1 e 2? Que partes apresentam nova informação? Resuma em suas palavras o que a Figura 3 apresenta.

6) Compare a Figura 4 com a Figura 1. Explique que novas informações são fornecidas pela Figura 4 e apresente uma interpretação dessas informações.

7) Compare a Figura 5 com a Figura 2. Explique que nova informação é fornecida pela Figura 4 e apresente uma interpretação.

8) Compare a Figura 6 com a Figura 3. Explique que nova informação é fornecida pela Figura 4 e apresente uma interpretação.

9) Das figuras que você criou, quais conclusões gerais podem ser retiradas sobre os fatores associados ao consumo de álcool entre estudantes universitários?

V. O APÊNDICE DE DADOS

O Apêndice de Dados deve consistir em múltiplas seções, uma para cada variável do banco de dados de análise. Salve também em *.pdf*.

Para cada variável, a seção deve fornecer:

- O nome da variável.
- O número de NAs no formato NA/Total (quantidade de NAs/número total de observações).
- Uma definição da variável.
- Possíveis valores da variável.
- Uma explicação do que cada valor da variável representa (codificação dos valores).
- Uma tabela apresentando a distribuição da frequência e a distribuição da frequência percentual. Se a variável categórica é ordenada, essa tabela deve também apresentar distribuição da frequência percentual acumulativa.
- Um gráfico de barra que ilustre a distribuição da frequência percentual.

VI. O ARQUIVO LEIA-ME

O arquivo Leia-Me é um documento que apresenta informações sobre os arquivos eletrônicos que você organizou para documentar o trabalho desta análise. Deve consistir em duas seções principais:

> A primeira seção deve apresentar um mapa de todos os arquivos incluídos na documentação da replicação, assim como as pastas e subpastas onde estão localizados.
>
> A segunda seção deve fornecer instruções passo-a-passo explicando como utilizar a documentação para (i) replicar todo o processamento de dados necessário para transformar o banco original no banco de análise, (ii) gerar todos os resultados que você apresenta no Apêndice de Dados, e (iii) reproduzir os seis gráficos de barra que você criou.

Essas instruções devem ser escritas em português claro e devem ser detalhadas o suficiente para que alguém não familiarizado com esse exercício consiga seguir e reproduzir tudo que você fez. Salve em *.pdf*.

VII. DOCUMENTAÇÃO ELETRÔNICA

Uma vez finalizado o exercício, os arquivos utilizados e criados devem estar alocados na hierarquia de pastas da seguinte maneira:

ExercicioAlcool/
Leia-Me.pdf
Relatorio.pdf
Dados/ (subpasta de **ExercicioAlcool/**)
DadosOriginais (subpasta de **Dados/**)
04291-0001-Data.dta
04291-0001-Codebook.pdf
DadosAnalise (subpasta e **Dados/**)
BancoAnalise.dta
Scripts/ (subpasta de **ExercicioAlcool/**)
tratamento.R
ApendiceDados.R
analise.R
Graficos/ (subpasta de **ExercicioAlcool/**)
Figura1.jpg
Figura2. jpg
Figura3. jpg
Figura4. jpg
Figura5. jpg
Figura6. jpg

Não deixe nada além disso nessas pastas

- Sem nenhuma pasta além dessas.
- Sem nenhum arquivo além desses

Teste seus arquivos para saber se rodam

- Encontre um computador que não seja o que você realizou a atividade e copie para ele a pasta inteira, com todo seu conteúdo.
- Siga os seguintes passos, sem modificar a ordem das pastas ou os arquivos:
  - Inicie o R;
  - Defina ***Scripts/*** como seu working directory
  - Rode o tratamento.R e após isso, confira se o arquivo BancoAnalise.dta que você tem na pasta **DadosAnalise/** foi alterado por uma nova cópia criada recentemente (confira a data e hora de modificação do arquivo na pasta)
  - Rode o ApendiceDados.R e verifique se gera os resultados esperados. Confira também se os gráficos na pasta **Graficos/** foram atualizados por novas versões recentes.

Rode o analise.R e confira se os seis arquivos de gráficos localizados na pasta **Graficos/** foram atualizados.

6. Outras atividades do TIER

Com o intuito de promover a importância da transparência e de difundir os valores da ética na ciência, o Projeto TIER desenvolve algumas atividades e eventos de promoção de seu protocolo e de boas práticas científicas. Todas podem ser encontradas em maiores detalhes em seu endereço da web[21].O Projeto TIER promove workshops voltados a professores e pesquisadores universitários, tanto para ouvi-los sobre suas práticas em prol da transparência e da replicabilidade na ciência, quanto para apresentá-los às ideias mais recentes sobre o protocolo (e colher comentários e impressões)[22]. Esses workshops são ótimas oportunidades de cooperação interuniversidades e de cooperação internacional, em termos de boas práticas científicas, reunindo professores de distintos campos das ciências sociais e diferentes países.

Há ainda o programa de *Fellowship,* voltado a jovens professores pesquisadores engajados com ensino e treinamento de métodos quantitativos sob uma lente

21 Ver: https://www.projecttier.org/ . Acesso em 20 de março de 2020.

22 Ver: https://www.projecttier.org/fellowships-and-workshops/fall-2019-faculty-development-workshop/ . Acesso em 20 de março de 2020.

transparente e replicável. As bolsas valem por um ano acadêmico e buscam incentivar que o recipiente desenvolva e difunda métodos de ensino de pesquisa transparente – como exercícios *soup-to-nuts* e *workshops*. Além disso, o TIER promoveu também nos meses de fevereiro e março de 2020 uma série de conferências no formato de webcast. A cada semana, um convidado fazia uma fala sobre os seus mais recentes desenvolvimentos e ideias sobre transparência e replicabilidade. Dentre esses "líderes da transparência", estavam Gary King, Dorothy Bishop e Scott Long. As apresentações estão disponíveis no site do TIER[23].

7. Considerações Finais

Os princípios e as práticas de transparência têm se tornado o padrão-ouro da ciência empírica. Iniciativas para difundir essas ideias e apresentar ferramentas para aumentar a transparência nas pesquisas e na produção científica sobre o assunto tem se tornado cada vez mais relevantes. Materiais de replicação ganham força enquanto elementos que compõe uma pesquisa concluída, e suas divulgações e compartilhamentos têm sido um requisito para publicação em diversos periódicos especializados, seja em território americano e, embora ainda de forma incipiente, aqui no Brasil.

Nesse trabalho, tivemos o objetivo de discutir formas de aumentar a transparência de pesquisas empíricas nas Ciências Sociais, principalmente através da documentação dos materiais de replicação. Apresentamos, ao longo do artigo, o surgimento do Protocolo TIER e seu enquadramento nas três dimensões do padrão de replicação (substantiva, pedagógica e transparência), uma vez que o protocolo permite a continuidade da produção acadêmica e a movimentação da roda da ciência, possibilita que jovens pesquisadoras sejam iniciadas no mundo na pesquisa empírica através de dados estatísticos e análise de dados, além de aumentar o nível de transparência dos trabalhos que o adotam. Além disso, apresentamos a estrutura organizacional do protocolo, os arquivos comportados e a hierarquia de pastas que pressupõem a documentação sugerida pelos criadores. Apresentamos, ainda, uma versão encurtada de um exercício *soup-to-nuts* para implementação do protocolo, ideal para aplicação em turmas de introdução a estatística e análise de dados. Esperamos que os alunos que sejam introduzidos a essas análises já dentro de um paradigma transparente e replicável tenham não só mais facilidade em aprender e em desenvolvê-las, mas que também carreguem consigo esses valores de boa ciência.

Com isso, visamos indicar um caminho menos tortuoso para que a pesquisadora alcance níveis mais altos de transparência em sua pesquisa. Além disso, esperamos ter difundido e apresentado de forma simples e didática o protocolo. Dessa maneira, esperamos contribuir com a agenda de pesquisa sobre transparência e replicabilidade, que vem crescendo na Ciência Social como um todo, e na Ciência Política brasileira em especial.

23 Ver: https://www.projecttier.org/fellowships-and-workshops/weekly-webcast-leaders-research-transparency/ . Acesso em 20 de março de 2020.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVAREZ, R. M; KEY, E. M; NÚÑEZ, L. (2018). "Research replication: Practical considerations". *PS: Political Science & Politics,* 51(2), 422-426.

BALL, R.; MEDEIROS, N. (2012). "Teaching Integrity in Empirical Research: A Protocol for Documenting Data Management and Analysis". *The Journal of Economic Education*, 43(2), 182-189.

BALL, R. (2018). Animal House in Alcohol-Free Dorms? TIER Project. Disponível em: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility. Acesso em 06 de outubro de 2020.

CASEY, K; GLENNERSTER, R; MIGUEL, E. (2012). "Reshaping institutions: Evidence on aid impacts using a preanalysis plan". *The Quarterly Journal of Economics*, 127(4), 1755-1812.

CHRISTENSEN, G.; SODERBERG, C. (2016). "Manual of best practices in transparent social science research". *Retrieved April*, 2, 2017.

CHRISTENSEN, G.; MIGUEL, E. (2018). "Transparency, reproducibility, and the credibility of economics research". *Journal of Economic Literature*, 56(3), 920-80.

DE LA GUARDIA, F. H.; STURDY, J. (2019). "Best Practices for Transparent, Reproducible, and Ethical Research". *IDB Technical Note* ; 1635.

FIGUEIREDO FILHO, D.; LINS, R.; DOMINGOS, A.; JANZ, N.; SILVA, L. (2019). "Seven Reasons Why: A User's Guide to Transparency and Reproducibility". *Brazilian Political Science Review*, 13(2).

JANZ, N. (2016). "Bringing the gold standard into the classroom: replication in university teaching". *International Studies Perspectives*, 17(4), 392-407.

KEY, Ellen M. (2016). "How are we doing? Data access and replication in political science. PS: Political Science & Politics, 2016, 49.2: 268-272

KING, G. (1995). "Replication, replication". *PS: Political Science & Politics*, 28(3), 444-452.

MIGUEL, E., CAMERER, C., CASEY, K., COHEN, J., ESTERLING, K. M., GERBER, A., LAITIN, D. (2014). "Promoting transparency in social science research". *Science*, 343(6166), 30-31.

MUNAFÒ, M. R., NOSEK, B. A., BISHOP, D. V., BUTTON, K. S., CHAMBERS, C. D., DU SERT, N. P., IOANNIDIS, J. P. (2017). "A manifesto for reproducible science". *Nature human behaviour*, 1(1), 1-9.

OLKEN, B. A. (2015). "Promises and perils of pre-analysis plans". *Journal of Economic Perspectives,* 29(3), 61-80.

OSINSKI, L. (2019). "PROOF course Open Science: The new default in science, 01-10-2019". *Presentation at Eindhoven University of Technology*. Disponível em: <https://www.projecttier.org/tier-classroom/course-materials/#modal230>. Acesso em 23 mar. 2020.

PARANHOS, R., FIGUEIREDO FILHO, D. B., da Rocha, E. C., da Silva Jr, J. A., & SANTOS, M. L. W. D. (2012). "Levando Gary King a sério: desenhos de pesquisa em Ciência Política". *Revista Eletrônica de Ciência Política*, 3(1-2).

PARANHOS, R., FIGUEIREDO FILHO, D. B., da ROCHA, E. C., CARMO, E. F. (2013). "A importância da replicabilidade na ciência política: o caso do SIGOBR". *Revista Política Hoje*, 22(2), 213-229.

PRZEWORSKI, A., SALOMON, F. (1988). "The Art of Writing Proposals: Some candid suggestions for Applicants to Social Science Research Council". *SSRC: New York*.

A map for transparency and replicability in empirical social science: the TIER Protocol

Amanda Domingos - Federal University of Pernambuco
Ian Rebouças Batista- Federal University of Pernambuco

Resumo

Transparência e replicabilidade são o padrão ouro da pesquisa científica. Mas, alcançar esse padrão pode ser um caminho longo e trabalhoso. Nesse artigo, apresentamos uma maneira de aumentar a transparência de sua pesquisa através da documentação. Argumentamos que o Protocolo TIER é uma importante ferramenta de transparência e de auxílio na organização de materiais para replicação. Trata-se de uma metodologia de documentação que compila os principais materiais produzidos numa pesquisa empírica social. Além disso, destacamos a adequação do protocolo nas três dimensões do padrão de replicação (substantiva, pedagógica e transparente), uma vez que contribui para o avanço da ciência, introduz jovens graduandas à análise de dados de uma forma prática e permite a transparência quanto aos procedimentos empregados na pesquisa. Por fim, discutimos atividades relacionadas ao Projeto TIER que contribuem com os objetivos do protocolo. Especificamente, apresentamos um exercício de introdução à análise de dados a ser realizado no R, onde a documentação do protocolo é requerida e sua aplicação, portanto, ilustrada.

Palavras-chave: Transparência; Replicabilidade; Documentação; Protocolo TIER

Abstract

Transparency and replicability are the gold standard in empirical research. But reaching that standard can be a long and tortuous journey. In this article, we present a way to increase the transparency of research through documentation. We argue that the TIER Protocol is an essential tool for transparency and assistance in organizing replication materials. It is a documentation methodology that compiles the main materials produced in empirical social research. Also, we highlight the adequacy of the protocol in the three dimensions of the replication standard (substantive, pedagogical, and transparent) – since it contributes to the advancement of science, introduces young undergraduate students to data analysis practically, and allows transparency of research procedures. Finally, we discuss activities related to the TIER Project that contribute to the objectives of the protocol. Specifically, we present a data analysis introduction exercise for R, where the protocol documentation is required and its application, therefore, exemplified.

Keywords: Replicability; Documentation; TIER Protocol

1. Introduction

Download the dataset, clean it, write scripts, program functions, execute commands and analyze results. If these are routine steps in your research, a protocol will help you organize and publish your results in a transparent and replicable way: the TIER Protocol.

Principles and practices of transparency and replicability have drawn attention in the scientific community, both in national (PARANHOS, 2012; PARANHOS et al., 2014; FIGUEIREDO et al., 2019; AVELINO and DESPOSATO, 2018) and international publications (CHRISTENSEN, 2016; MUNAFÓ et al., 2017; CHRISTENSEN and MIGUEL, 2018; STOCKEMER, KOEHLER E LENZ, 2018; DE LA GUARDIA and STUDY, 2019). Also, such practices have been encouraged through inter-university initiatives such as the *Berkeley Initiative for Transparency in the Social Sciences* (BITSS) and the *Teaching Integrity in Empirical Research* (TIER), which promote courses, workshops, and events focused on the exchange of experiences and ideas in favor of transparency.

Here we understand transparency regarding the procedures adopted in empirical research, such as the availability of the materials used and the indication of the steps taken until reaching the results. In addition to the benefits – such as allowing public access to methodological choices and enabling that the work is replicated and more rigorously evaluated –, transparency also allows: (i) more accurate hypothesis testing, (ii) collaboration with other researchers with similar interests, and (iii) increase in the author's credibility. We argue that a transparent researcher, who documents their work as it is done, has an easier process in elaborating her research. Rigorous documentation brings benefits in terms of scientific reputation, given that it contributes to good practices in science and generates practical benefits, such as increased citations (FIGUEIREDO et al., 2019).

Given this, some instruments have been used to disseminate and implement transparency and replicability in the Social Sciences, with three leading practices: disclosure plan, preregistration, and open data and materials (MIGUEL et al., 2014). This study aims to focus on the third kind. In this paper, we present and suggest adopting the TIER Protocol, a methodology for the documentation of the files used throughout the development of the research (dataset, scripts, outputs), and later accessibility of these materials for replication. We believe that adopting a specific rule regarding the organization of research files increases transparency and favors a study's replicability.

The paper is divided as follows. In the next section, we present the gold-standard of scientific research regarding the transparency of procedures and replicability of results and point out why the reader should aim towards these values and practices. Next, we present the TIER Protocol, highlighting its essential aspects. Then, we emphasize the benefits and incentives of adopting a documentation and disclosure strategy to make research transparent. Lastly, we list some activities developed by the TIER project that corroborate and seek to disseminate good practices in transparency and replicability in the Social Sciences and present an introductory exercise to the protocol. The conclusion wraps up the paper.

2. Why a transparency protocol?

Among the characteristics of good scientific research are "conceptual innovation, methodological rigor, and rich, substantive content" (PRZEWORSKI and SALOMON, 1998, p. 1). In the Social Sciences, however,

two other elements have been added to these over the past few decades: transparency and replicability, held as the gold-standard of empirical research (KING, 1995; PARANHOS et al., 2014; JANZ, 2016). There are many benefits of transparency both for the discipline as a whole and the researcher in particular. Adoption of transparent practices allows the wheel of science to spin faster by allowing more precise hypothesis testing, facilitating collaboration with other researchers with similar interests, and increasing the credibility of the research, as well as removing the usual attention given to the statistical significance and increasing the relevance of research *per se* (DE LA GUARDIA and STUDY, 2019).

Regarding benefits to authors and journals, Figueiredo et al. (2019) present seven reasons to adopt transparent procedures in empirical research in the Social Sciences: (I) to prevent honest mistakes or even frauds in empirical analyses; (II) to making writing scientific papers easier, since it allows testing other hypotheses with the same dataset; (III) transparent procedures allow reviewers to have access to data and scripts and therefore enables them to give more precise recommendations on how to improve the paper; (IV) by ensuring systematization of raw and analyzed data, it enables the author's scientific work to continue; (V) it contributes to data analysis learning, by allowing students to have contact with data and statistical procedures used in real life; and (VII) it increases the impact of academic work since empirical evidence shows that papers that adopts transparent procedures and practices are cited twice as often as those who are not transparent.

On the other hand, there are many costs in not adopting transparency and replicability procedures. Studies that do not maintain public procedures may contribute to lowering confidence in their results, and it may be an obstacle to publishing in high-impact factor journals that request replication materials – diminishing the study's range (MUNAFO et al., 2017; DE LA GUARDIA and STUDY, 2019). Thus, someone could argue that adopting transparency principles and practices are too much work. But it is worth remembering that even the gold-standard of replication does not require that an individual reproduce the results presented in books or scientific papers, only recommends that is the author makes available as much information as needed so that a researcher who desires to can reproduce them (KING, 1995).

Despite that, the practice of transparency is not the rule. Roughly 67% of papers published in the *American Political Science Review* between 2013 and 2014 are not replicable (KEY, 2016). In Brazil, only 5% of papers published between 2012 and 2016 in Social Sciences journals were apt for complete replication (AVELINO e DESPOSATO, 2018)[12]. The standard is still obscurity, not transparency. Consequently, it is not unusual to find researchers who have not made public the procedures and practices that led them to their scientific products' conclusions. We ask the dear reader: how many times, in the production of your empirical research, did you make an effort to remember the paths that took you to the result you found? More than that, what have you done to be transparent? If you thought a bit about that, this article could help you. Over the next sections, we will suggest an organization protocol to facilitate transparent practices and discuss the documentation process of empirical research.

3. The TIER Protocol

The TIER Protocol is a documentation methodology for replication materials in empirical research. Simply put, the protocol gives a general indication for the

1 The report was a result of the project "*Research Transparency in Brazilian Political and Social Science: A First Look*", by the authors along with the *Berkeley Initiative for Transparency in the Social Sciences* (BITSS). For more information, visit: < https://www.bitss.org/projects/research-transparency-in-brazilian-political-and-social-science-a-first-look/>. Last acessed in: 20 March 2010.

2 The final report is publicly avaliable at: <https://osf.io/f279z/>. Last acessed in 20 March 2020.

organization of files used in a study, the construction of folders, and what information must be made available in each of them. Initially, the protocol was developed to deal with a problem in the classroom. The papers handed in by the students of Introduction to Statistics, discipline taught by Richard Ball for economics undergraduate students at Haverford College (USA), had a similar issue to those in scientific papers published in great academic journals: they were not transparent. When handing in the discipline's final papers with wrong results, it was practically impossible for the professor to understand at what point in the data analysis the procedures were incorrect conducted, which hindered adequate grading by the professor and, consequently, the student's learning process.

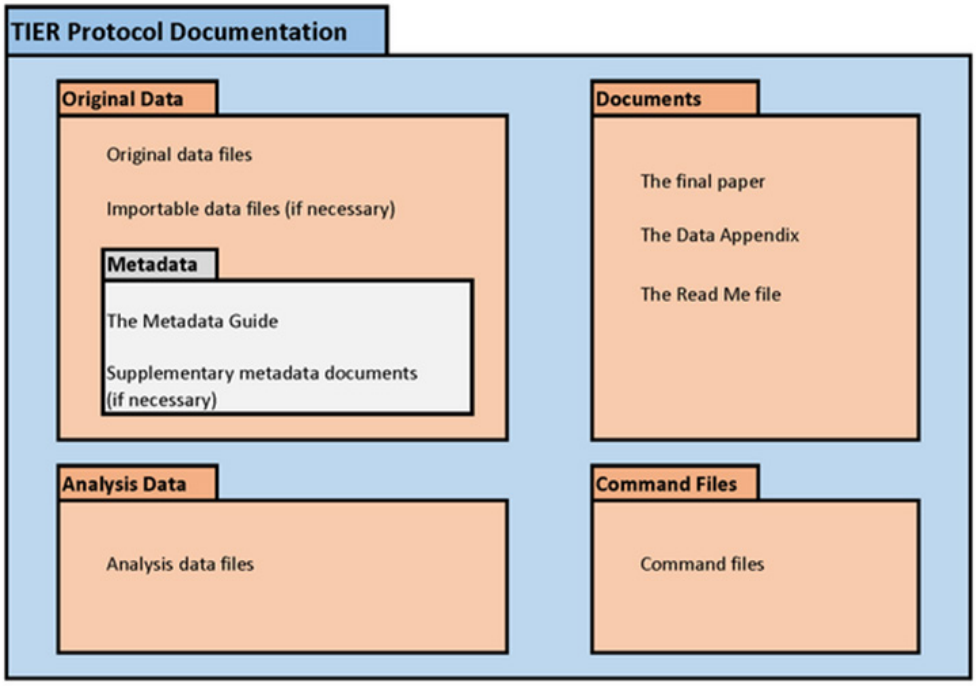
To guarantee a standard in the final class project delivery, enabling the professor to grade efficiently (having access to scripts and data in an intelligible and organized way), Ball elaborated, with the College's librarian, Norm Medeiros, the TIER Protocol (BALL and MEDEIROS, 2012). The protocol consists of a specific organization and hierarchy of folders that separate information on a given research (raw data, analyzed data, scripts, etc.) and facilitates the replication of the findings. Despite its pedagogical origin, the TIER Protocol goes beyond the classroom. Several researchers in the field of Social Sciences can use the protocol to increase the level of transparency in their research. Thus, we argue that adopting the TIER Protocol results in the collaboration of the three dimensions in the replication standard: substantive, pedagogical, and transparency (PARANHOS et al., 2013).

The TIER Protocol has a general guideline, as illustrated in Figure 1, and has versions that improve the organization and operationalization of the work's replication, as with version 4.0 in Figure 2. We will discuss the two arrangements next, but some procedural clarifications of TIER's organizational methodology purposes help in understanding.

The folders that will comprise the work's replication material must be created at the beginning of the research. They will stay with the researcher throughout the whole process: treatment of the original dataset, running analyses, writing the paper and appendices. The researcher must save and work with the files in the proper folders and, in their scripts or command file, referencing the working directories according to the folder hierarchy, making it essential that the folder hierarchy is maintained throughout the project. For instance, in the script where you treat the original dataset and create the set for analysis, you must load it from the folder *OriginalData* and save the treated dataset in the folder *AnalysisData* (Figure 1). The idea is that when using software such as R, the script has the lines needed both to open the original data and also save the analysis data in the correct folders[3].

3 As a replicability measure, ideally one would adopt a relative working directory. For more information, see: https://ssc.wisc.edu/sscc/pubs/R_intro/book/1-10-paths-and-working-directories.html. Last accessed on: 07 October 2020.

Figure 1: Overall view of the documentation suggested by the TIER Protocol



Source: Project TIER.4

But what folder organization should you adopt since the structures proposed in Figures 1 and 2 are different? The protocol's main goal is to provide a guideline for organizing and making accessible all digital material used during empirical research. It is important that this guideline is coherent and , if not intuitive, justified and presented. Therefore, the folders' exact configuration and files are up to you. Figure 1 is the most generic example, from which four other versions were elaborated, and Figure 2 is the most recent formatting suggested by TIER Project (4.0). Therefore, once the protocol's argument on research documentation and transparency convinces the researcher/reader, she is encouraged to organize her material based on a protocol that works for her – that is, in the hierarchy of folders that she finds convenient.

For individual protocols to work for any future replicator, a few basic guidelines that compound TIER's message are required. The first one is creating ReadMe files (usually in .txt, to facilitate reading regardless of the operational system), which will guide the replication activity from the beginning. One of the premises of a transparent replication is that this kind of files are the first to be read because, in addition to presenting and explaining the adopted folder configuration that the replicator will find when reading, these are the text documents that say which files to read first and what will be the sequence of the replication work. Therefore, it is useful that it appears in the first folder, as shown in Figure 2. The elaboration of a clear and direct ReadMe file shows the researcher's concern with the future replication of their findings. Thus, it must offer a complete overview of how the folders are organized and replication steps.

For instance, for a paper using the protocol inspired by what is shown in Figure 2, the ReadMe file might contain a text similar to this: "Initially, go to the *Treatment* subfolder, located in the *Scripts* folder, and open the

4 Original template used by the TIER Project. Available at: https://www.projecttier.org/tier-protocol/specifications-3-0/#overview-of-the-documentation. Last accessed on 22 January 2021.

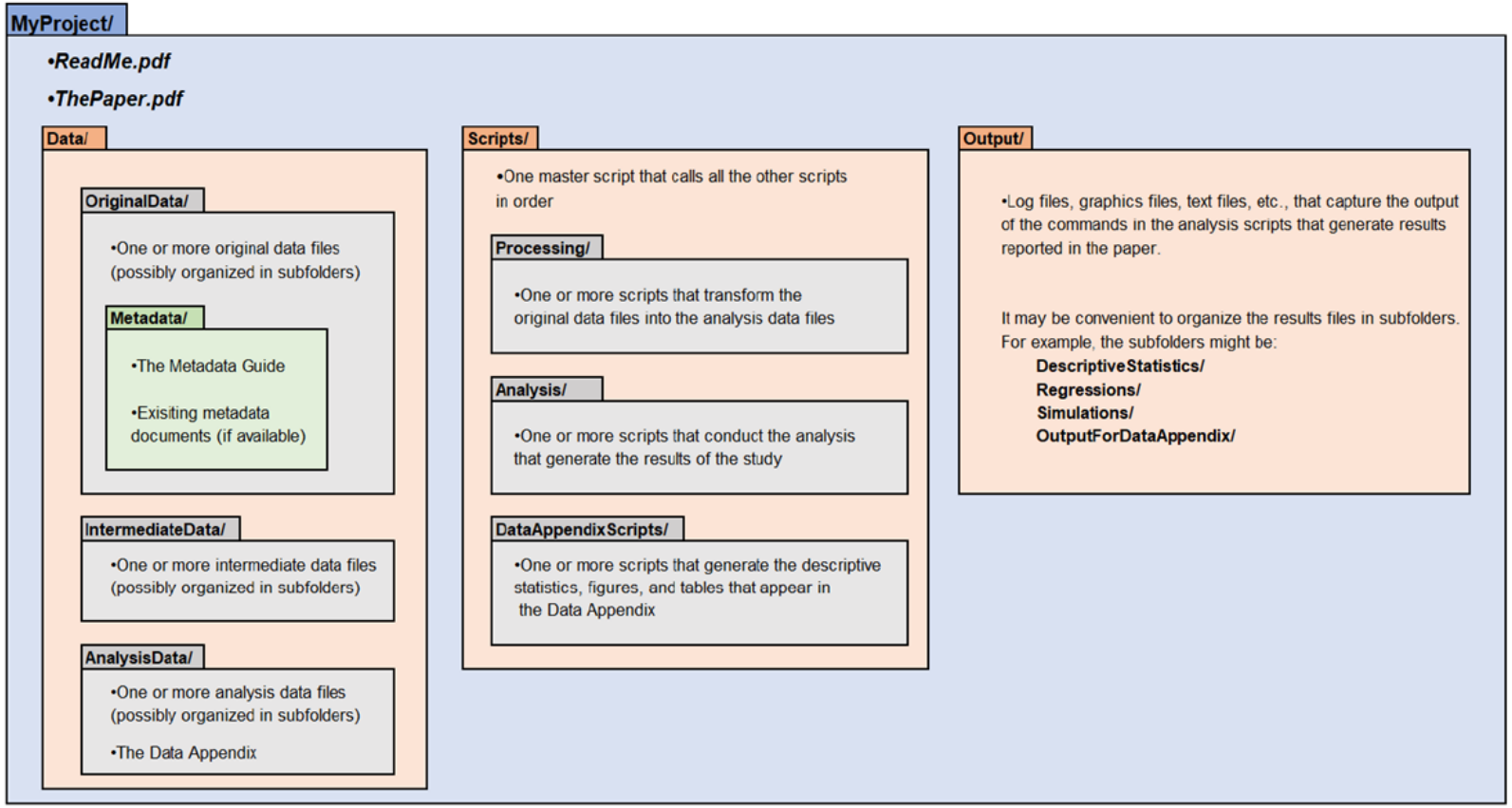
Treatment.R script. With it, you will treat Database. dta, which is the *OriginalData* subfolder, within the *Data* folder, and will generate NewDatabase.dta, which will be saved in the subfolder *AnalysisData*. Next, also in the *Scripts* folder, go to the *Analysis* subfolder and open the Analysis.R script. With it, you will generate Graphs 1, 2, and 3 and you will save it in the *Results* folder". ReadMe files can also contain information on the software used for data analysis, web addresses where the original data were downloaded, or even the researcher's contact information.

Like ReadMe, but with a different function, the TIER Protocol considers it fundamental that the researcher creates a Data Appendix (or codebook). In this document, the variables present in the analysis dataset are described in detail, including their type, how they were measured, and how they were built (if it was already in the original dataset or if it was manipulated by the researcher while treating the data). Descriptive statistics on each variable can also be included. Also, in this document, the research gives more in-depth information on the software's version and which statistical packages were used.

Another TIER Protocol's basic guideline on any of its versions, and that must also be present in a possible personal protocol inspired by it, is the separation of files by types. In both Figures 1 and 2, there are folders for datasets, scripts, and documents. We will analyze these figures in more detail now.

Figure 1, the protocol's general guideline, has four folders in total: *OriginalData; AnalysisData; CommandFiles; Documents*. Within the *OriginalData* folder, we find the original dataset, directly downloaded from whatever source, the converted dataset, should the researcher have had to convert the original downloaded dataset (note that, even then, it is crucial to keep the dataset in the original format), and a subfolder called *Metadata*, where a Metadata Guide (a Data Appendix for the original dataset, including the bibliographical reference, the DOI identifier, the date of its download, and the web address), and supplementary documents from this original dataset (if there should be any). In the *AnalysisData* folder, there is the treated dataset(s), the one with the researcher will conduct the analysis. We found the final paper, the Data Appendix, and the ReadMe file in the Documents folder. Lastly, in the *CommandFiles* folder, there will be all the scripts used in the research, from the treatment of the original dataset up to the analyses carried out.

Figure 2: TIER Protocol version 4.0 (2020)



Source: Project TIER. Tradução dos autores

Figure 2 is based on a preliminary elaboration of the protocol's version 4.0, presented in November 2019. In the main folder, there are immediately two files, the final paper, the ReadMe file, and three folders: *Data*, *Scripts*, and *Results*. There are three subfolders in the Data folder: *OriginalData* (with the *Metadata* folder contained within); *IntermediateData*, where there will be datasets that might have been converted to a different format; and *AnalysisData*, where the *DataAppendix* will also be present. In the *Scripts* folder, there is a "master" script, which combines all the work's scripts and three subfolders: *Treatment*, *Analysis*, and *DataAppendixScript*, each folder containing the script for the corresponding step. Lastly, in the *Results* folder, all the files that were created throughout the analysis, such as tables, graphs, and figures, which can also be subdivided into folders.

Both versions of the protocol share the same ideal of research transparency and documentation, although adopting different folder organizations and hierarchies. Understanding the protocol's goals and the logic behind documentation, the researcher interested in transparency and replicability can, if she does not adopt one of the main versions, create her own TIER Protocol version. A version that makes sense to her and benefits her workflow while also being clear and accessible to anyone interested in replicating their research results.

After pointing out all the material following the TIER Protocol, availability must occur in the most accessible way possible. Since data transparency is a growing issue, we have at our disposal a range of digital resources that offer an adequate structure. The *Open Science Framework*[5] (OSF), *Dataverse*[6], and *Github*[7] are three examples of suitable repositories for sharing all work material, including the possibility of subdividing the folders of the protocol. In this paper, we consider a quantitative approach, given that the nature of the protocol benefits this type of work. But there are also digital repositories for papers with qualitative methods[8], such as the *Qualitative Data Repository*[9], the UK data archive[10], and the National Anthropological Archives[11].

4. Documentation, TIER Protocol, and the dimensions of the replication standard

In 1995, the year the "*verification/replication*"[12] issue was released, the reality of scientific production in the Social Sciences was different. The dossier proposed to researchers to make available more details on the empirical choices conducted in their research, in footnotes and methodological appendices. But these efforts were met with the limits imposed by journals and editors in publishing these details. On the other hand, availability and storage of information were a problem: the facility of hard drives and clouds did not exist; the digital archive related to publications from the *Inter-University Consortium for Political and Social Research* (ICPSR)[13], for example, would only be released sometime after the issue (ALVAREZ et al., 2018).

With the development of suitable online platforms and repositories, information disclosure has become eas-

5 See: https://osf.io . Last accessed on 20 March 2020.

6 See: https://dataverse.org/ . Last accessed on 20 March 2020.

7 See: https://github.com/ . Last accessed on 20 March 2020.

8 For more information on repositories that accept qualitative data, see: https://www.sesync.org/list-of-qualitative-data-repositories. Last accessed on 24 March 2020.

9 See: https://qdr.syr.edu/. Last accessed on 24 march 2020.

10 See: https://www.ukdataservice.ac.uk/manage-data. Last accessed on 24 March 2020.

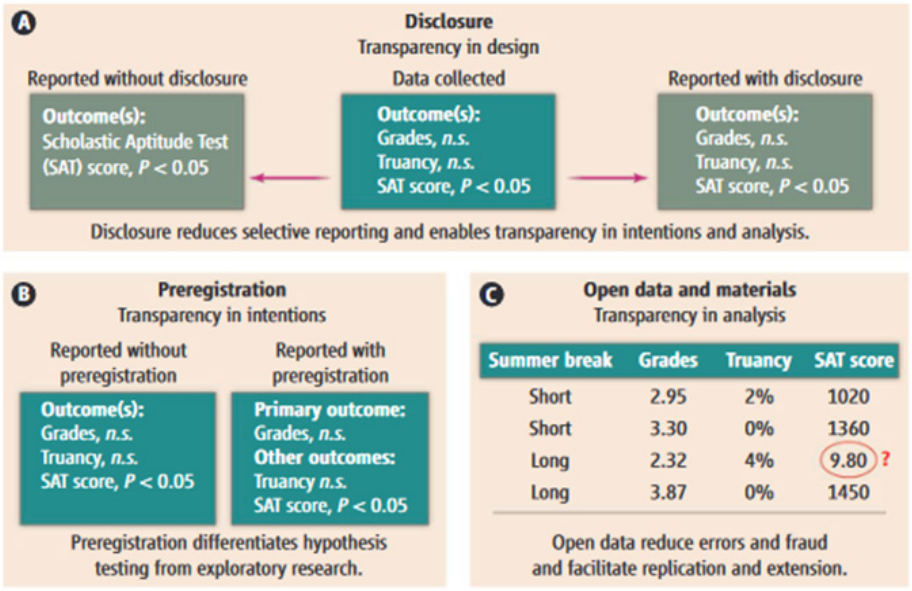
11 See: http://anthropology.si.edu/naa/search.html. Last accessed on 24 March 2020.

12 The dossier is available at: < https://www.cambridge.org/core/journals/ps-political-science-and-politics/issue/DAEAAA8412CD56791EAFCC75082A4C15>. Last accessed on 24 March 2020.

13 Available at: <www.icpsr.umich.edu/icpsrweb/deposit/pra/index.jsp>. Last accessed on 24 March 2020.

ier every day. Consequently, several efforts have been undertaken to disseminate and implement the principles and practices of transparency and replicability in the Social Sciences. These efforts have been concentrated on three leading practices: disclosure plans, preregistration, and open data and materials (MIGUEL et al., 2014), as illustrated in Figure 3.

Figure 3: Three mechanisms to increase transparency in scientific reports (Demonstration with the research question "do shorter summer breaks improve educational outcomes?", where "n.s" means P > 0,05.)



Source: Miguel et al. (2014), p.30.

Disclosure plans arise to argue for the production of systematic reports that seek to ensure that researchers document and disclose the way they proceeded regarding data collection and analysis – this is a strategy widely used in medicine and implemented by some academic journals in Political Science and Psychology [14] (MIGUEL et al., 2014). Preregistration is a public document where the researcher specifies variables, data processing procedures, and analysis techniques (OLKEN, 2015). Preregistration aims to diminish publication bias, prevent the display of selective research results and reinforce its credibility (CASEY et al., 2012). Lastly, public data and materials enable other researchers to reproduce the research results, test remaining hypotheses and identify muddy results or even frauds (MIGUEL et al., 2014; MUFANO et al., 2017; CHRISTENSEN and MIGUEL, 2018).

The TIER Protocol fits into the third practice. More specifically, in the documentation that precedes data and replication materials' disclosure. The adoption of replicability policies by high impact factor journals pushes researchers to systematically document their research files. Despite demand, the materials made available are substandard. Stockemer, Koehler, and Lenz (2018) analyzed all the papers published in 2015 in the three main political behavior journals[15] and found that for 25% of them the code was so poorly organized that replication was impossible.

Well-organized replication materials enhance the expansion of related studies and enable more accessible

14 According to research by the authors, journals that have adopted policies towards this type of procedure are: *Journal of Experimental Political Science, Management Science,* and *Psychological Science.* See: Miguel et al. (2014), p. 30.

15 The authors analyzed all the papers present in the 2015 volumes of the following academic journals: *Electoral Studies*, *Party Politics*, and *Journal of Elections, Public Opinion & Parties*.

analyses (ALVAREZ et al., 2018). At the most basic level of preparation towards becoming a transparent researcher, it is recommended to adopt some sort of protocol for data organization to have reusable and humanly understandable data (FIGUEIREDO et al., 2019; OSINSKI, 2019).

Alvarez, Key, and Nuñez (2018) give some recommendations to increase data organization, clarity, and usage. The authors provide tips on how to archive materials related to the research in a useful and secure way, such as how to produce ReadMe files that are helpful for future replications, clues on how the files in replication materials should be named, as well as indicating how to report software (used or created by the author), scripts, and outputs generated by the research. Table 1 summarizes the authors' suggestions.

Table 1 - Recommendations for creating replication materials (Alvarez et al., 2018)

Processes	Recommendations
Archive	Replication materials must be stored in permanent files that will be available for long periods. Currently, the most trustworthy are those with institutional guarantees, whether in consortiums or universities; Authors must save their files so that they can be used in the future (e.g. comma-separated values file), rather than software-specific formats.
ReadMe files	The file must be sufficiently clear and detailed, but not overly. Some items must be included, such as (a) reference to the paper or associated publication; (b) a brief explanation of the files and the types of files included, e.g., original data, processed data, scripts, etc.; (c) an indication of the order the scripts should be executed; (d) a software list, software packages, and the operating system used to produce the results in the paper; (e) if uncommon extensions were used, the author must indicate how to proceed with those files.
File names	The files' names must be easy to comprehend and give information on their content. This is especially important for replication materials that include several scripts and data files. In case some files must be used in a certain order, they must be named in a way that includes the order (e.g., "1_ProcessedData, "2_Estimation", etc.).
User-created Software and Packages	When the author has created a statistical package, these must be (ideally) archived in a stable location, such as CRAN for R. If possible, the author should include a copy of the software or package in the replication materials.
Operating-system compatibility	Ideally, authors must ensure that their replication materials work in the most common operating systems. If it is not possible, they should indicate in which system they do work.
Scripts	(a) should contain a brief description of what they do, what they depend on, what packages are required, and what are the outputs; (b) lines without specification of what they do must be avoided, given that this indication helps in identifying specific areas of data management and processing and identifying possible errors or problems; (c) there should be good planning and recoding the data in the middle of the analysis should be avoided, unless extremely necessary.
Outputs	Replication materials should contain tables, graphs, and figures in the sequence that they appear in the paper.
Excluded outputs	Usually, replication materials have scripts or data that were not used in the paper or were linked to an online appendix. In this case, the author should differentiate them from the ones in the paper.
Intermediate outputs	Sometimes papers present intermediate outputs that lead to something essential in the research. In that case, they should be present in the replication materials.
Random sequences	For simulation or estimation strategies studies that use randomization, authors must always include the random sequence of numbers used to produce the results in the paper.
Directories and Folders	The directories' paths must be easily identifiable in the script, facilitating alteration by the users. It recommended that the directory is added only at the beginning, as modifying them throughout the script can create confusion.
Computing time	Authors must indicate (in the code and ReadMe file) if any script takes too long to calculate. This is important so that the user is not faced with a long estimation and believes they are having a problem.
Parallel Computing	More and more, authors are using parallel computing in their scripts to speed up the computational process. Unfortunately, this does not work on every computer. Therefore, they should indicate at which point in the script the strategy was used.
. Warnings and errors	It is not uncommon to find scripts that produce errors or warnings – that are usually unrelated to the inadequate use of packages or incompatibility with the operating system. In this case, the authors should indicate the reasons why these are no cause for concern.

Source: Alvarez et al. (2018), p. 425-426.

The suggestions by Alvarez et al. (2018) are, undoubtedly, a path towards producing more transparent and replicable research. Despite this, the number of steps required and the lack of more specific guidelines on implementing them make the work more difficult. Consequently, we defend adopting the TIER Protocol as a tool for the organization of replication materials.

We believe that the adoption and wide circulation of the TIER Protocol offers a clear benefit to the academic community. By providing a way to standardize replication materials, researchers in different fields intuitively know what they should find in each folder. In other words, the organization of replication materials will become more organic for researchers and more intuitive for replicators. Consequently, the TIER Protocol emerges as a less tortuous path to adopting the replicability standard.

These benefits are deeply rooted in the three dimensions of replicability: substantive, pedagogical, and transparent. In the substantive dimension, the replication standard allows research to be falsifiable and, thus, for science to move faster (PARANHOS et al., 2012; PARANHOS et al., 2013). Similarly, the TIER Protocol's adoption enabled the papers by Ball's students to be evaluated more cautiously and for students to received more precise comments on the problems faced in the production of the paper. For researchers who have decided to adopt the protocol, the organization strategy allows their papers to receive a better review by their peers, whether colleagues or journal reviewers.

Regarding the pedagogical dimension, the replication standard allows students to be initiated in the world of research and data analysis by having contact with real-world data and problems (KING, 1995). More recently, Ball's students solve *soup-to-nuts* exercises (more on that below), leading them to implement the protocol and go through the whole empirical research process with statistical data, from data collection to writing the report with statistical analyses. In like manner, the TIER hub in Brazil, based in Recife, has begun the production of workshops for undergraduate students in Political Science and International Relations, intending to disseminate the use of the protocol and get them initiated in empirical research.

Lastly, in the transparency dimension, the replication standard allows procedures and data to be publicly available to all researchers that are interested in the theme (KING, 1995; PARANHOS et al., 2013; FIGUEIREDO, 2019). The protocol enables that the availability of necessary information for replication is humanly intelligible and accessible so that any researcher can replicate the study's results without assistance from the author. For Ball's students, it enabled that the results were replicated without their presence. For science, it is the way that conflicting studies, discussions, and hypotheses advance towards producing knowledge.

Therefore, we believe that adopting the protocol can be a useful tool in the path of social scientists who aim to reach higher levels of transparency in their research, aiding not only those who are finding out the value of transparency now, but offering a clear map on how to proceed for those who were interested, but did not know how. In the next section, we present a *soup-to-nuts* exercise for replication, which introduces the protocol's application.

5. Exercise for the Application of the Protocol

In this section, we present the first *soup-to-nuts* exercise developed by the TIER initiative, to demonstrate the protocol's applicability (BALL, 2018).

The *soup-to-nuts* exercises were designed initially to be used in Introduction to Statistics of Introduction to Data Analysis classes[16]. In these activities, the aim is to lead the students to conduct a full empirical research process, from downloading a dataset through treatment, data analysis, and delivery of results. The steps' description is detailed and originally set to be reviewed in the classroom and then concluded by the students at home. In this activity, the students must build their folder hierarchy at the outset of the project working, therefore, with the TIER Protocol throughout the whole journey, given that they will be required to hand in the results with the complete and organized documentation. In addition to facilitating the professor's job when grading the exercises, as we have mentioned in Richard Ball's experience, the protocol helps students in the workflows, thus being a powerful example of this organizational methodology's potential. Therefore, the soup-to-nuts exercises are proper tutorials for students to work with data analysis in a transparent and replicable manner.

This paper will use a shortened version of the "Alcohol Consumption in American Universities" exercise as an example of the protocol[17]. We recommend that you dedicate some time to elaborate it, understand how the protocol contributes to the workflow, and facilitate the replication of a study. The complete version in Portuguese can be found in the OSF site[18] and the original version on the TIER website[19]. The solved exercise, with the report model and scripts are also available in the addresses indicated. The version presented below should be solved in R.

16 See: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility . Last accessed on 20 March 2020.

17 Ball (2018) is the author of this exercise's original version, and Jenna Krall is credited as collaborator. The translation to Portuguese was done by members of the TIER hub Recife. This shortened version is by the authors of this paper, with authorization by the author of the original.

18 Complete version of the exercise in Portuguese: https://osf.io/mc2e9/?view_only=32e9fa640c83422ba7d6e9ab91e65743. Last accessed on 06 October 2020.

19 Complete original: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility . Last accessed on 10 October 2020.

Replication Activity (TIER Protocol): Alcohol Consumption in American Universities

In this activity, you will explore if alcohol consumption is lower among students who live in housing where the practice is forbidden in relation to those who live in accommodation where the rule does not exist. The data comes from a study conducted in 2001 in American universities, by the *Harvard School of Public Health*. This activity will require you to use the data from that research to create some bar graphs and to answer a series of questions on their interpretation, comparing those who can drink in their housing and those who should not. Results to be handed in will not only be the report with graphs and answers but all the material used for processing and analysis conducted in this activity, following the documentation indicated.

I. PREPARATION

IA. Create a folder hierarchy to keep and organize your work.

Create a folder entitled **AlcoholExercise/.**

Inside **AlcoholExercise/**, create the following folders and subfolders:

- **Data/** (a subfolder of **AlcoholExercise/**)
  - **OriginalData/** (a subfolder of **Data/**)
  - **AnalysisData/** (uma subpasta **Data/**)
- ***Scripts/*** (a subfolder of **AlcoholExercise/**)
- ***Graphs/*** (a subfolder of **AlcoholExercise/**)

Save this folder hierarchy in the location you selected to save your work for this exercise.

I.B Install the packages

Two packages will be useful in this activity. Install them before beginning your work.

- *haven* (to read a dataset in *.dta* format)
- *tidyverse* (to facilitate processing procedures and creation of graphs)

II. DOWNLOAD AND EXAMINE THE DATABASE AND THE CODEBOOK YOU WILL USE IN THIS EXERCISE.

II.A. Find and explore the website where the data for this exercise are available.

The research data are archived at the *Inter-University Consortium for Political and Social Research* (ICPSR).

Go to the ICPSR's website (www.icpsr.umich.edu) and look for *Harvard School of Public Health College Alcohol Study, 2001*.

When you have found the page for this study, read the information provided.

II.B. Download the dataset and codebook.

- Before downloading the dataset, you will need to create an account at the ICPSR's website.

Once that is done, log in and go to the study's page: https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4291

Click on the download button. You will see a menu with formats to choose from, including SAS, SPSS, Stata, and ASCII. In this menu, select Stata (with the *haven* package, R will read the *.dta* format).

The dataset's file will be downloaded in a zipped folder, *ICPSR_04291-V2.zip*.

When you extract the files, you will see that there are several documents, including a subfolder named **DS0001**. For this exercise, you will only use two files that came in the zipped folder and both are in the **DS0001** subfolder:

- *04291-0001-Data.dta*. This file contains the research data in *.dta* format. This file will be referred to as "the original dataset" in this exercise.
- *04291-0001-Codebook.pdf*. This is the codebook for the study used. It will be referred to as "the metadata for the original dataset".

You must save copies of these two files in the folder **OriginalData/**.

II. C. Examine the codebook and the dataset.

Examine the original codebook and dataset to familiarize yourself with the information contained in them. For this exercise, you will need five variables for this study. In the original dataset, these variables are named *A6, B8, B9, C13,* and *G11*. For each of these variables, analyzing the codebook, answer the following questions (the answers to these questions do not need to be in the final report, but they are key for you to understand the nature of the data you will be dealing with):

- How is the variable defined? What does it represent?
- What class is the variable (continuous, ordinal, categorical, discrete)?
- Is it qualitative or quantitative?
    - If it is quantitative, find the mean, the standard-deviation, the three quartiles, and the minimum and maximum values.
    - If it is qualitative categorical, how many categories are there, what do they represent, and what is the proportion of observations per category? For each

categorical variable, decide if the categories are or not ranked.

III. WRITING THE SCRIPTS

Working directory conventions

For all *scripts*, adopt the following working directory conventions:

- When all your scripts are executed, R's working directory must be set to the **Scripts/** folder.
- Do not write commands in your scripts that change the working directory. For instance, avoid the *setwd()* function.
- When you need to open or save a file in folder other than **Scripts/**, specify the folder's location in question using a relative directory path.

For this exercise, you will write three scripts:

III.A. *treatment.R*

- The purpose of this script is to modify the original dataset (*04291-001-Data.dta*), to prepare it for the analysis you will conduct.
    - Modifications include creating a new variable set and excluding a few observations in the dataset.
- Save the new dataset in a file named *analysis.RData*.
    - This file will be called analysis dataset.

Creating the script

- Open RStudio.
- Open a new R script.
- At the top of the new script, write a comment indicating that R's working directory should be the **Scripts/** folder. For example:

    # When this script is executed, the working directory should be the **Scripts/** folder.
- Save this new script in your **Scripts/** folder with the name *treatment.R*.
    - To avoid losing work, constantly save the script as you write it.
- Load the *haven* and *tidyverse* packages.
- Open (import) the *04291-0001-Data.dta* dataset.
    - Use the *read_dta()* function in the *haven* package.
    - Since your working directory is the **Scripts/** folder, you need to use a relative directory path to specifiy that the *04291-0001-Data.dta* file must be imported from **Original-Data/**.
- For this activity, we want to consider only students who live on campus or in fraternities/sororities. Therefore, exclude all cases in which the student's residence is "*Off campus house/apt*" or "*other*". Also exclude all the cases where the variable representing the student's residence is NA.
- Create a variable called *drunk*, which is:
    - Equal to 0 if the student drank enough to get drunk fewer three times in the past 30 days.
    - Equal to 1 if the student drank enough to get drunk three times or more in the past 30 days.
    - Equal to NA if the variable indicating how often the student drank enough

to get drunk in the past 30 days is NA.

- Create a variable called *hsdrunk*, which is:
    - Equal to 0 if the student drank five or more drinks in a row on two or less occasions during their last year of High School
    - Equal to 1 if the student drank five or more drinks in a row on three or more occasions during their last year of High School
    - Equal to NA if the variable indicating how many times the student had five or more drinks in a row during their last year of High School is NA
- Create a variable called *free*, which is:
    - Equal to 0 if the student does not live in alcohol-free housing
    - Equal to 1 if the student lives in alcohol-free housing
    - Equal to NA if the variable indicating whether or not the student lives in alcohol-free housing is NA.
- Create a variable called *volfree*, which is:
    - Equal to 1 if free=1 (the student lives in alcohol-free housing) and the student requested to live in alcohol-free housing
    - Equal to 0 if free=1 (the student lives in alcohol-free housing) and the student was assigned to live in alcohol-free housing
    - Equal to NA in all other cases
- Create a variable called *housing*, which is:
    - Equal to 1 if free=0 (the student does not live in alcohol-free housing)
    - Equal to 2 if free=1 (the student lives in alcohol-free housing) and the student was assigned to live in alcohol-free housing
    - Equal to 3 if free=1 (the student lives in alcohol-free housing) and all on-campus housing is alcohol-free
    - Equal to 4 if free=1 (the student lives in alcohol-free housing) and the student requested to live in alcohol-free housing
    - Equal to NA in all other cases
- Exclude all variables other than the six you just created.
- Exclude all cases where the values for *hsdrunk, drunk,* or *housing* are NA.
- Save the new dataset in the **AnalysisData/** folder, naming the file *Analysis.RData*.
    - You save the dataset with the *save()* function.
    - Here you will also need to indicate a relative directory path to specify the AnalysisData/folder's location.

Remember to save treatment.R in your **Scripts/** folder.

III.B. *DataAppendix.R*

The DataAppendix.R script will contain commands that generate the information that will appear in the Data Appendix (more info below).

Creating the script

- Open RStudio.
- Open a new R script.
- Open the data analysis file, analysis.RData (which, after having written and run with with treatment.R, should be saved in the **AnalysisData/** folder).
- Generate the following outputs for each of the five variables in the analysis dataset:
    - The number of NAs and the total

number of observations (NAs included)
    - A frequency distribution and a percent frequency distribution
    - A bar graph that presents the percent frequency distribution.

Remember to write comments explaining which lines produce which information and graph. Each line that produces information or graph that is part of the Data Appendix must be preceded by a comment that describes what is expected from this command. For example, above a line that generates a table showing the frequency distribution of the *housing* variable, you should write a comment such as "The following line generates a table showing the frequency distribution of the *housing* variable".

Remember to save DataAppendix.R in your **Scripts/** folder.

III.C. *analysis.R*

The analysis.R script generates the bar graphs for your main analysis in this activity, using the data from your analysis dataset (analysis.RData).

Creating the script

- Open RStudio.
- Open a new R script.
- Open analysis.RData (which, after having written and run with with treatment.R, should be saved in the **AnalysisData/** folder).
- Generate a bar graph, with two bars, where:
    - One bar represents students living in alcohol-free housing
    - One bar represents students not alcohol-free housing
    - The height of each bar is equal to the proportion of students (of the student group that the bar represents) that drank enough to get drunk three or more times in the past 30 days.
    - Save this graph to the **Graphs/** folder (subfolder of **AlcoholExercise/**), with the name Figure1.jpg. (Use a relative directory path to specify the location of the **Graphs/** subfolder).
- Generate a bar graph, with two bars, where:
    - One bar represents students who live in alcohol-free housing because they requested it
    - One bar represents students who live alcohol free housing because they were assigned to it
    - the height of each bar is equal to the proportion of students (of the student group that the bar represents) who drank enough to get drunk three or more times in the last thirty days
    - Save this graph in the **Graphs/** folder (subfolder of **AlcoholExercise/**), with the name Figure2.jpg.
- Generate a bar graph, with four bars, where:
    - Each bar represents students in one of the four categories defined by the *housing* variable
    - The height of each bar is equal to the proportion of students (in the group represented by the bars) who drank enough to get drunk three times or more in the past 30 days
    - Save this graph in the **Graphs/** folder (subfolder of **AlcoholExercise/**), with the name Figure3.jpg.
- Generate two bar graphs (side-by-side), each with two bars, where:
    - One of the graphs represents only students who had five or more drinks in a

row on three or more occasions during their last year of High School
    - One of the graphs represents only students who had five or more drinks in a row on less than three occasions during their last year of High School
    - And, in each of these graphs,
        - One bar representing students who live in alcohol-free housing
        - One bar representing students who do not live in alcohol-free housing
        - The height of each bar is equal to the proportion of students (in the group represented by the bars) who drank enough to get drunk three or more times in the past 30 days
        - Save this graph in the **Graphs/** folder (subfolder of **AlcoholExercise/**), with the name Figure4.jpg.
- Generate two bar graphs (side-by-side), each with two bars, where:
    - One of the graphs represents only students who had five or more drinks in a row on three or more occasions during their last year of High School
    - One of the graphs represents only students who had five or more drinks in a row on less than three occasions during their last year of High School
    - And, in each of these graphs,
        - One bar that represents students who live in alcohol-free housing because they requested it
        - One bar that represents students who live in alcohol-free housing because they were assigned to it
        - The height of each bar is equal to the proportion of students (in the group represented by the bars) who drank enough to get drunk three or more times in the past 30 days
        - Save this graph in the **Graphs/** folder (subfolder of **AlcoholExercise/**), with the name Figure5.jpg.
- Generate two bar graphs (side-by-side), each with four bars, where:
    - One of the graphs represents only students who had five or more drinks in a row on three or more occasions during their last year of High School
    - One of the graphs represents only students who had five or more drinks in a row on less than three occasions during their last year of High School
    - And, in each of these graphs,
        - Each bar represents students in one of the four categories defined by the *housing* variable
        - The height of each bar is equal to the proportion of students (in the group represented by the bars) who drank enough to get drunk three or more times in the past 30 days
        - Save this graph in the **Graphs/** folder (a subfolder of **AlcoholExercise/**), with the name Figure6.jpg.

Write comments indicating which lines produced

which graphs. For instance, before the line that generates Figure 4, write a comment such as "The following command line generates Figure 4".

Remember to save analysis.R in your **Scripts&Relatorio/** folder.

IV. THE REPORT

The report you must hand in for this activity will consist of presenting the figures generated and the answers to a series of questions that ask you to interpret the bar graphs you created. Save to *.pdf*.

Present the Figures and write a caption under each graph that indicates the number and an informative title. For instance, an appropriate caption for Figure 1 would be: "Figure 1: Rates of Heavy Drinking in Alcohol-Free vs. Not Alcohol-Free Housing".

Questions

1) Describe the sample in the AnalysisDataset.dta file that you created. What is the unit of analysis? How many observations are there? Describe the method used to create the sample and what criteria were considered to include or exclude groups of individuals.

Note that this question is about your cleaned and processed dataset, rather than the original dataset that you downloaded from ICPSR. However, to answer, you will need to turn to the original dataset's codebook.

2) Explain in your own words what Figure 1 shows. Was it what you expected? If not, how is Figure 1 different from what you expected?

3) Explain in your own words what Figure 2 shows. Does what you see in Figure 2 somehow clarifies what you saw in Figure 2? Explain.

4) How many individuals are represented in Figure 1? (That is, how many observations were used to generate the figure?). And how many individuals are represented in Figure 2? If the number of individuals represented in Figures 1 and 2 is not the same, explain: are some individuals shown in Figure 1 not shown in Figure 2? If so, which individuals were these? And are some individuals shown in Figure 2 not present in Figure 1? If so, which ones?

5) Considering Figure 3, which sections contain information already presented in Figures 1 and 2? Which sections show new information? Summarize in your words what Figure 3 presents.

6) Compare Figure 4 to Figure 1. Explain which new information is given by Figure 4 and give an interpretation for that information.

7) Compare Figure 5 to Figure 2. Explain which new information is given by Figure 5 and interpret it.

8) Compare Figure 6 to Figure 3. Explain which new information is given by Figure 6 and interpret it.

9) Of the figures you created, what general conclusions can be drawn regarding the factors associated with alcohol consumption among college students?

V. THE DATA APPENDIX

The Data Appendix should consist of multiple sections, one for each variable of the analysis dataset. Save it to *.pdf* as well.

For each variable, the section must have:

- The name of the variable.
- The number of NAs in the NA/Total format (number of NAs/total number of observations).
- A definition of the variable.

- Possible values for the variable.
- An explanation of what each value for the variable represents (codification of values).
- A table presenting the distributions of the frequency and percent frequency. If the categorical variable is ranked, that table must also present the cumulative percent frequency distribution.
- A bar graph that illustrates the distribution of the percentage frequency.

VI. THE READ ME FILE

The Read Me file is a document that gives information on the electronic files that you organized to document the work in this analysis. It should consist of two main sections:

> The first section should present a map of all the files included in the replication documentation and the folders and folder where they are.
>
> The second section should offer step-by-step instructions, explaining how to the documentation to (i) replicate all the necessary data processing to transform the original dataset into the analysis dataset, (ii) generate all the output you present in the Data Appendix, and (iii) reproduce the six bar graphs that you created.

These instructions must be written in clear language and detailed enough for someone unfamiliar with this exercise to follow and reproduce everything you have done. Save to *.pdf*.

VII. ELECTRONIC DOCUMENTATION

Once the exercise is finished, the files used and created must be allocated in the folder hierarchy in the following manner:

AlcoholExercise/
ReadMe.pdf
Report.pdf
Data/ (subfolder of **AlcoholExercise/**)
OriginalData (subfolder of **Data/**)
04291-0001-Data.dta
04291-0001-Codebook.pdf
AnalysisData (subfolder of **Data/**)
AnalysisDataset.dta
Scripts/ (subfolder of **AlcoholExercise/**)
treatment.R
DataAppendix.R
analysis.R
Graphs/ (subfolder of **AlcoholExercise/**)
Figure1.jpg
Figure2. jpg
Figure3. jpg
Figure4. jpg
Figure5. jpg
Figure6. jpg

Do not leave anything else to these folders

- No folders in addition to these.
- No files in addition to these.

Test your files to see if they run

- Find a computer other than the one where you conducted this activity and copy the whole folder into it, with all its contents.
- Go through the following steps, without changing the order of the folders or files:
  - Start R;
  - Define **Scripts/** as your working directory
  - Run treatment.R and, after that, check if the file AnalysisDataset.dta that you have in the **AnalysisData/** folder was altered by a new copy recently created (check the date and

time of modification of the file in the folder)

- Run DataAppendix.R and check if it generates the expected outputs. Also check of the graphs in the **Graphs/** folder were updated by new, recent versions.

Run analysis.R and check if the six graph files located in the **Graphs/** folder were updated.

6. Other TIER activities

Intending to promote the importance of transparency and disseminate the values of ethics in science, the TIER Project develops some promotion activities and events for its protocol and good scientific practices. All details can be found at their web address[20]. The TIER Project promotes workshops for faculty members and researchers to debate practices towards transparency and replicability in science and present to them the most recent ideas on the protocol (and gather comments and impressions)[21]. These workshops are excellent interuniversity and international cooperation opportunities for good scientific practices, bringing together researchers from different Social Sciences and different countries.

There is also the Fellowship program for young professors and researchers engaged in quantitative methods teaching and training with a transparent and replicable view. The scholarships last one academic year and seek to encourage the recipient to develop and disseminate transparent research teaching methods – such as *soup-to-nuts* exercises and workshops. Also, in February and March 2020, TIER promoted a series of conferences in webcast format. Each week, a guest would speak on their most recent developments and ideas on transparency and replicability. Among these "transparency leaders" were Gary King, Dorothy Bishop, and Scott Long. The presentations are available on TIER's website[22].

7. Conclusions

The principles and practices of transparency have become the gold-standard of empirical science. Initiatives to disseminate these ideas and present tools to increase transparency in research and scientific production have become more and more relevant. Replication materials gained strength as elements that compose finished research, and its publication and sharing have become a requirement for publication in several specialized journals, whether in the United States or, incipiently, in Brazil.

In this paper, we aimed to discuss ways to increase transparency in empirical research in the Social Sciences, mainly through documentation of replication materials. We presented, throughout the paper, the development of the TIER Protocol and its framing in the three dimensions of the replication standard (substantive, pedagogical, and transparency), since the protocol enables young researchers to be initiated in the world of empirical research through statistical data and data analysis, as well as increases the level of transparency of the works that adopt it. In addition, we presented the organizational structure of the protocol, the files included, and the folder hierarchy that follows the documentation suggested by the creators. We also presented a shortened version of a *soup-to-nuts* exercise to implement the protocol, ideal for application in Introduction to Statistics and Data Analysis classes. We hope

20 See: https://www.projecttier.org/ . Last accessed on 20 March 2020.

21 See: https://www.projecttier.org/fellowships-and-workshops/fall-2019-faculty-development-workshop/ . Last accessed on 20 March 2020.

22 See: https://www.projecttier.org/fellowships-and-workshops/weekly-webcast-leaders-research-transparency/ . Last accessed on 20 March 2020.

that the students who are introduced to these analyses already within a transparency and replicability paradigm have an easier time learning and developing them and carrying with them these values of good science.

Thus, we set about indicating a less tortuous path so that researchers can reach higher levels of transparency in their production. Also, we hope to have disseminated and presented the protocol in a simple and didactic manner. In this way, we hope to contribute with the transparency and replicability agenda, which has been growing in the Social Sciences as a whole, and in Brazilian Political Science in particular.

BIBLIOGRAPHICAL REFERENCES

ALVAREZ, R. M; KEY, E. M; NÚÑEZ, L. (2018). "Research replication: Practical considerations". *PS: Political Science & Politics,* 51(2), 422-426.

BALL, R.; MEDEIROS, N. (2012). "Teaching Integrity in Empirical Research: A Protocol for Documenting Data Management and Analysis". *The Journal of Economic Education*, 43(2), 182-189.

BALL, R. (2018). Animal House in Alcohol-Free Dorms? TIER Project. Disponível em: https://www.projecttier.org/tier-classroom/soup-nuts-exercises/#shorter-exercises-for-teaching-transparency-and-reproducibility. Acesso em 06 de outubro de 2020.

CASEY, K; GLENNERSTER, R; MIGUEL, E. (2012). "Reshaping institutions: Evidence on aid impacts using a preanalysis plan". *The Quarterly Journal of Economics*, 127(4), 1755-1812.

CHRISTENSEN, G.; SODERBERG, C. (2016). "Manual of best practices in transparent social science research". *Retrieved April*, 2, 2017.

CHRISTENSEN, G.; MIGUEL, E. (2018). "Transparency, reproducibility, and the credibility of economics research". *Journal of Economic Literature*, 56(3), 920-80.

DE LA GUARDIA, F. H.; STURDY, J. (2019). "Best Practices for Transparent, Reproducible, and Ethical Research". *IDB Technical Note* ; 1635.

FIGUEIREDO FILHO, D.; LINS, R.; DOMINGOS, A.; JANZ, N.; SILVA, L. (2019). "Seven Reasons Why: A User's Guide to Transparency and Reproducibility". *Brazilian Political Science Review*, 13(2).

JANZ, N. (2016). "Bringing the gold standard into the classroom: replication in university teaching". *International Studies Perspectives*, 17(4), 392-407.

KEY, Ellen M. (2016). "How are we doing? Data access and replication in political science. PS: Political Science & Politics, 2016, 49.2: 268-272

KING, G. (1995). "Replication, replication". *PS: Political Science & Politics*, 28(3), 444-452.

MIGUEL, E., CAMERER, C., CASEY, K., COHEN, J., ESTERLING, K. M., GERBER, A., LAITIN, D. (2014). "Promoting transparency in social science research". *Science*, 343(6166), 30-31.

MUNAFÒ, M. R., NOSEK, B. A., BISHOP, D. V., BUTTON, K. S., CHAMBERS, C. D., DU SERT, N. P., IOANNIDIS, J. P. (2017). "A manifesto for reproducible science". *Nature human behaviour*, 1(1), 1-9.

OLKEN, B. A. (2015). "Promises and perils of pre-analysis plans". *Journal of Economic Perspectives,* 29(3), 61-80.

OSINSKI, L. (2019). "PROOF course Open Science: The new default in science, 01-10-2019". *Presentation at Eindhoven University of Technology*. Disponível em: <https://www.projecttier.org/tier-classroom/course-materials/#modal230>. Acesso em 23 mar. 2020.

PARANHOS, R., FIGUEIREDO FILHO, D. B., da Rocha, E. C., da Silva Jr, J. A., & SANTOS, M. L. W. D. (2012). "Levando Gary King a sério: desenhos de pesquisa em Ciência Política". *Revista Eletrônica de Ciência Política*, 3(1-2).

PARANHOS, R., FIGUEIREDO FILHO, D. B., da ROCHA, E. C., CARMO, E. F. (2013). "A importância da replicabilidade na ciência política: o caso do SIGOBR". *Revista Política Hoje*, 22(2), 213-229.

PRZEWORSKI, A., SALOMON, F. (1988). "The Art of Writing Proposals: Some candid suggestions for Applicants to Social Science Research Council". *SSRC: New York*.

Uma Introdução à Bibliometria e Cientometria: o caso do presidencialismo latino-americano

Amanda Sangalli - UFPR [1]
Samira Kauchakje - UDESC, UFPR [2]

Resumo

Introdução: O que são, para que servem e qual a aplicabilidade da bibliometria e da cientometria na ciência política? Nesse artigo apresentamos suas diferenças, vantagens e limitações a partir da literatura e da prática de aplicação. **Método**: Por meio das coleções bibliográficas Web of Science e SciELO Citation Index, foram selecionados 244 artigos científicos com enfoque analítico no presidencialismo de países latino-americanos e construídas redes bibliométricas com o VOSviewer. **Resultados:** A aplicabilidade foi demonstrada com a medição de coautoria, cocitação, acoplamento bibliográfico e coocorrência de palavras. Permitiram identificar, respectivamente, a internacionalização da agenda de estudos, as influências científicas fundamentais da temática, as frentes de pesquisa e os conceitos debatidos atualmente. **Discussão:** As principais vantagens da metodologia são: i) confiabilidade: por sua transparência e facilidade de replicação; e ii) amplitude: pela capacidade de analisar múltiplas dimensões da ciência. É necessário atentar-se que o método apresenta uma aproximação da realidade, mas esta tem limitações por se restringir ao "viés" da indexação existente nas plataformas, o que acarreta o número proporcionalmente maior de artigos publicados em países de língua inglesa com incentivos à pesquisa e às publicações científicas. É cabível manipular outras coleções bibliográficas ou inserir textos manualmente para maior precisão. Uma dificuldade enfrentada foi a manipulação de metadado, que nem sempre estão padronizados e podem variar entre diferentes coleções.

Palavras-chave: Bibliometria; cientometria; indicadores bibliométricos; presidencialismo; ciência política.

1 Graduada em Ciências Sociais (PUCPR, UNSAM). Mestranda em Ciência Política pela Universidade Federal do Paraná (UFPR). Pesquisadora do Laboratório de análise do Campo Científico (LaCC/UFPR).

2 Cientista política. Graduada em Ciências Sociais/área Ciência Política (UFPR); Mestra em Ciência Política (UFPR); Doutora em Ciência Política (UFSCAR); Doutora em Educação (UNICAMP). Professora no departamento de Governança Pública da Universidade do Estado de Santa Catarina (UDESC) e professora colaboradora no Programa de Pós-Graduação em Ciência Política da Universidade Federal do Paraná (UFPR).

Abstract

Introduction: What are they, what are they for and what is the applicability of bibliometrics and scientometrics in political science? In this article we present its differences, advantages and limitations from the literature and the application practice. **Method:** Through the Web of Science and SciELO Citation Index bibliographic collections, 244 scientific articles were selected with an analytical focus on the presidentialism of Latin American countries and built bibliometric networks with VOSviewer. **Results:** The applicability was demonstrated with the measurement of co-authorship, co-citation, bibliographic coupling and co-occurrence of words. They allowed to identify, respectively, the internationalization of the study agenda, the fundamental scientific influences of the theme, the research fronts and the concepts currently debated. **Discussion:** The main advantages of the methodology are: i) reliability: due to its transparency and ease of replication; and ii) breadth: due to the ability to analyze multiple dimensions of science. It is necessary to note that the method presents an approximation of reality, but it has limitations because it is restricted to the "bias" of indexing existing on the platforms, which results in the proportionally greater number of articles published in English-speaking countries with incentives for research and scientific publications. It is possible to manipulate other bibliographic collections or to insert texts manually for greater precision. One difficulty faced was the manipulation of metadata, which is not always standardized and can vary between different collections.

Keywords: Bibliometrics; scientometrics; bibliometric indicators, presidentialism; political science.

1. Introdução

O crescimento exponencial da produção científica mundial tornou inviável aos pesquisadores conhecê-la em sua totalidade. Ainda que os artigos científicos tenham o papel de apresentar as inovações de um determinado campo do conhecimento e que sejam um veículo rápido de comunicação entre cientistas, o volume de informações registradas dificulta o processo de sistematização da literatura e de compreensão dos avanços científicos e tecnológicos (GRANOVSKI, 2001). A relação entre tempo de trabalho e quantidade de informação, requer que o(a) pesquisador(a) priorize procedimentos automatizados de compilação de bibliografias. Deste modo, viabiliza mapear o conhecimento e direcionar a leitura ao hierarquizar documentos segundo sua relevância na comunidade ou sua atualidade.

A bibliometria e a cientometria são procedimentos dedicados a mensurar o desenvolvimento da ciência – seja de uma área do conhecimento, disciplina ou temática. São especialmente utilizadas por instituições de pesquisa públicas e privadas como meio de avaliação e de financiamento das produções científicas. Essa não é a única utilidade, pesquisadores interessados na evolução ou estruturação da ciência, isto é, que se dedicam a compreender sua trajetória sócio-histórica, igualmente operam tais estratégias metodológicas (SILVA, BIANCHI, 2001). Também pontuamos a usabilidade para desenho de carreiras, pois permite identificar quais são os principais centros de produção sobre determinado objeto de pesquisa, facilitando o encontro de parcerias e o planejamento de quais caminhos percorrer ao longo da carreira. Ambas possibilitam conhecer, não apenas o que já foi produzido sobre um assunto em décadas passadas, como também, quais são os temas emergentes da literatura e as lacunas de pesquisa.

O artigo está dividido em três seções mais a conclusão. A primeira seção é uma breve revisão metodológica. Apresentamos o que são, para que servem e quais as diferenças da bibliometria e cientometria, bem como as definições e aplicações de alguns indicadores bibliométricos. Na segunda seção, aplicamos a metodologia ao estudo de um tema de interesse da ciência política: o presidencialismo latino-americano. Explicamos os procedimentos para a construção de um banco de dados confiável e replicável, além de quais são os protocolos que uma pesquisa do tipo precisa seguir. Em terceiro momento, discutimos os resultados obtidos com este estudo, dando ênfase a aplicabilidade do método, por isso apresentamos uma análise suscinta do tema. Na conclusão, discutimos os obstáculos enfrentados nessa pesquisa, com recomendações e alertas para estudos futuros.

2. Bibliometria e cientometria: o que são e qual a diferença?

O campo da Ciência da Informação (CI) teve início no século XIX, a partir de uma aproximação das áreas de Sociologia da Ciência e História Social da Ciência, mas somente no século XX é que os estudos de métrica da informação foram sistematizados. A partir da demanda de entender a atividade científica como um fenômeno social (HAYASHI, 2012), é que se originaram os conceitos clássicos de bibliometria e cientometria. Há uma vasta discussão teórica a respeito desses conceitos e de suas aplicações, dificultando a delimitação dos espaços desses subcampos. Não sendo o ideal para este artigo exaurir esse embate histórico3, faremos uma discussão teórica fundamentada nas definições de Tague-Sutcliffe (1992) e de Van Raan (1997).

3 Para aprofundamentos no debate histórico ver a revisão bibliográfica de Sancho (1990) e o debate conceitual de Vanti (2002) e Hayashi (2012)

Quadro 1. Bibliometria e cientometria: diferenciação de propósito, objetos de estudo, variáveis, medidas e objetivos.

Campo de estudos	Embasamento teórico	Objetos de estudo	Variáveis	Medidas	Objetivos
Bibliometria	É um campo de estudos que desenvolve modelos e medidas matemáticas para mensurar quantitativamente a produção, disseminação e uso de informação registrada. Sua aplicação fornece uma base sólida para tomada de decisão e previsão (TAGUE-SUTCLIFFE, 1992).	Registros impressos, livros, documentos, revistas, artigos científicos, autores e usuários (VANTI, 2005).	Número de acessos (circulação) e de citações, frequência de palavras, extensão de frases (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998).	Ranking, frequência, distribuição (MACIAS-CHAPULA, 1998)	Alocar recursos: pessoas, tempo, dinheiro; compreender a produção e divulgação do conhecimento (adaptado de VANTI, 2002)
Cientometria	A análise cientométrica relaciona aspectos quantitativos a contextos sociais e econômicos visando interpretar os avanços do desenvolvimento científico. Se baseia não apenas em métodos matemáticos, como também sociológicos e comportamentais (VAN RAAN, 1997).	Áreas do conhecimento, campos científicos e tecnológicos, profissionais de um mesmo campo de atuação, disciplinas, assuntos, patentes, dissertações e teses (VANTI, 2005).	Fatores que diferenciam as disciplinas e os subcampos disciplinares; como os/as cientistas se comunicam (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998).	Análise de conjuntos; análise de correspondência; coocorrência de termos, expressões, palavras-chave; citação e cocitação de documentos, autores, instituições, países; acoplamento bibliográfico (adaptado de VANTI, 2002)	Identificar domínios de interesse, onde os estudos sobre determinado assunto estão concentrados; compreender como e o quanto os cientistas se comunicam (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998)

Fonte: elaboração das autoras a partir da literatura.

Em sentido teórico, podemos definir a bibliometria como um campo de estudos que se dedica ao desenvolvimento de técnicas para sistematizar informações bibliográficas. É comum a confusão na interpretação do conceito, sendo considerada pela literatura, ora metodologia, ora técnica. Conforme Gänzel (2003), há três grupos interessados na aplicação da bibliometria: os bibliometristas, os interessados em informação científica e os responsáveis pela gestão e elaboração de políticas científicas. O primeiro grupo, realiza pesquisas de desenvolvimento metodológico no campo da CI. Ou seja, consideram-na parte fundamental na execução de uma pesquisa. O segundo grupo é multidisciplinar, composto por pesquisadores de diferentes áreas, e busca empregar técnicas da bibliometria – indicadores bibliométricos –, para informar-se sobre determinado assunto. É nesse grupo que este artigo se encaixa ao tratar de conhecer e mapear o tema de presidencialismo latino-americano. O terceiro grupo também utiliza indicadores, mas se difere do grupo anterior em objetivo, pois tem o propósito de avaliar a ciência e pensar políticas de desenvolvimento. De acordo com essa diferenciação entre grupos podemos presumir que, em sentido prático, a bibliometria é uma metodologia para cientistas da informação, ao mesmo tempo que é um conjunto de técnicas para pesquisadores de outras áreas do conhecimento.

Enquanto a bibliometria é fundamentalmente quantitativa, aderir a cientometria como método é adotar estratégias metodológicas tanto no estilo quantitativo, como qualitativo, para a coleta e interpretação de dados. Para além dos indicadores bibliométricos, outros métodos aplicados na análise cientométrica são: o uso da estatística, modelos de rede sociológica, survey psicológico e outros métodos de entrevistas (VAN RAAN, 1997). Por se preocupar com a dinâmica da ciência, sua atividade social, circulação e consumo (SANTOS, KOBASHI, 2009), a filosofia e sociologia da ciência e a linguística, muitas vezes, também estão presentes nesse tipo de análise (VAN RAAN, 1997).

2.1 Indicadores bibliométricos

À luz da teoria apresentada, assumimos que a bibliometria é a análise quantitativa dos dados de bibliografia e a cientometria a análise das relações do campo científico, havendo um ponto de convergência entre elas que é o estudo de documentos bibliográficos e o emprego de indicadores bibliométricos.

Os indicadores são aplicados a casos com um grande volume de metadados bibliográficos. Eles podem ser utilizados para análise de desempenho, medir a quantidade de produções de um grupo de pesquisadores, periódicos, instituições ou países, também para mapeamento científico e compreensão da dinâmica de comunicação num determinado campo do conhecimento (ZUPIC, ČATER, 2015). Existem variados indicadores que podem ser aplicados em análises de bibliografias (SANCHO, 1990)4, os selecionados para esse exercício metodológico foram as técnicas de cocitação, acoplamento bibliográfico, coautoria e co-word – apresentadas no quadro 2 com suas funções. Se entendidas suas limitações e aplicados segundo suas funcionalidades específicas, eles são úteis para compreender o processo de criação, reprodução e difusão da ciência (MUGNAINI, JANNUZZI, QUONIAM, 2004).

4 Sancho (1990) descreve a usabilidade dos seguintes indicadores: número e distribuição de publicações; produtividade dos autores; colaboração nas publicações; conexões entre documentos e autores; referências ou citações conjuntas entre documentos; número de citações (impacto dos trabalhos); indicador de impacto das revistas; índice de imediatismo (rapidez com que são citados os documentos); influência das revistas; associações temáticas e coocorrência de termos.

Quadro 2. Descrição, vantagens e desvantagens dos indicadores bibliométricos selecionados para o estudo de caso.

Indicador	Unidade de análise	Unidade de observação	Objetivo	Medida	Explicação	Vantagem	Desvantagem
Cocitação	Documentos	Citação conjunta	Conectar documentos com base na aparição conjunta de referências em um dado documento.	Relevância	Serve para identificar os alicerces científicos. Encontra as referências fundamentais de um determinado campo do conhecimento. É uma sofisticação da técnica de citação, pois além de estimar a influência no campo (número de citações), conecta as unidades de acordo com suas semelhanças.	Filtra os trabalhos mais importantes de uma dada área.	Não é ideal para o mapeamento de pesquisas recentes. As citações levam tempo para serem acumuladas, portanto, as novas publicações não podem ser conectadas diretamente, mas por meio de clusters e base de conhecimento. Além disso, ao analisar a cocitação de autores, aparece apenas a informação do primeiro autor da publicação, dificultando análise de redes de autores.
Acoplamento bibliográfico	Documentos	Referências compartilhadas	Conectar documentos baseado no número de referências que eles compartilham entre si.	Semelhança	É semelhante à cocitação, mas inversamente proporcional. Serve para identificar, com base na bibliografia fundamental, quais são as referências atuais que citam essa bibliografia, isto é, apresenta as frentes de pesquisa num dado campo. Com essa técnica é possível medir aproximação entre instituições e países, o que pode mostrar se as agendas de pesquisa são locais, regionais ou globais.	Não requer citações acumuladas. Pode ser usado para novas publicações que ainda não são citadas, campos emergentes e subcampos menores.	Ele só pode ser usado por um tempo limitado (até um intervalo de cinco anos). Ele não identifica inerentemente os trabalhos mais importantes por contagem de citações como cocitação; é difícil saber se as publicações mapeadas são importantes ou não.
Coautoria	Países	Colaboração	conecta países com base no número de coautorias em publicações.	Internacio-nalização	Na mesma linha de identificação das amplitudes das agendas, a técnica de coautoria pode mostrar efetivamente a internacionalização das agendas, as colaborações entre instituições e os colégios invisíveis de autores. A conexão ocorre por meio das colaborações explícitas como coautoria nos documentos.	Permite identificar as colaborações acadêmicas e a estrutura social do campo.	A colaboração em trabalhos nem sempre é reconhecida como coautoria.
Co-word	Termos/ palavras	Frequência de termos/ palavras em títulos e resumos	conecta termos pelo número de ocorrência simultaneamente em títulos e resumos.	Temáticas, afinidades teóricas ou metodológicas	Também conhecido como coocorrência de termos, esse indicador detecta os termos centrais das obras contidas no banco de dados bibliográfico e as relaciona de acordo com a frequência em que aparecem conjuntamente nos documentos. Sua utilidade principal é identificar a dinâmica de relação ou organização da literatura em grupos temáticos e encontrar conceitos centrais. Quando aplicado longitudinalmente, essa técnica é capaz de demonstrar as transformações de agenda, mostrando quais conceitos eram frequentes em anos passados e quais são frequentes atualmente.	Utiliza o conteúdo real dos documentos para análise (outros métodos só usam metadados bibliográficos).	As palavras podem aparecer de diferentes formas e podem ter diferentes significados.

Fonte: adaptado de Zupic e Čater (2015).

3. Aplicação da metodologia: o tema presidencialismo latino-americano

3.1 Construindo uma base de dados bibliográfica5

O processo de compilação de bibliografias exige cautela na escolha das bases bibliográficas de consulta e na seleção dos documentos para análise. Dividiremos este tópico em duas etapas. Na primeira justificamos a escolha das bases, a consulta de documentos e os filtros de busca. Na segunda expomos os critérios de inclusão e exclusão de documentos.

ETAPA 1 – Escolhemos analisar um tema que a disciplina de ciência política – especificamente, mas não exclusivamente – se dedica a compreender: o presidencialismo. É essencial que as bases de consulta sejam contemplativas para a disciplina, por isso, convém navegar por diversas bases antes de escolher qual utilizar. Também é possível, e por vezes imprescindível, mesclar bases, como será o caso apresentado nesse artigo. Porém, é necessário levar em consideração que isso demandará mais tempo do(a) pesquisador(a), porque nem sempre os dados bibliográficos estarão padronizados e com o mesmo ordenamento das variáveis. Isso demandará ajustes, logo, maior tempo de dedicação.

Para o nosso caso, realizamos a compilação dos artigos científicos na coleção principal da *Web of Science* (WoS)[6] e na coleção SciELO *Citation Index (SCI)*[7]. Ambas estão indexadas na plataforma da WoS, mas encontram-se separada para consulta. A primeira, produzida pelos Estados Unidos, é uma base referencial de busca de bibliografia, conhecida por dispor de um conjunto de documentos com alto impacto científico. Sua coleção contém artigos com o maior número de citações, os periódicos mais bem avaliados e apresenta uma grande diversidade de áreas de pesquisa, inclusive a ciência política e áreas afins, por isso é adequada para este estudo. A SciELO, desenvolvida no Brasil, é uma rede cooperativa direcionada principalmente para a divulgação científica latino-americana, hispânica e lusitana, de idioma português e espanhol. Selecionamos essa coleção, porque o recorte da pesquisa tem interesse regional (América Latina), assim, os documentos que não fossem contemplados pela coleção principal, poderiam estar contemplados na coleção SCI.

Por se tratar de uma área temática com consenso terminológico, o termo de busca nas plataformas foi "*presidentialism*". Quem escreve sobre o presidencialismo usualmente utiliza esse termo no título, resumo ou palavras-chave (TRP) do artigo científico, o que facilita a localização deles nas bases bibliográficas. Geralmente, nessa etapa é necessário formular uma string de busca. Essa, deve ser abrangente, de maneira a incluir o maior número de documentos relacionados ao tema de estudo, mas objetiva, a ponto de excluir aqueles que fogem da temática. Por isso é recomendado estudar a lógica de busca booleana, uma técnica que possibilita combinar termos utilizando operado-

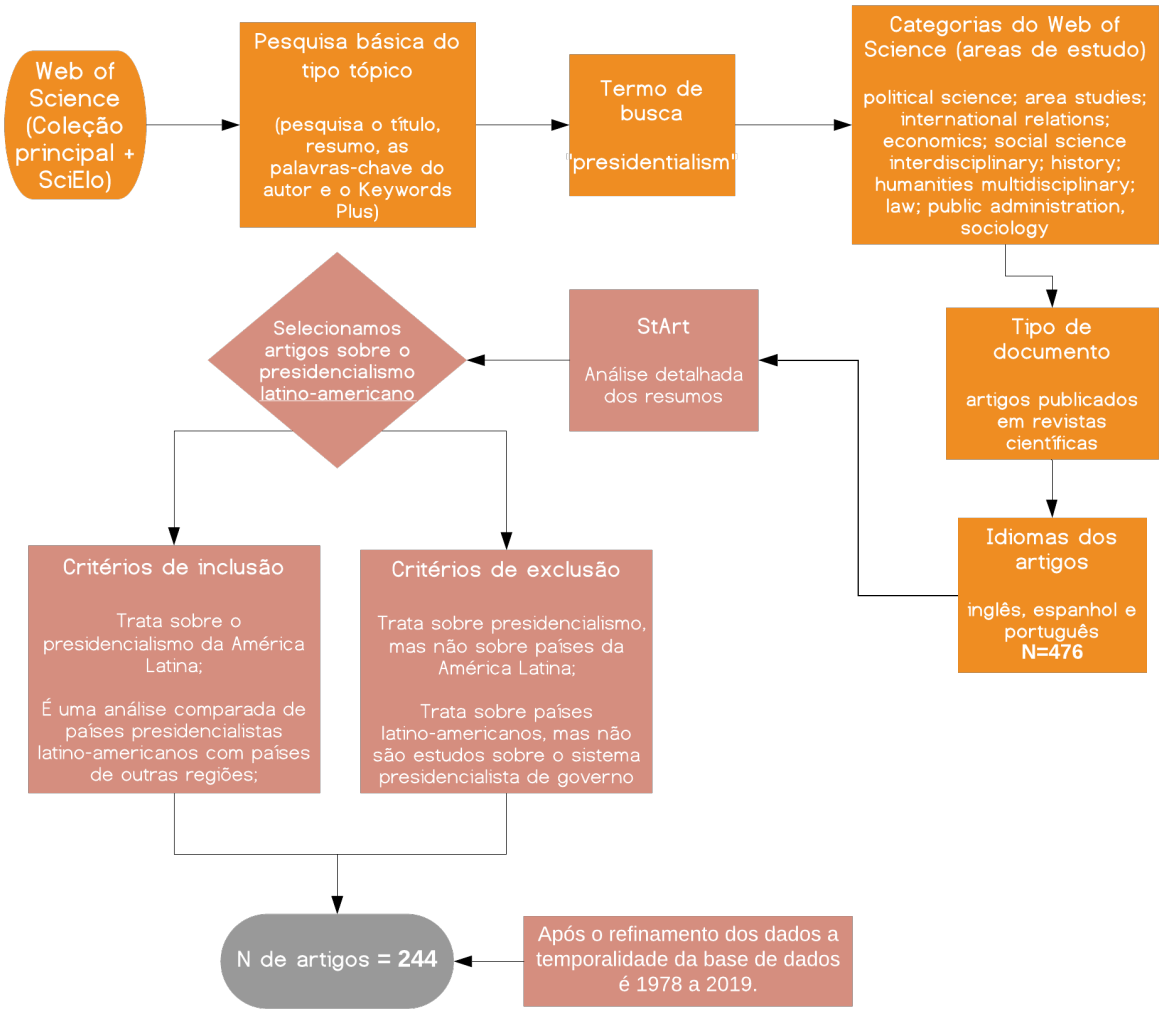
5 Toda a parte metodológica aqui presente, consta também no trabalho de conclusão de curso de Sangalli (2019), assim como alguns resultados, que estão presentes neste artigo apenas como exemplo de aplicação da metodologia e não para discussão efetiva da temática em questão.

6 "A Principal Coleção do Web of Science é a base de dados de citações líder em todo o mundo. Ela contém registros de artigos desde os periódicos de maior impacto em todo o mundo—incluindo periódicos de acesso aberto—anais de conferência e livros. A cobertura de alguns títulos data de 1900." Disponível em: https://images.webofknowledge.com/WOKRS517B4/help/pt_BR/WOS/hp_whatsnew_wos.html

7 "O SciELO Citation Index (Scientific Electronic Library Online) foi projetado para facilitar a descoberta, o uso e a atribuição de literatura acadêmica em ciências, ciências sociais, artes e humanidades publicada nos principais periódicos de acesso aberto da América Latina e do Caribe bem como Espanha, Portugal e África do Sul." Disponível em: http://images.webofknowledge.com/WOKRS5132R4.2/help/pt_BR/SCIELO/hp_whatsnew_scielo.html

res como AND, OR, NOT, entre outros, para limitar a busca nos documentos. A escolha do(s) termo(s) de busca é fundamental para esse tipo de pesquisa e demanda um razoável conhecimento prévio de palavras ou conceitos-chave do tema ou disciplina que se pretende estudar[8]. O fato de o termo escolhido estar em inglês é porque, além do idioma original do artigo, os TRPs também estão indexados em inglês nas plataformas, possibilitando encontrar textos de diversos idiomas.

Figura 1. Fluxograma da construção do banco de dados – tema presidencialismo. Buscas nas plataformas realizadas em 23 de março de 2019.



Fonte: Sangalli (2019)

Na figura 3 apresentamos um fluxograma de cada etapa de seleção dos artigos da WoS e da SCI. Foi selecionado desde a primeira publicação sobre o tema de presidencialismo registrada nas coleções. Primeiro fizemos a coleta na coleção principal WoS e na sequência aplicamos o mesmo procedimento para a coleção SCI (ver anexo 1).

Filtramos a busca selecionando áreas que poderiam conter publicações do interesse dessa pesquisa, sendo elas áreas afins da ciência política. Os filtros da base aplicados foram: *political science; area studies; international rela-*

8 "Preparar uma estratégia de busca, bem como a seleção de um banco ou uma base de dados a ser consultada para responder a uma pergunta específica, pode exigir do usuário determinados conhecimentos (idiomas, mecanismos de busca do banco de dados, instrumentos de auxílio para identificação da terminologia, lógica booleana e outros recursos disponibilizados). A implementação da estratégia de busca requer, portanto, conhecimentos técnicos e de conteúdo específicos para operacionalização da mesma" (SAKS, 2005, p.2).

tions; economics; social science interdisciplinary; history; humanities multidisciplinary; law; public administration; e, sociology. Outro filtro aplicado foi de idioma. Selecionamos apenas artigos de idiomas inglês, espanhol e português, por serem os mais frequentes e de domínio das pesquisadoras[9]. Neste momento de filtragem obtivemos 476 artigos, somando as duas coleções (figura 3).

O primeiro artigo sobre presidencialismo encontrado está indexado na WoS, datado em 1973, e os artigos mais recentes da base de dados foram publicados em março de 2019 – mês que ocorreu a coleta dos dados. É sempre importante fazer a menção de quando ocorreu a coleta dos dados para maior transparência da pesquisa, afinal as bases são constantemente atualizadas, o que pode influenciar diretamente no número de documentos encontrados por futuros pesquisadores que se dediquem a replicar ou prosseguir com a pesquisa.

ETAPA 2 – Após o download do banco de dados sobre presidencialismo, o procedimento seguinte foi refinar a seleção aplicando o recorte de estudos sobre a América Latina. Realizamos a leitura detalhada dos resumos dos artigos selecionados na etapa 1 para selecionar apenas aqueles que analisavam países latino-americanos. Utilizamos a ferramenta *State of the Art through Systematic Review* (StArt) nesse processo. Ela é útil para elaborar revisões sistemáticas, porque exige um protocolo de planejamento, execução e sumarização da pesquisa (ver anexo 2). Entretanto, utilizamo-la apenas para ler os resumos e excluir os documentos duplicados. Quando não foi possível identificar nos resumos se havia a América Latina como interesse, realizamos a leitura dinâmica do artigo para classificá-lo. Nessa análise utilizamos os seguintes critérios de inclusão e exclusão de artigos (quadro 1):

Quadro 3. Critérios de inclusão e exclusão de textos sobre presidencialismo para refinar o banco de dados.

Natureza do critério	Inclusão	Exclusão
Temático	a) É um estudo sobre o presidencialismo; b) É um estudo que compara o presidencialismo com outros sistemas de governo.	a) Não é um estudo sobre presidencialismo.
Regional	c) É um estudo de um ou mais países da América Latina; d) É uma análise comparada de países presidencialistas latino-americanos com outros países de outras regiões.	b) É sobre presidencialismo, mas não é sobre países da América Latina; c) É sobre países latino-americanos, mas não são estudos sobre o sistema de governo presidencialista.

Fonte: Sangalli (2019)

Foram escolhidos apenas aqueles cuja discussão central eram os países presidenciais da América Latina, restando 244 artigos para a análise bibliométrica. A temporalidade inicial se alterou para 1978 quando aplicado o recorte regional e permaneceu até março de 2019.

Verificamos que as revistas que os publicaram pertencem a seis áreas distintas (tabela 1). Para essa classificação lemos o escopo de cada revista em seus websites. Dividem-se entre revistas específicas, como: Ciência Política e Relações Internacionais; Direito; Economia; Administração Pública; e, História; e revistas interdisciplinares consideradas como Ciências Sociais – são revistas que permitem publicações de Sociologia, Antropologia, Ciência Política e História – em seus escopos.

9 Havia sete artigos em outros idiomas, os quais realizamos a leitura de seus resumos na própria plataforma e identificamos que não tinham como objeto de estudo o presidencialismo na América Latina – nosso foco de interesse - por isso foram excluídos manualmente, restando apenas artigos escritos nos três idiomas anteriormente mencionados.

Os artigos estão distribuídos em 101 periódicos científicos (tabela 1 e anexo 3), 58% deles estão em revistas com escopo direcionado à área de Ciência Política e Relações Internacionais, 28% em revistas interdisciplinares de Ciências Sociais, 7% em revistas de Direito e 7% em outras áreas (Adm. Pública, Economia e História).

Tabela 1. Área das revistas científicas do banco de dados, número de revistas de cada área e número total de artigos publicados em cada área sobre o tema presidencialismo latino-americano. Base de dados Wos e SCI, busca realizada em 23 de março de 2019.

Áreas	Revistas		Artigos	
	N	%	N	%
Ciência Política e Relações Internacionais	43	43	141	58
Ciências Sociais (Interdisciplinar)	32	32	67	28
Direito	10	10	18	7
Administração Pública	5	5	7	3
Economia	6	5	6	2
História	5	5	5	2
TOTAL	101	100	244	100

Fonte: elaboração Sangalli (2020)

Como esperado, as discussões sobre presidencialismo são publicadas majoritariamente em revistas de Ciência Política e Relações internacionais (n=141 artigos) e de Ciências Sociais (n=67), dada sua relevância para o campo como apresentado por Basabe-Serrano e Huertas Hernández (2018). Mas encontramos também publicações no Direito, em específico pelas análises das constituições latino-americanas (GARCÍA, 2014; DAZA, 2015; BENENTE, 2017; ORTIZ ORTIZ, 2018); na Administração Pública, pelo interesse em discutir capacidades estatais e burocracia (BERSCH, PRAÇA, TAYLOR, 2017; LIMA, 2018); na Economia, em especial sobre *accountability* (NAKAGUMA, 2015); e na História, com discussões em perspectiva histórica sobre o sistema de governo (SAN FRANCISCO, 2003; TIMMERMANN, 2021).

3.2 Visualização bidimensional dos dados

O software *VOSviewer* possibilita a criação, visualização e exploração dos dados em formato de mapas de rede. Permite visualizar uma grande quantidade de itens ou metadados, sendo capaz de gerar mapas bidimensionais. Consideramos essa ferramenta vantajosa por ser gratuita e pela facilidade de manuseio. Entretanto, ela é mais indicada para quem está iniciando o contato com esse estilo de análise, pois não apresenta opções detalhadas sobre as medidas de distância das conexões em rede. Para maiores detalhes métricos recomendamos o uso do software *Gephi*. Apesar disso, o *VOSviewer* é eficiente na construção de mapas de proximidade entre autores, revistas, instituições de pesquisa e países. As conexões são feitas pelo número de links[10] ou citações mútuas, que são representados por arestas, onde a espessura delas reflete a força de relação entre dois itens. Quanto menor a distância entre os itens plotados no espaço bidimensional, maior a relação entre eles. De acordo com a intensidade dessas ligações é que ocorre a conformação de *clusters*, isto é, grupos de itens com aspectos comuns que são representados por uma mesma coloração.

Há três tipos de visualizações fornecidas pelo VOS-

10 "Um link é uma conexão ou relação entre dois itens. Exemplos de links são links de acoplamento bibliográfico entre publicações, links de coautoria entre pesquisadores e links de co-ocorrência entre termos. Um mapa normalmente inclui apenas um tipo de link" (VAN ECK, WALTMAN, 2017, p. 4)

viewer para os mapas: por rede, por sobreposição e por densidade (*network visualization*, *overlay visualization*, e a *density visualization*, respectivamente). A primeira visualização apresenta os itens por rótulo, padrão de coloração e um "nó", sendo o tamanho do rótulo e do nó o peso de cada item. A cor de um item é determinada pelo *cluster* ao qual pertence. As arestas entre itens representam os links, assim, quanto mais próximos estiverem os itens no espaço bidimensional, mais forte são suas relações. A segunda visualização é idêntica à anterior, exceto por sua coloração. Existem duas formas de clusterização, por citação (maior/menor) ou por ano da publicação (antigo/atual). Uma legenda de cores indica como os itens estão mapeados. A terceira visualização tem duas opções possíveis, a *item density*, estabelece em cada ponto do mapa a densidade dos itens, que também é indicada por cores, no entanto, é definida pelo peso e número de itens vizinhos. E a densidade por cluster (*cluster density*), onde a cor de um ponto é definida pela mistura dos clusters vizinhos e o peso é atribuído à cor, de acordo com o número de itens pertencentes ao cluster que se encontram na margem de vizinhança (VAN ECK, WALTMAN, 2017)[11].

O software permite criar mapas baseados em dados bibliográficos a partir de formatos específicos de banco de dados, são eles: Web of Science, Scopus, PubMed, RIS e Crossref JSON. Assim, os dados coletados da coleção SCI precisaram de um tratamento para que pudessem ser reconhecidos. Foi padronizada a ordem das variáveis através da planilha Microsoft Excel e os dados provenientes da SCI incorporados manualmente à base da coleção WoS. Após isso, o arquivo foi salvo em formato de documento de texto para a leitura no software.

11 Outra função de visualização, são as técnicas de layout e clustering do mapa. Dentre as técnicas temos os métodos de normalização, força de associação, fracionamento e LinLog/modularidade, os três corrigem as forças dos links entre os itens para uma melhor visualização bidimensional. Para compreender suas diferenças ver Van Eck e Waltman (2009, 2017). Também é possível modificar o layout utilizando os mecanismos de atração e repulsão, que serve apenas para melhorar a visualização do mapa bidimensional e não altera seu resultado, devendo o valor de repulsão ser sempre menor que o de atração. Mais informações sobre manuseamento do VOSviewer estão disponíveis no manual elaborado por Van Eck e Waltman (2017).

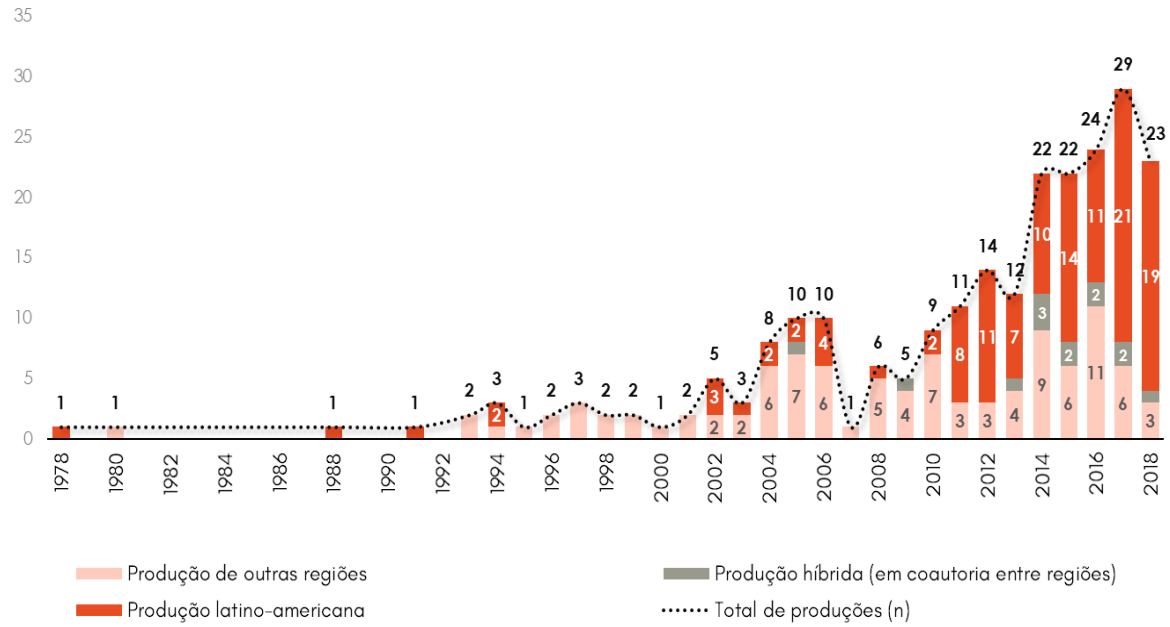
4. Resultados

4.1 Produção e colaborações: países, gênero e coautoria

O número de publicação de artigos sobre o tema em análise é crescente ao longo dos anos, por mais que oscile (gráfico 1). De 1978 até 1991 encontramos apenas 4 artigos sobre a temática nas plataformas. A partir de 1992 até 2003 foram publicados em média 2,16 artigos por ano, totalizando 26 artigos. Em 2004 aumenta o número de publicações, deste ano até 2013 a média é de 8,6 artigos ao ano. De 2014 até 2018 é quando temos o maior número de publicações (n=120), nesses últimos anos a média aumentou para 24 artigos.

Quando examinamos a origem institucional de autores dos artigos, percebemos que, nas bases em questão, havia uma predominância de estudos originários de países externos à América Latina até 2010 (gráfico 1), que chamaremos de estudos latino-americanistas. São poucas as produções latino-americanas até então. Isso pode ser justificado pelo fato de a ciência política da região ser jovem, pois a estruturação da pós-graduação ocorreu a partir dos anos 1980 (BARRIENTOS DEL MONTE, 2013) e somente após sua institucionalização e autonomização, por volta dos anos 2000 (LEITE, CODATO, 2013), é que a disciplina dispôs de incentivos e investimentos para publicações de artigos científicos. Em 2011 há um aumento expressivo de artigos, no qual a produção latino-americana passa a ser predominante para a temática. Pelo contexto da disciplina, esse crescimento também ocorre para outros temas da ciência política, como demonstrado por Basabe-Serrano e Huertas Hernández (2018), tais como partidos, democracia, políticas públicas, entre outros.

Gráfico 1. Número de artigos científicos sobre o presidencialismo latino-americano por ano de publicação e origem institucional de autoras(es) no banco de dados (1978-2018). Base de dados Wos e SCI, busca realizada em 23 de março de 2019.



Fonte: elaboração Sangalli (2020)

A proporção de cada tipo de produção no conjunto de dados é de 51% de produção latino-americana e 43% de outras regiões. Somente 6% é produção híbrida (em coautoria entre autor(a) latino-americano e autor(a) de outra região), concentradas na última década. Quando verificamos as plataformas separadamente, a maior parte dos artigos estão indexados na WoS (66%) e a maioria das produções da América Latina está na plataforma SCI (29%), isto é explicado pelo próprio alcance das plataformas (tabela 2).

Tabela 2. Proporção dos tipos de produção separados por coleção bibliográfica e a distribuição de artigos em número e porcentagem. Artigos sobre o presidencialismo latino-americano (1978-2019) nas bases de dados WoS e SCI. Busca realizada em 23 de março de 2019.

	produção latino-americana		produção híbrida		produção de outras regiões		TOTAL	
	N	%	N	%	N	%	N	%
Web of Science	53	22	11	5	97	40	161	66
SciELO Citation Index	72	29	3	1	8	3	83	34
TOTAL	125	51	14	6	105	43	244	100

Fonte: elaboração Sangalli (2020)

A partir da divisão entre tipo de produção (gráfico 1 e tabela 2), mostramos quais são os centros de pesquisa sobre a temática em questão (tabelas 3 e 4). Estados Unidos e Brasil são os países quem mais produziram artigos, compõe 25% e 22% dos artigos da base dados, respectivamente. A produção do México (9%) está em terceiro lugar, na sequência aparecem Argentina, Chile e Reino Unido (6% cada país). Ainda que as produções latino-americanas sejam maioria, nota-se a predominância de estudos estadunidenses distribuído por todo o período analisado, ainda que exista predominância nas primeiras décadas (1980-1990).

Além de ser o segundo polo de pesquisas sobre o assunto, o Brasil também é o país mais estudado, 24% dos artigos contribuem sobre o país. Isso ocorre devido

a agenda de estudos sobre o presidencialismo de coalizão, inaugurada por Abranches, em 1988 (SANGALLI, 2019). A produção sobre o México está em segundo lugar (8%) contribuindo sobre a estrutura do presidencialismo mexicano e as reformas constitucionais do país (HERNANDEZ RODRIGUEZ, 1978, 1994; HERNANDEZ CHAVES,1994; PANIZZA, PHILIP, 2005; NEGRETTO, 2006; MAYER-SERRA, 2017). Em terceiro lugar está a produção sobre a Argentina (7%) concentrada na discussão sobre democracia, populismo, autoritarismo e seus efeitos no sistema de governo presidencialista (PHILIP, 1998; LARKINS, 1998; QUINCHE RAMÍREZ, 2013; TEDESCO, DIAMINT, 2014; TARTAKOFF, 2019), também há discussões sobre crise política e econômica no caso argentino (MUSTAPIC, 2005; MALAMUD, 2015; PEÑA, DAVIES, 2017). Enquanto os estudos que analisam o presidencialismo chileno focam no debate de democracia (RABKIN, 1996; SAN FRANCISCO, 2003; CAMERON, 2010) e de coalizões políticas RENIU, ALBALA, 2012; HUNEEUS 2012; SIAVELIS, 2016; BORGES, TURGEON, 2019; CHAISTY, POWER, 2019).

A presença de produções de treze países de outras regiões do mundo em uma temática específica como a analisada, demonstra que essa é uma agenda de estudos internacionalizada. Como certificado na figura 4, a coautoria com produção em maior intensidade é entre Estados Unidos e Brasil, com força do link igual a 6. Na sequência há a ligação entre Brasil e Inglaterra (3) e Argentina e Espanha (2). É particularmente interessante visualizar a distância do Brasil dos demais países latino-americanos no mapa bidimensional, demonstra maior colaboração com países do norte global[12]. Há ainda, um grupo de países da América Latina mais próximos entre si no mapa – conformam o cluster amarelo –, indicando laços científicos no estudo da temática entre Chile, Colômbia, Venezuela e Equador.

Tabela 3 e 4. Países que mais produziram e países mais estudados na temática de presidencialismo latino-americano. Proporção e número de artigos (WoS e SCI, 1978-2019). Busca realizada em 23 de março de 2019.

País de produção	Artigos	
	N	%
Estados Unidos	69	25
Brasil	61	22
México	24	9
Argentina	17	6
Chile	17	6
Reino Unido	17	6
Colômbia	13	5
Espanha	8	3
Canadá	7	3
Uruguai	6	2
Venezuela	4	1
Portugal	5	2
Noruega	4	1
Alemanha	4	1
Equador	3	1
França	3	1
Suécia	2	1
Turquia	2	1
Austrália	2	1
Coreia do Sul	2	1
Finlândia	1	0
Irlanda	1	0
China	1	0
Total	273*	100

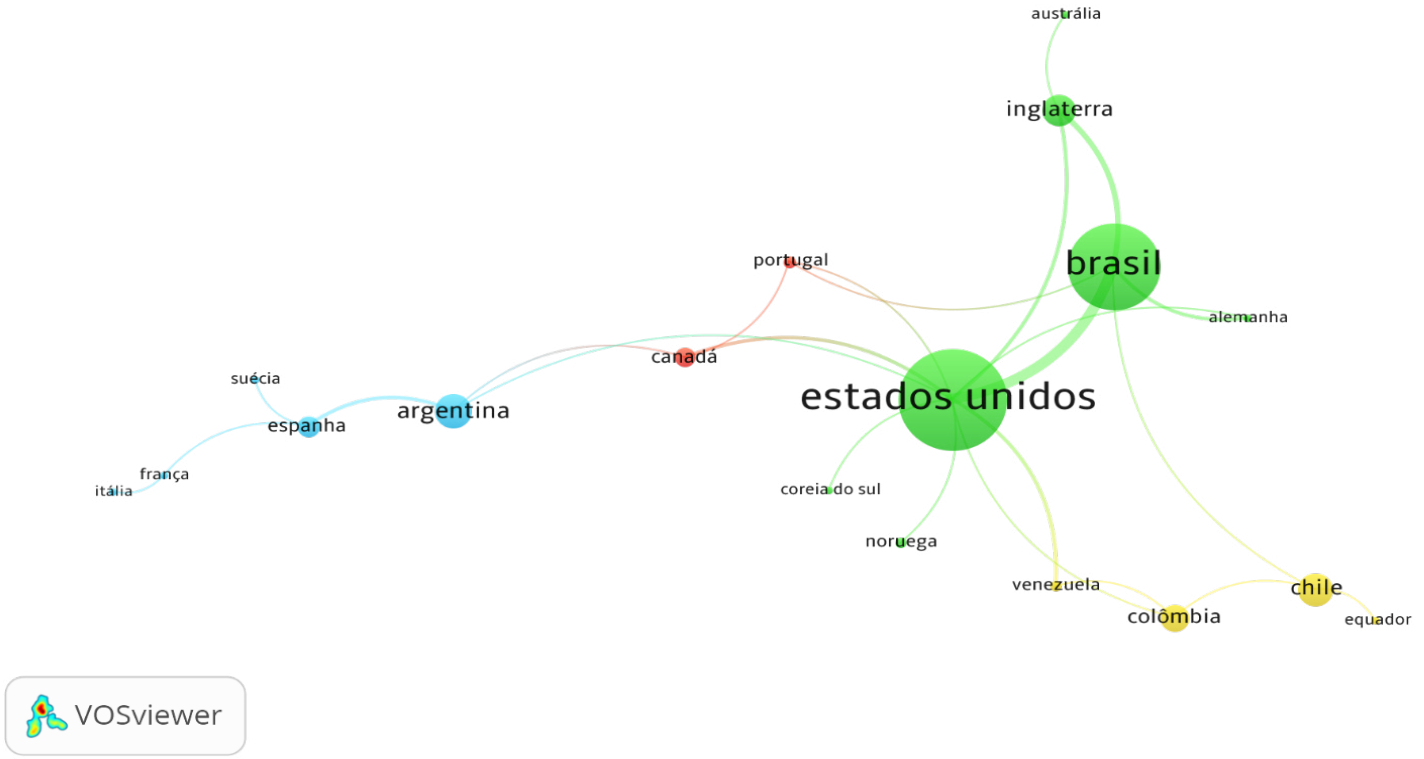
País estudado	Artigos	
	N	%
Brasil	73	24
México	24	8
Argentina	22	7
Chile	18	6
Venezuela	14	5
Equador	13	4
Uruguai	11	4
Peru	9	3
Colômbia	8	3
Paraguai	6	2
Rep. Dominicana	4	1
Costa Rica	2	1
Nicarágua	1	0
Bolívia	1	0
Região agrupada*	96	32
Total	302**	100

Fonte: elaboração Sangalli (2020)

*esse total representa o número de autores, pois a medida é a origem institucional destes, assim os artigos em coautoria contaram para ambos os países.

12 O país apresenta apenas uma força de ligação com o Chile.

Figura 2. Mapa de rede de coautoria de países: 18 itens* (mínimo de 1 documento e 0 citações. Método de contagem: completa. Método de normalização: LinLog/modularidade (atração 1, repulsão -1); Método de visualização: network visualization, peso por número de documentos e coloração por cluster. Quanto maior o número de documentos publicados pelo país, maior é o rótulo e o nó do item. Quanto mais próximos os itens no espaço bidimensional e mais espessa a aresta de conexão, mais forte a relação de coautoria entre eles. (Bases de dados: WoS e SCI, busca realizada em 23 de março de 2019. Banco de dados sobre presidencialismo latino-americano).



Fonte: elaboração Sangalli (2020)
*para obter esse mapa foi necessário construir um tesauro.

No que se refere ao perfil de autores e autoras,

> a pesquisa bibliométrica vem demonstrando que boa parte do conhecimento em Ciência Política é produzido por grupos de pessoas com características semelhantes, o que em si invisibiliza outros atores e, consequentemente, seus discursos e visões. Existe um déficit de representatividade de gênero, raça e origem regional, apenas para mencionar algumas das clivagens mais recorrentes em estudos bibliométricos. A expansão da Ciência Política para todo o globo não foi acompanhada de uma democratização da divulgação de conhecimento, impactando não só na invisibilidade de temas e abordagens epistemológicas de grupos historicamente minoritários na disciplina; como também nas carreiras e na estrutura de incentivos às mesmas (LENINE, MÖRSCHBÄCHER, 2020).

Os nossos dados sobre a temática também revelam disparidade (tabela 5 e gráfico 2). Dividimos a literatura em três grupos para medir a intensidade de produção por gênero e coautoria: i) produção de homens: sem a presença de mulheres em nenhuma posição de autoria; ii) produção mista: com presença de mulheres em alguma instância de autoria (1ª, 2ª ou 3ª) e homens nas demais posições; e, iii) produção de mulheres: aquelas com presença de mulheres em todas as instancias de autoria (tabela 5).

Tabela 5. Proporção das produções por gênero nas coleções bibliográficas e a distribuição de artigos sobre presidencialismo latino-americano em número e porcentagem. Busca realizada em 23 de março de 2019.

	Produção de homens		Produção mista		Produção de mulheres		TOTAL	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Web of Science	113	46	21	8	27	11	161	66
SciELO Citation Index	62	26	9	4	12	5	83	34
TOTAL	175	72	30	12	39	16	244	100

Fonte: elaboração Sangalli (2020)

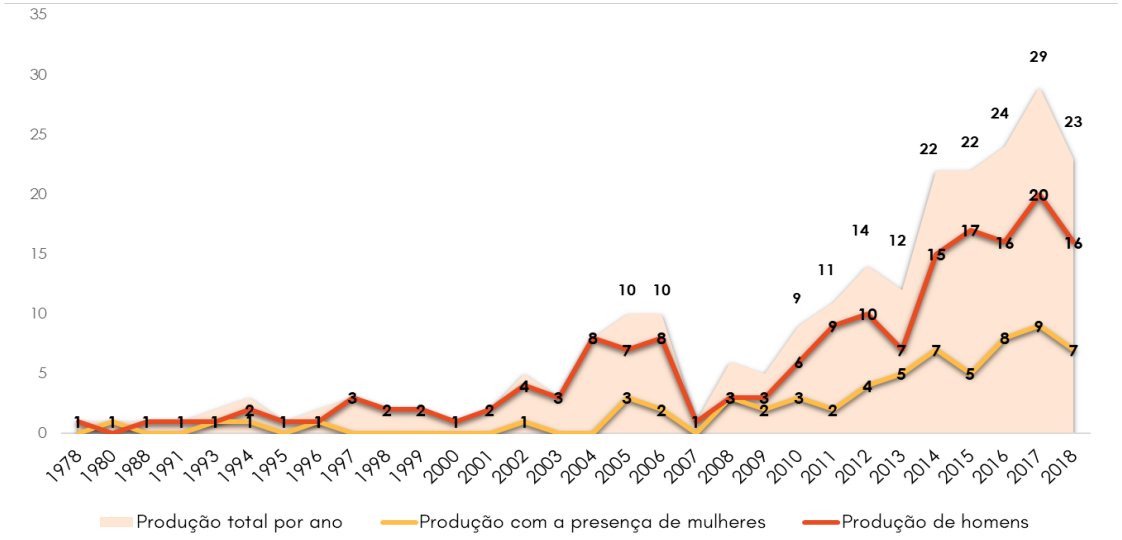
Como sustentado por Freidenberg (2019), a ciência política ainda enfrenta resistências à igualdade de gênero. A produção sobre a temática também releva essa disparidade de gênero (tabela 5). No conjunto de dados 72% dos artigos são escritos por homens, à medida que a produção de mulheres representa 16% dos artigos, sendo ainda mais baixa a porcentagem de colaboração entre gênero, 12% de produção mista. Ainda quando observamos cada base, ambas apresentam maior proporção de artigos escritos por homens, não demonstrando um viés entre as plataformas.

Dentre os três pilares das ciências sociais (sociologia, antropologia e ciência política) a ciência política na América Latina, e em específico a brasileira, ainda é a mais desigual na distribuição de cargos de docência, tanto em relação a gênero, como raça (CANDIDO, FERES JÚNIOR, CAMPOS, 2019). Para Rocha-Carpiuc (2016, apud Rocha-Carpiuc, Madeira, 2019, p.552),

> as mulheres encontram-se sub-representadas nos postos acadêmicos mais importantes (universidades e sistemas nacionais de investigação e avaliação). Os dados apontam que ao longo da trajetória acadêmica, as mulheres têm maiores dificuldades de galgar tais postos. Quanto à produção acadêmica, as mulheres têm maior presença como apresentadoras em congressos do que como autoras em artigos nas revistas acadêmicas mais prestigiadas.

De acordo com a autora ainda existem obstáculos maiores para as mulheres publicarem artigos científicos. Realmente, como verificamos no nosso caso de estudo, a produção de mulheres é menor. Ainda que, no gráfico 2, seja perceptível que na última década houve um aumento da presença de mulheres na autoria dos artigos em relação aos anos anteriores.

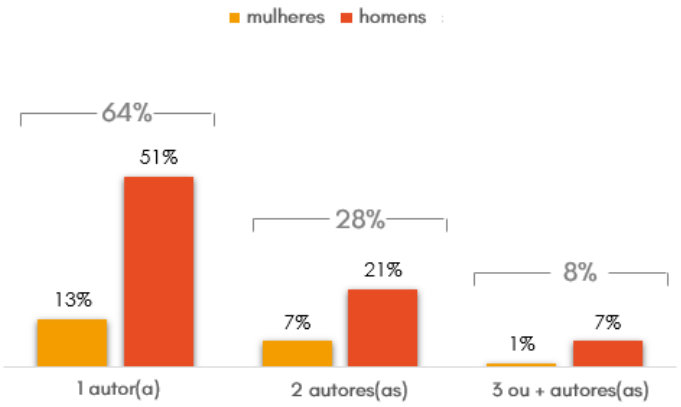
Gráfico 2. Presença de gênero na produção anual sobre o presidencialismo latino-americano (WoS e SCI, 1978-2018, busca realizada em 23 de março de 2019).



Fonte: elaboração Sangalli (2020)

Esses estudos que investigam a desigualdade de gênero na ciência vêm demonstrando que há lacunas nas políticas de avaliação que desconsideram as diferentes socializações de gênero[13]. Em contextos de maternidade, por exemplo, a produção científica de mães pesquisadoras é afetada pela dupla ou tripla jornada de trabalho, dificultando a produção e publicação de artigos. É fundamental que estudos bibliométricos levem em consideração que nem sempre produtividade é sinônimo de qualidade, por isso, mesmo que a produção seja baixa é necessária uma análise de impacto, que combine múltiplos fatores – que não apenas a citação – e considere o contexto social que o(a) pesquisador(a) está inserido. Assim seria possível uma avaliação de pesquisadores mais democrática, que promova incentivos financeiros ou simbólicos para condições diferentes de atuação profissional.

Gráfico 3. Instâncias de autoria por gênero nos artigos sobre presidencialismo latino-americano (WoS e SCI, 1978-2019, busca realizada em 23 de março de 2019).



Fonte: elaboração Sangalli (2020).

13 ANDRADE, Rodrigo de Oliveira. Maternidade no currículo: projeto de pesquisadoras estima impacto das atividades como mãe na carreira científica. Pesquisa Fapesp, São Paulo, ed. 269, 2018. Disponível em: https://revistapesquisa.fapesp.br/2018/07/19/ maternidade-no-curriculo/. Acesso em: 05 de dezembro de 2020.

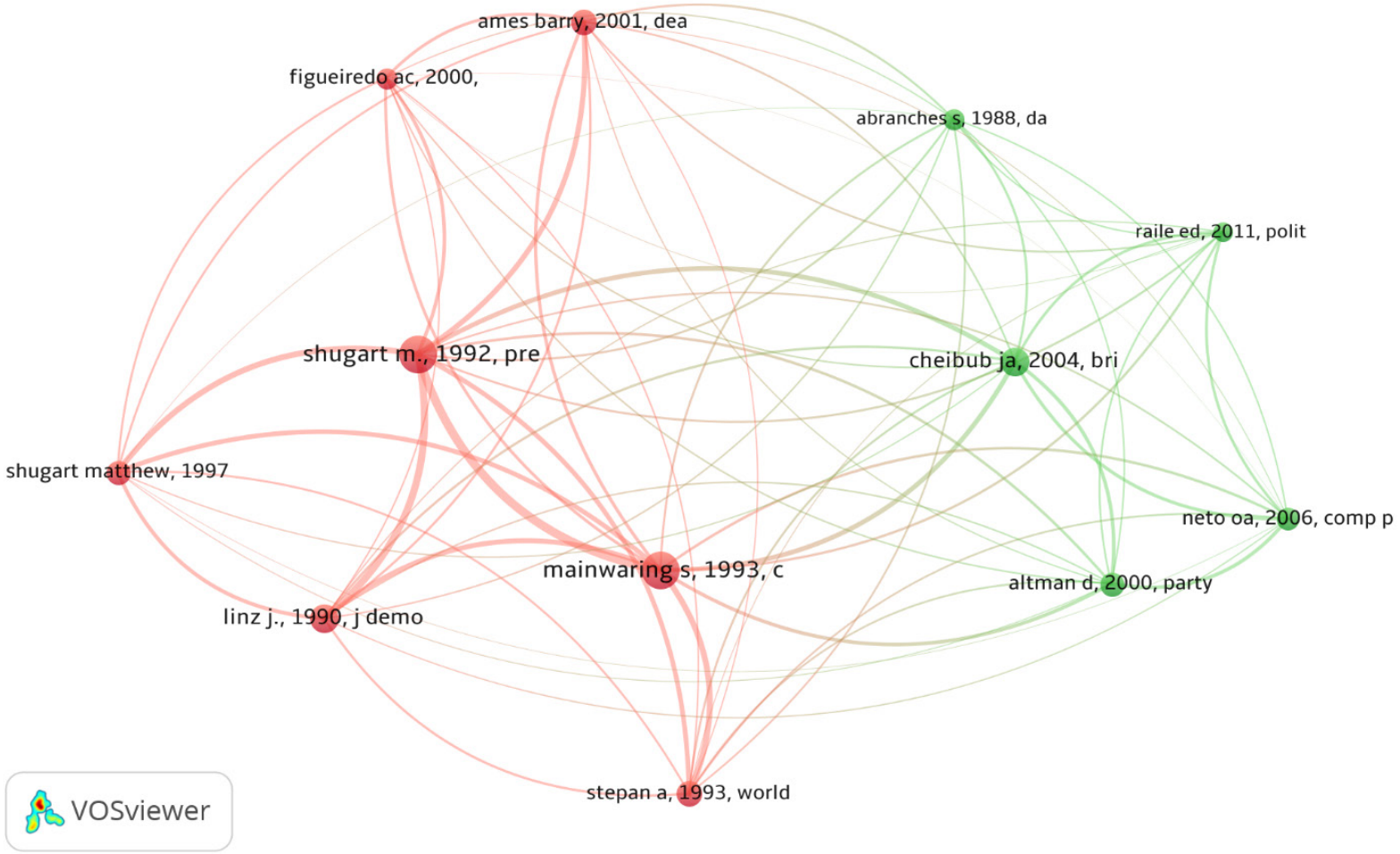
A maior parte da produção é de autoria única (64% dos artigos), 28% deles são em coautoria entre dois autores(as) e 8% escrito em colaboração entre 3 ou mais autores(as). Mesmo que numericamente menor, a produção de mulheres se concentra na mesma instância de autoria que a produção de homens, não aparentando desnivelamento ou indicando marginalização nesse sentido.

4.2 Cocitação: conhecendo os documentos fundamentais da temática

A cocitação é mais vantajosa do que apenas a citação isolada do artigo, pois apresenta uma cadeia de artigos que são citados juntos, mostrando afinidades de discussão para além da relevância para a comunidade científica. No mapa de cocitação, os textos que aparecem não são necessariamente os que estão no nosso banco de dados, mas sim as referências comuns entre os artigos do nosso banco, podendo aparecer artigos de discussões paralelas ao tema central de presidencialismo latino-americano.

Identificamos duas tradições intelectuais que moldam a estrutura dessa área de estudos (figura 4). No cluster vermelho encontramos textos clássicos sobre sistemas de governo, que discutem as vantagens e desvantagens do presidencialismo e o compara com o parlamentarismo. A maioria desses estudos foram escritos na década de 1990 e no início dos anos 2000. Podemos considerar com base na força total do link (quadro 2), que o livro de Shugart e Carey (1992) e os artigos de Mainwaring (1993) e de Linz (1990), pertencentes ao cluster vermelho, são as referências mais mencionadas nos estudos sobre a temática, com força de conexão >100. Enquanto no cluster verde, encontramos estudos específicos sobre presidencialismo de coalizão, sendo os estudos de Cheibub, Pzeworski e Saiegh (2004), Altman (2000) e Abranches (1988) os mais influentes desse cluster, com força de conexão >60.

Figura 3. Mapa de rede de cocitação de documentos com os 12 artigos sobre presidencialismo latino-americano mais citados juntos do banco de dados e suas relações. Os documentos são conectados de acordo com o número de citações recíprocas. Método de contagem: completa. Método de normalização: LinLog/modularidade (atração 2, repulsão 1); Método de visualização: network visualization, peso pelo total de força do link. Quanto mais próximos os itens no espaço bidimensional, mais forte são suas relações. A coloração ocorre pela conformação de clusters. Quanto maior a quantidade de links, maior é o rótulo e o nó do item (Busca realizada na WoS e SCI em 23 de março de 2019).



Fonte: elaboração Sangalli (2020)

Quadro 4. Ranking dos 12 documentos fundamentais sobre presidencialismo latino-americano segundo o total de aparições conjuntas (links) nas referências dos artigos do banco de dados (Busca realizada na WoS e SCI em 23 de março de 2019).

Ranking	Referência	Espécie	Citações	Links	Força total do link	Cluster
1	Shugart, M. S., & Carey, J. M. (1992). Presidents and assemblies: Constitutional design and electoral dynamics. Cambridge University Press.	livro	65	11	166	**vermelho**
2	Mainwaring, S. (1993). Presidentialism, multipartism, and democracy: the difficult combination. Comparative political studies, 26(2), 198-228.	artigo	54	11	156	**vermelho**
3	Linz, J. J. (1990). The perils of presidentialism. Journal of democracy, 1(1), 51-69.	artigo	41	11	102	**vermelho**
4	Shugart, M. S., & Mainwaring, S. (1997). Presidentialism and democracy in Latin America: Rethinking the terms of the debate. In Presidentialism and democracy in Latin America (pp. 12-54). Cambridge University Press.	capítulo de livro	34	10	81	**vermelho**
5	Cheibub, J. A., Przeworski, A., & Saiegh, S. M. (2004). Government coalitions and legislative success under presidentialism and parliamentarism. British Journal of Political Science, 565-587.	artigo	33	11	102	**verde**
6	Stepan, A., & Skach, C. (1993). Constitutional frameworks and democratic consolidation: Parliamentarianism versus presidentialism. World Politics, 1-22.	artigo	32	11	88	**vermelho**
7	Ames, B. (2001). The deadlock of democracy in Brazil. University of Michigan Press.	livro	26	11	84	**vermelho**
8	Altman, D. (2000). The politics of coalition formation and survival in multiparty presidential democracies: The case of Uruguay, 1989-1999. Party Politics, 6(3), 259-283.	artigo	25	11	79	**verde**
9	Abranches, S. (1988). Presidencialismo de coalizão: o dilema institucional brasileiro. Dados, 31(1), 5-38.	artigo	23	11	61	**verde**
10	Figueiredo, A. C., & Limongi, F. (2000). Presidential power, legislative organization, and party behavior in Brazil. Comparative Politics, 151-170.	artigo	21	11	66	**vermelho**
11	Raile, E. D., Pereira, C., & Power, T. J. (2011). The executive toolbox: Building legislative support in a multiparty presidential regime. Political Research Quarterly, 64(2), 323-334.	artigo	21	10	59	**verde**
12	Neto, O. A. (2006). The presidential calculus: Executive policy making and cabinet formation in the Americas. Comparative Political Studies, 39(4), 415-440.	artigo	20	11	68	**verde**

Fonte: elaboração Sangalli (2020).

4.3 Acoplamento bibliográfico e co-word: identificando as frentes de pesquisa e os conceitos emergentes da temática

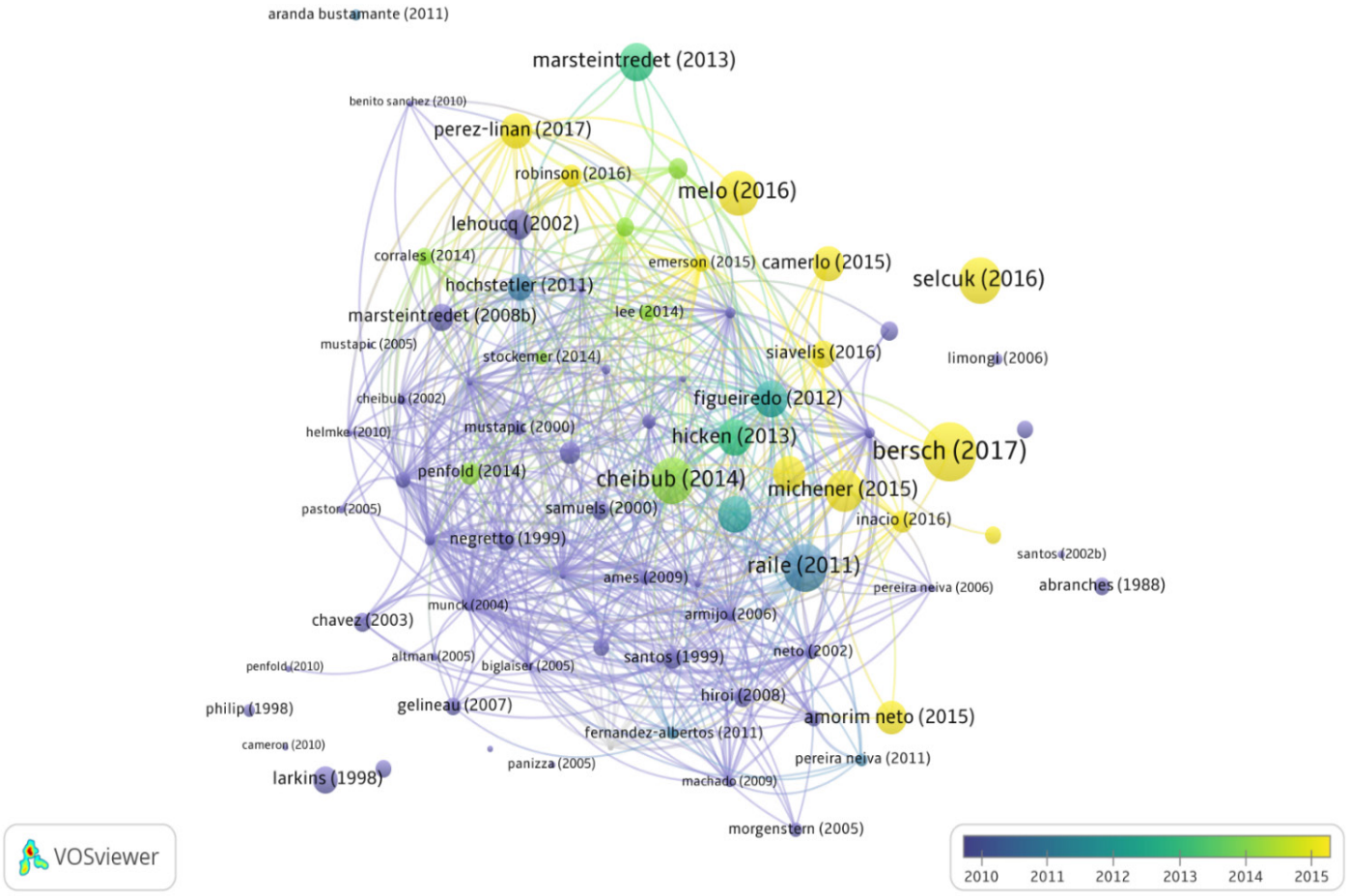
O acoplamento bibliográfico é a técnica mais recomendada para identificar as frentes de pesquisa. Sua aplicabilidade é para períodos curtos de publicação (até 5 anos), por isso não é aconselhável basear sua interpretação no número de citações. Entretanto, uma alternativa é utilizar a citação normalizada[14]. A normalização corrige o fato de que documentos mais antigos tiveram mais tempo para receber citações do que documentos mais recentes (VAN ECK, WALTMAN, 2017), atribuindo um peso ponderado de citação dado o tempo de publicação do artigo.

Criamos um mapa de sobreposição (figura 5) onde o enforque de análise são os artigos com coloração amarela. Quanto maior o número de referências que os artigos compartilham, mas forte a conexão entre eles. Assim, quanto mais próximos no espaço bidimensional, maiores são suas afinidades teóricas e/ou metodológica, indicando uma relação implícita entre eles (GRÁCIO, 2016). Os estudos mais recentes, isto é, as frentes de pesquisa, encontram-se no quadro 3 em formato de ranqueamento pelo peso de citação normalizada. Foi possível identificar dois clusters temáticos, o cluster 1 é sobre crise política e o cluster 2 é sobre administração pública e burocracia.

De modo complementar, o mapa de co-word (figura 6) mostra quais são os conceitos principais dessas frentes de pesquisa. Conectando palavras que aparecem juntas nos textos, identificamos que os conceitos emergentes (cor amarela) que são: a) Impeachment: está conectado com palavras como crise, democracia, participação e estabilidade, apontando qual a descendência do debate sobre impeachment. A emergência desse conceito está relacionada com a conjuntura latino-americana, uma vez que casos recentes de impeachment ocorreram na região, nos países Equador, Paraguai, Guatemala e Brasil; b) Ministério, gabinete e agency: esses termos estão conectados entre si e a outros termos como governo de coalizão e governabilidade. É presumível, portanto, que eles são centrais nos estudos sobre capacidades burocráticas e nos estudos que analisam a relação entre os poderes Legislativo e Executivo.

14 "O número normalizado de citações de um documento é igual ao número de citações do documento dividido pelo número médio de citações de todos os documentos publicados no mesmo ano e incluídos nos dados que são fornecidos ao VOSviewer" (VAN ECK, WALTMAN, 2017, p. 32, tradução nossa).

Figura 4. Mapa de sobreposição de acoplamento bibliográfico com 79 itens (mínimo de 3 citações e máximo de 50 citações). Método de contagem: completa. Método de normalização: força de associação (atração 2, repulsão 0); Método de visualização: overlay visualization, peso por citação normalizada e coloração por ano de publicação. Quanto maior o número de citação normalizada, maior é o rótulo e nó do item, quanto mais recente é a publicação, mais amarelada é a coloração do item (Banco de dados sobre presidencialismo latino-americano. Busca realizada na WoS e SCI em 23 de março de 2019).



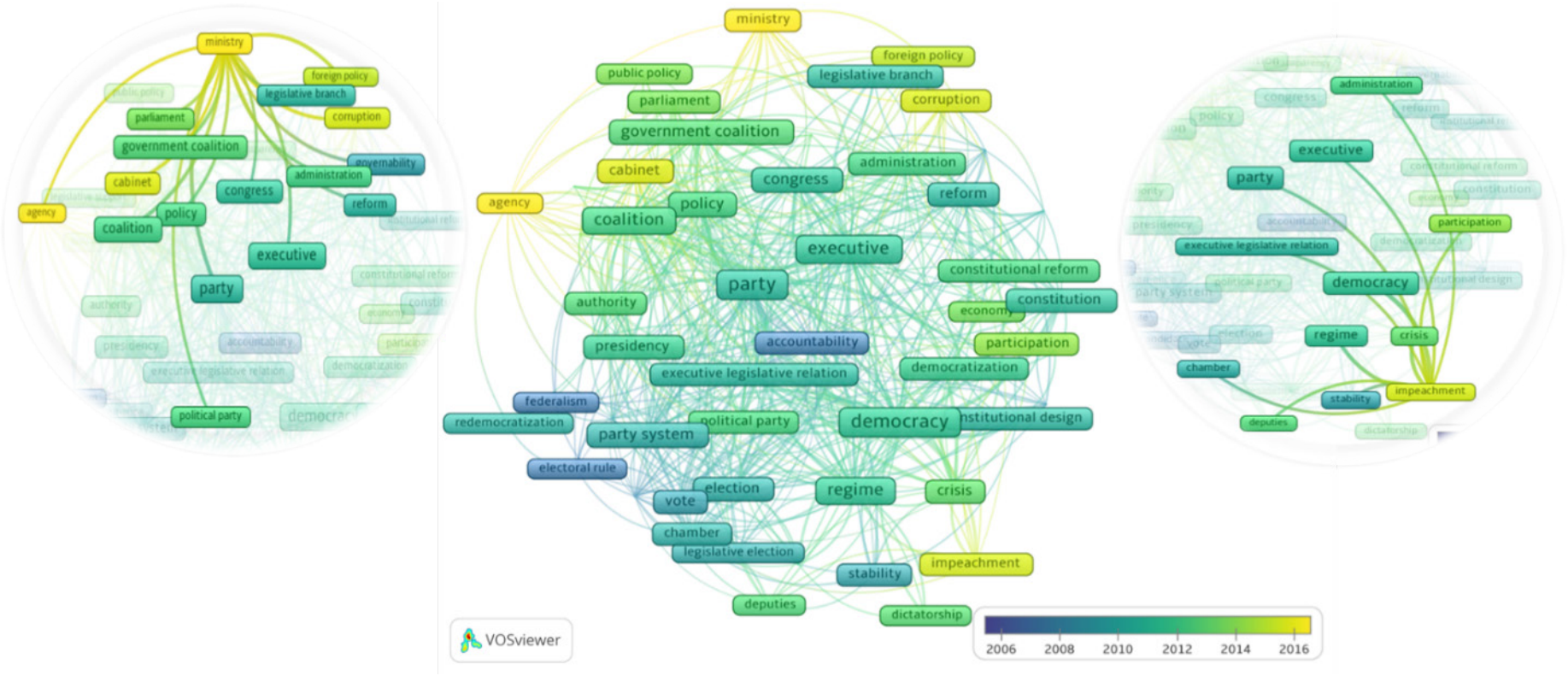
Fonte: elaboração Sangalli (2020).

Quadro 5. Ranking das frentes de pesquisa: os 10 artigos mais recentes sobre presidencialismo latino-americano, com maior peso de citação normalizada (Busca realizada na WoS e SCI em 23 de março de 2019).

Ranking	Autoria	Artigo	Score: ano de publicação	Revista científica	Citações	Weigths: citação normalizada	Links	Força total do link	Cluster
1	Bersch, K., Praça, S., Taylor, M. M.	State Capacity, Bureaucratic Politicization, and Corruption in the Brazilian State	2017	Governance	57	10.80	34	55	2
2	Selçuk, O.	Strong presidents and weak institutions: populism in Turkey, Venezuela and Ecuador	2016	Southeast European and Black Sea Studies	3	6.70	5	5	1
3	Melo, M. A.	Latin America's New Turbulence: Crisis and Integrity in Brazil	2016	Journal of Democracy	5	6.14	49	75	1
4	Michener, G.	How cabinet size and legislative control shape the strength of transparency laws	2015	Governance	18	5.33	50	87	2
5	Pérez-Liñán, A., Polga-Hecimovich, J.	Explaining military coups and impeachments in Latin America	2017	Democratization	10	4.05	42	116	1
6	Martínez-Gallardo, C., Schleiter, P.	Choosing Whom to Trust: Agency Risks and Cabinet Partisanship in Presidential Democracies	2015	Comparative Political Studies	48	3.76	40	72	2
7	Camerlo, M., Pérez-Liñán	Minister Turnover, Critical Events, and the Electoral Calendar in Presidential Democracies	2015	The Journal of Politics	14	3.76	38	60	2
8	Amorim Neto, O., Malamud, A.	What Determines Foreign Policy in Latin America? Systemic versus Domestic Factors in Argentina, Brazil, and Mexico, 1946–2008	2015	Latin American Politics and Society	11	3.45	28	42	2
9	Siavelis, P. M.	Cabinets and Informal Advisory Networks in Multiparty Presidential Systems	2016	Presidential Studies Quarterly	9	2.23	51	94	2
10	Inácio, M., Llanos, M.	The Institutional Presidency in Latin America: A Comparative Analysis	2016	Presidential Studies Quarterly	20	1.67	52	110	2

Fonte: elaboração Sangalli (2020)

Figura 5. Mapa de rede de coocorrência com os 56 termos* mais frequentes (mínimo de 5 aparições) nos títulos e resumos dos 244 artigos sobre presidencialismo latino-americano e suas relações recíprocas. Os termos são conectados de acordo com o número de aparições conjuntas em texto. Método de contagem: binária; Método de normalização: LinLog/modularity (atração 3, repulsão 0); Método de visualização: overlay visualization: peso por ocorrência. Quanto maior a frequência do termo, maior seu peso, portanto, maior seu rótulo. Quanto mais espessa a aresta, maior a intensidade de conexão entre os termos. A proximidade entre termos no espaço bidimensional representa maior força em suas relações. A coloração está associada ao ano de publicação do artigo que contém o termo, quanto mais recente, mais amarelada a cor do item. Ao lado do mapa estão duas imagens de zoom nos conceitos emergentes com suas conexões. (Busca realizada na WoS e SCI em 23 de março de 2019)



Fonte: elaboração Sangalli (2020).

*Para obter esse mapa foi necessário construir um tesauro.

5. Conclusão

Este artigo diferenciou os conceitos e aplicações dos métodos de bibliometria e cientometria e propôs demonstrar como aplicá-los ao estudo de um tema de interesse da ciência política. As principais vantagens de utilizar os métodos de métrica da informação são sua confiabilidade e amplitude. A coleta de dados quando bem executada economiza o tempo do(a) pesquisador(a) na organização dos dados e facilita sua transposição para softwares de redes bibliométricas (HAYASHI, 2013), além de garantir a replicação pela transparência das etapas de seleção. O contato bibliométrico com uma área do conhecimento apresenta uma vasta gama de utilidades. Em primeiro momento, permite analisar múltiplas dimensões da ciência, como: quem escreve, o que escreve, onde divulga, como a comunidade se comunica, permitindo a compreensão do perfil e do comportamento da comunidade científica.

Em segundo plano, consideramos valioso que um(a) estudante-pesquisador(a) em início de carreira saiba utilizar tal metodologia, porque para além de medir e mensurar a atividade científica, a aprendizagem da cientometria pode servir como meio de planejamento de carreira. Indicadores bibliométricos como o de coautoria pode auxiliar na identificação de quais são os países e instituições que mais se dedicam ao estudo de uma temática, sendo possível, por exemplo, que esse(a) estudante se organize para pleitear bolsas de estudos de pós-graduação em instituições que possam contribuir mais significativamente na sua formação e futura atuação profissional. O acoplamento bibliográfico e o co-word também são vantajosos nesse sentido, pois visualizar quais são as frentes de pesquisa e os conceitos emergentes, facilita conhecer os debates atuais e a se informar sobre as lacunas dos estudos já realizados, sendo um ponto essencial para elaborar projetos inovadores. Também é possível realizar uma análise bibliométrica para encontrar as revistas que publicam com maior frequência sobre determinado tema, descomplicando o processo de escolha do melhor periódico para submeter um artigo. Mais além, os bancos de dados bibliográficos têm os endereços eletrônicos dos autores, viabilizando o contato com outros cientistas com afinidades teórico-metodológicas para futuras parcerias científicas.

A nossa aplicação metodológica contribuiu para exemplificar o uso do software VOSviewer na criação de redes com dados de bibliografia. Os indicadores aplicados se mostraram vantajosos para o conhecimento da temática. Identificamos que o perfil de autores que escrevem sobre presidencialismo latino-americano é majoritariamente masculino e proveniente dos Estados Unidos. No entanto, de maneira agrupada é possível dizer que a produção latino-americana é maior em quantidade de artigos no nosso banco de dados, do que a de outras regiões. A presença de coautoria com vários países externos à região demonstra que a temática é de interesse internacional, além de regional. Ficou evidente que nas bases analisadas a relação mais forte de coautoria sobre o tema é entre pesquisadores estadunidenses e brasileiros.

A principal limitação deste estudo é que ele mensura apenas as produções indexadas aos acervos bibliográficos selecionados, não abrangendo exaustivamente todas as produções sobre o assunto. Dentre as limitações da busca bibliográfica, enfatizamos um "viés oculto" na indexação, pois as principais bases científicas apresentam uma porcentagem elevada de periódicos de instituições do hemisfério norte e acaba por reproduzir a desigualdade existente entre a produção científica de países economicamente mais desenvolvidos ou menos (KAU-

CHAKJE, ROSA, 2020). Desta forma, é preciso que estudos futuros também explorem outras plataformas bibliográficas ou adicionem manualmente artigos não contemplados nesta pesquisa. Outra limitação, é que o filtro de seleção dos documentos não se restringiu apenas a estudos publicados em revistas da área de ciência política, por considerarmos que a temática é de interesse multidisciplinar, o que pode interferir na interpretação das transformações da temática para a disciplina de ciência política de maneira isolada.

Como vimos nesse artigo, a bibliometria se ocupa de avaliar a produtividade e a cientometria as relações do campo. Por um lado, esse tipo de estudo é rico em detalhes e proveitoso a nível exploratório. Por outro, houve resistência aos estudos bibliométricos por seu caráter exclusivamente quantitativo, que não permite análises mais aprofundadas dos conteúdos dos documentos (HAYASHI, 2013). À luz disso, surgiu no debate teórico-metodológico uma nova abordagem do conceito, a neo-bibliometria. Essa abordagem apresenta um caminho qualitativo a ser percorrido conjuntamente com os procedimentos quantitativos e essa combinação proporciona tanto uma aproximação, como um delineamento teórico e metodológico dos estudos. Enfatizamos, entretanto, que mesmo que os indicadores bibliométricos sejam de caráter quantitativo, foi fundamental uma aproximação qualitativa das pesquisadoras com o banco de dados para melhor interpretação dos achados. Deste modo, concluímos que a nova roupagem do termo de bibliometria surgiu mais como uma aprimoração transparente do processo de conhecimento das informações, do que como um novo conceito e aplicação.

Reconhecemos também que para ir além dos dados descritivos que apresentamos, seria importante contextualizar os resultados – tanto com a conjuntura da disciplina ou do campo em análise, como do desenvolvimento da ciência e das políticas de financiamento – para melhor interpretação. Isto posto, reforçamos a importância da metodologia para a análise de outras temáticas e vertentes teóricas centrais da área de ciência política como Estado e Governo, Instituições Políticas, Cultura Política, Comportamento Político e Políticas Públicas para uma perspectiva comparada. Como também aconselhamos aprofundar essa pesquisa na busca de compreender a relação entre capacidades estatais e a governabilidade em regimes presidencialistas para produzir uma análise sistemática da literatura. Ou ainda, dar sequência aos estudos sobre gênero na ciência política latino-americana.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRANCHES, Sérgio H. H. (1988), "Presidencialismo de coalizão: o dilema institucional brasileiro", *Dados–Revista de Ciências Sociais*, vol. 31, no. 1: 1-39.

ALTMAN, David. (2000), "The politics of coalition formation and survival in multiparty presidential democracies: The case of Uruguay, 1989-1999", *Party politics*, vol. 6 no.3: 259-283.

BARRIENTOS DEL MONTE, Fernando. (2013), "La ciencia política en América Latina: una breve introducción histórica", *Convergencia*, vol. 20, np. 61: 105-133.

BASABE-SERRANO, Santiago; HUERTAS-HERNANDEZ, Sergio. (2018), "El estado de la investigación en Ciencia Política sobre América Latina", *Revista Española de Ciencia Política-RECP*, no. 47: 153-170.

BENENTE, Mauro. (2017), "Presidencialismo y poder económico en la reforma constitucional Argentina de 1994", *Revista de Derecho*, no. 47: 73-104.

BERSCH, Katherine; PRAÇA, Sérgio; TAYLOR, Matthew M. (2017), "State capacity, bureaucratic politicization, and corruption in the Brazilian state", *Governance*, vol. 30, no. 1: 105-124.

BORGES, André; TURGEON, Mathieu. (2019), "Presidential coattails in coalitional presidentialism", *Party Politics*, vol. 25, no. 2: 192-202.

CAMERON, Maxwell A. (2010), "The State of Democracy in the Andes: Introduction to a thematic issue of Revista de Ciencia Política", *Revista de ciencia política (Santiago)*, vol. 30, no. 1: 5-20.

CANDIDO, Marcia Rangel; FERES JÚNIOR, João; CAMPOS, Luiz Augusto. (2019), "Desigualdades na elite da Ciência Política brasileira", *Civitas-Revista de Ciências Sociais*, vol. 19, no.3: 564-582.

CHAISTY, Paul; POWER, Timothy J. (2019), "Flying solo: Explaining single-party cabinets under minority presidentialism", *European Journal of Political Research*, vol. 58, no. 1: 163-183.

CHÁVEZ, Alicia Hernández. (1994), "Mexican presidentialism: A historical and institutional overview", *Estudios Mexicanos*, vol. 10, no.1: 217-225.

CHEIBUB, José Antonio; PRZEWORSKI, Adam; SAIEGH, Sebastian M. (2004), "Government coalitions and legislative success under presidentialism and parliamentarism", *British Journal of Political Science*, vol. 34, no. 4: 565-587.

DAZA, Javier Duque. (2015), "¿Presidentes legislando? Las facultades legislativas en el presidencialismo colombiano 1991-2014", *Revista Estudios Socio-Jurídicos*, vol. 17, no. 1: 79-123.

FREIDENBERG, Flavia. (2019), "La ceguera de género en la ciencia política latinoamericana", *La gaceta de ciencia política*, no. 1: 26-44. Disponível em: https://gacetadecpol.files.wordpress.com/2019/03/gaceta-2019-final-1.pdf . Acesso em: 22 ago. 2020.

GÄNZEL, Wolfang. (2003), *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. Course Handouts.

GARCÍA, Jean-René. (2014), "¿Presidencialismo o Poder Ejecutivo ambivalente?: De la pertinencia de la noción jurídica para el estudio de los sistemas políticos de América Latina", *Cuestiones constitucionales*, no. 31: 57-79.

GRANOVSKI, Yuri V. (2001), "Is it possible to measure science? V. V. Nalimov's research in scientometrics", *Scientometrics*, vol. 52, no. 2: 127-150

HAYASHI, Carlos. R. Massao. (2013), "Apontamentos sobre a coleta de dados em estudos bibliométricos e cientométricos", *Filosofia e Educação*, vol. 5, no. 2: 89–102. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rfe/article/view/8635396 . Acesso em: 15 nov. 2020.

HAYASHI, Maria Cristina. (2012), "Sociologia da Ciência, Bibliometria e Cientometria: Contribuições para a análise da produção científica", Anais Eletrônico – IV EPISTED – Seminário de Epistemologia e Teorias da Educação, Faculdade de Educação/Unicamp. Disponível em: https://www.marilia.unesp.br/Home/Graduacao/PETBiblioteconomia/soc-da-ciencia-pet.pdf . Acesso em: 20 jun. 2020

HERNÁNDEZ, Rogelio. (1978), "¿Presidencialismo u oligarquía? Algunas reflexiones sobre la estructura política mexicana", *Problemas del Desarrollo*, vol. 9, no. 35:53-85.

HUNEEUS, Carlos. (2012), "Variedades de governos de coalizão no presidencialismo. Chile, 1990-2010", *Dados-Revista de Ciências Sociais*, vol. 55, no. 4: 877-910.

KAUCHAKJE, Samira; ROSA, Maria Arlete. (2020), *Revisão sistemática da literatura: Como Fazer? vol. 1*, Curitiba, Editora UTP.

LARKINS, Christopher. (1998) "The judiciary and delegative democracy in Argentina", *Comparative Politics*, vol. 30, no. 4: 423-442.

LEITE, Fernando; CODATO, Adriano. (2013), "Autonomização e institucionalização da Ciência Política brasileira: o papel do sistema Qualis-Capes", *Agenda Política*, vol. 1, no.1: 1-21.

LENINE, Enzo; MÖRSCHBÄCHER, Melina. (2020), "Pesquisa bibliométrica e hierarquias do conhecimento em Ciência Política", *Revista Brasileira de Ciência Política*, no. 31: 123-160.

LIMA, Iana Alves de. (2018) "O impacto do arranjo institucional brasileiro no controle político sobre a burocracia", *Cadernos EBAPE. BR*, vol. 16, no. 4: 656-666.

LINZ, Juan J. (1990), "The perils of presidentialism", *Journal of democracy*, vol. 1, no.1: 51-69.

MAINWARING, Scott. (1993), "Presidentialism, multipartism, and democracy: The difficult combination", *Comparative political studies*, vol. 26, no. 2: 198-228.

MALAMUD, Andrés. (2015), "Social revolution or political takeover? The Argentine collapse of 2001 reassessed", *Latin American Perspectives*, vol. 42, no.1: 11-26.

MARCIAS-CHAPULA, Cesar A. (1998), "O papel da informetria e da ciênciometria e sua perspectiva nacional e internacional", *Ciência da Informação*, vol. 27, no. 2: 134-140.

MAYER-SERRA, Carlos Elizondo. (2017), "Reforma de la Constitución: la economía política del Pacto por México", *Revista mexicana de ciencias políticas y sociales*, vol. 62, no. 230: 21-49.

MUGNAINI, Rogério; JANNUZZI, Paulo de Martino; QUONIAM, Luc. (2004), "Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal", *Ciência da informação*, vol. 33, no. 2: 123-131.

MUSTAPIC, A. M. (2005), "Instability without collapse. The resignation of the presidents: Argentina in 2001", *Desarrollo Economico-Revista De Ciencias Sociales*, vol. 45, no. 178: 263-280.

NAKAGUMA, Marcos Yamada. (2015), "Choosing the form of government: Theory and evidence from Brazil", *The Economic Journal*, vol. 125, no. 589: 1986-2023.

NEGRETTO, Gabriel L. (2006), "Minority presidents and democratic performance in Latin America", *Latin American Politics and Society*, vol. 48, no. 3: 63-92.

ORTIZ ORTIZ, Richard. (2018), "Los problemas estructurales de la Constitución ecuatoriana de 2008 y el hiperpresidencialismo autoritario", *Estudios constitucionales*, vol. 16, no. 2: 527-566.

PANIZZA, Francisco; PHILIP, George. (2005), "Second generation reform in Latin America: Reforming the public sector in Uruguay and Mexico", *Journal of Latin American Studies*, vol. 37, no. 4: 667-691.

PEÑA, Alejandro Milcíades; DAVIES, Thomas Richard. (2017), "Responding to the street: government responses to mass protests in democracies", *Mobilization: An International Quarterly*, vol. 22, no. 2: 177-200.

PHILIP, George. (1998), "The New Populism, Presidentialism and Market-Orientated Reform in Spanish South America", *Government and Opposition*, vol. 33, no. 1: 81-97.

QUINCHE RAMÍREZ, Manuel Fernando. (2013), "Presidentialism, conventionality control and democracy in Andean countries", *Co-herencia*, vol. 10, no. 19: 167-206.

RABKIN, Rhoda. (1996), "Redemocratization, electoral engineering, and party strategies in Chile, 1989-1995", *Comparative Political Studies*, vol. 29, no. 3: 335-356.

RENIU, Josep Ma; ALBALA, Adrián. (2012), "Los gobiernos de coalición y su incidencia sobre los presidencialismos latinoamericanos: el caso del Cono Sur", *Estudios políticos (México)*, no. 26: 161-214.

RÍOS GÓMEZ, Claudia; HERRERO SOLANA, Víctor. (2005), "La producción científica latinoamericana y la ciencia mundial: una revisión bibliográfica (1989-2003)", *Revista interamericana de Bibliotecología*, vol. 28, no. 1: 43-61.

ROCHA-CARPIUC, Cecilia; MADEIRA, Rafael Machado. (2019), "Desigualdade de gênero, internacionalização e trajetórias acadêmicas na Ciência Política: Evidências no Brasil e no Uruguai", *Civitas-Revista de Ciências Sociais*, vol. 19, no. 3: 545-563.

SAN FRANCISCO, Alejandro. (2003), "La gran convención del partido liberal democrático en 1893: Un hito en la reorganización del balmacedismo después de la guerra civil chilena de 1891", *Historia (Santiago)*, no. 36: 333-377.

SANCHO, Rosa. (1990), "Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Revisión bibliográfica", *Revista española de documentación científica*, vol. 13, no. 3-4: 842-865.

SANGALLI, Amanda F. (2019), "Presidencialismo latino-americano: uma análise descritiva e cientométrica da temática (1978-2019)", Trabalho de Conclusão de Curso (Graduação em Ciências Sociais) – Pontifícia Universidade Católica do Paraná, Curitiba, p.1-71. Disponível em: https://www.academia.edu/43909793/PRESIDENCIALISMO_LATINO_AMERICANO_UMA_AN%C3%81LISE_DESCRITIVA_E_CIENTOM%C3%89TRICA_DA_TEM%C3%81TICA_1978_2019_

SANGALLI, Amanda F. (2020), "Mapeando las producciones sobre el presidencialismo latinoamericano (1978-2019)", Jornadas de Ciencia Política, área temática: política latinoamericana, Facultad de Ciencias Sociales de la Universidad de Buenos Aires, Argentina. Disponível em: https://www.academia.edu/44239452/Mapeando_las_producciones_sobre_el_presidencialismo_latinoamericano_1978_2019_ . Acesso em: 20 set. 2020.

SANTOS, Raimundo Nonato Macedo dos; KOBASHI, Nair Yumiko. (2009), "Bibliometria, cientometria, infometria: conceitos e aplicações", Grupo de Trabalho 7: Produção e Comunicação da Informação em CT&I, Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, vol. 2, no. 1: 155-172.

SHUGART, Matthew Soberg; CAREY, John M. (1992), *Presidents and assemblies: Constitutional design and electoral dynamics*, Cambridge University Press.

SIAVELIS, Peter M. (2016), "Cabinets and Informal Advisory Networks in Multiparty Presidential Systems", *Presidential Studies Quarterly*, vol. 46, no. 3: 569-591.

SILVA, José Aparecido da; BIANCHI, Maria de Lourdes Pires. (2001), "Cientometria: a métrica da ciência", *Paidéia (Ribeirão Preto)*, vol. 11, no. 21: 5-10.

TAGUE-SUTCLIFFE, Jean. (1992), "An introduction to informetrics", *Information Processing & Management*, vol. 28, no. 1: 1-3.

TARTAKOFF, Laura Y. (2019), "Ecuador, Argentina, and Paraguay: At Last Immunity to Authoritarian Rule?", *Society*, vol. 56, no. 1: 15-22.

TEDESCO, Laura; DIAMINT, Rut. (2014), "Latin American Democracy. What to Do with the Leaders?", *Bulletin of Latin American Research*, vol. 33, no. 1: 31-45.

TIMMERMANN, Andreas. (2012), "El concepto de" gobierno moderado" como hilo conductor en el constitucionalismo temprano de Hispanoamérica", *Estudios de historia moderna y contemporánea de México*, no. 44: 3-48.

VAN RAAN, A. F. J. (1997), "Scientometrics: state-of-the-art", *Scientometrics*, vol. 38, no. 1: 205-2018

VANTI, Nadia Aurora Peres. (2002), "Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento", *Ciência da informação*, vol. 31, no. 2: 152-162.

VANTI, Nadia Aurora Peres. (2005), "Os links e os estudos webométricos", *Ciência da informação*, vol. 34, no. 1: 78-88.

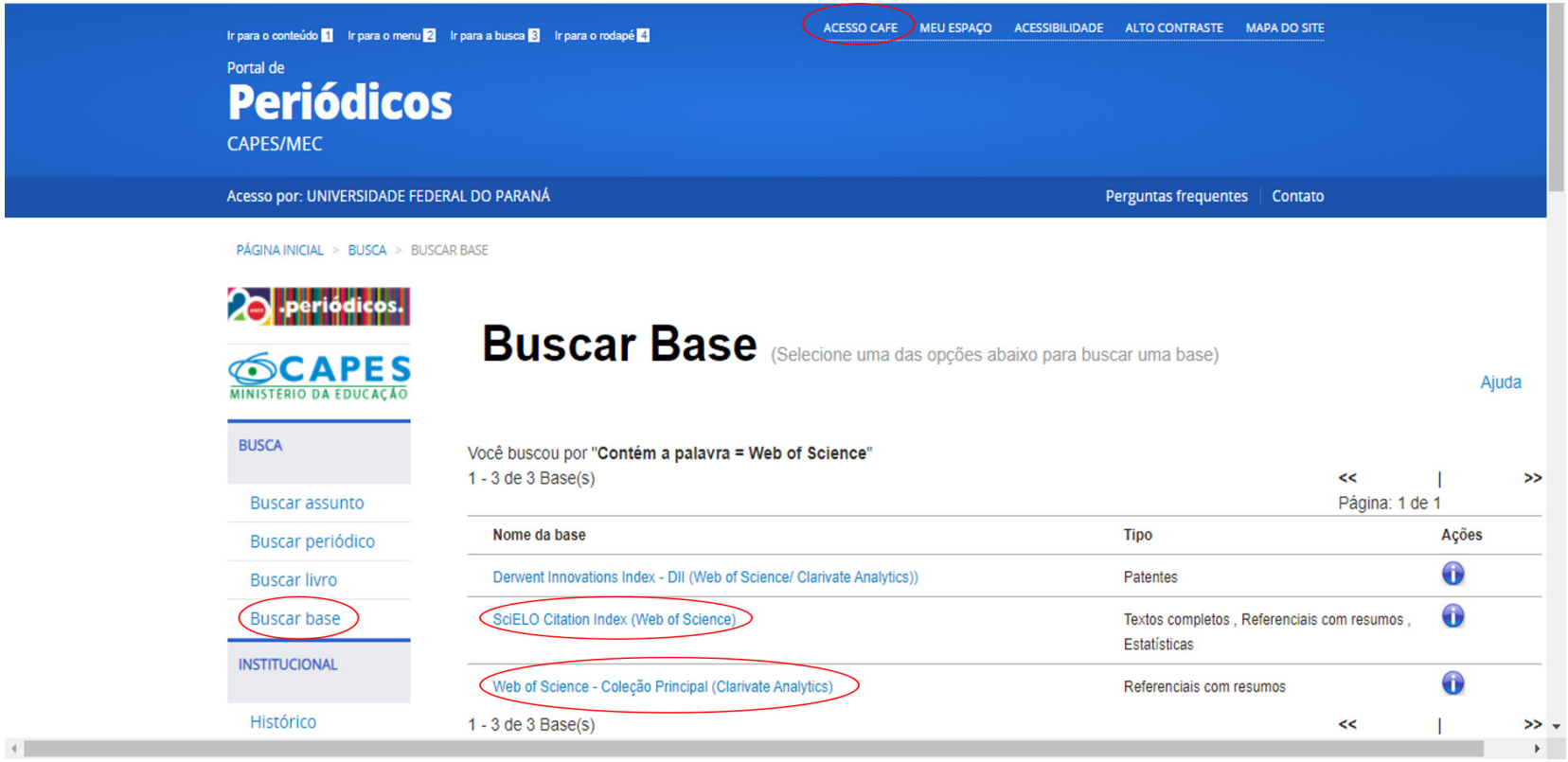
VANTI, Nadia Aurora Peres. (2011), "A cientometria revisitada à luz da expansão da ciência, da tecnologia e da inovação", *Ponto de Acesso*, vol. 5, no. 3: 5-31.

WALTMAN, L.; VAN ECK, N. VOSviewer manual: manual for VOSviewer version 1.6.6, Centre for Science and Technology Studies, Leiden University, The Netherlands, p. 1-48. Disponível em: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.6.pdf . Acesso em: 10 jun. 2019.

ZUPIC, Ivan; ČATER, Tomaž. (2015), "Bibliometric methods in management and organization", *Organizational Research Methods*, vol. 18, no. 3: 429-472.

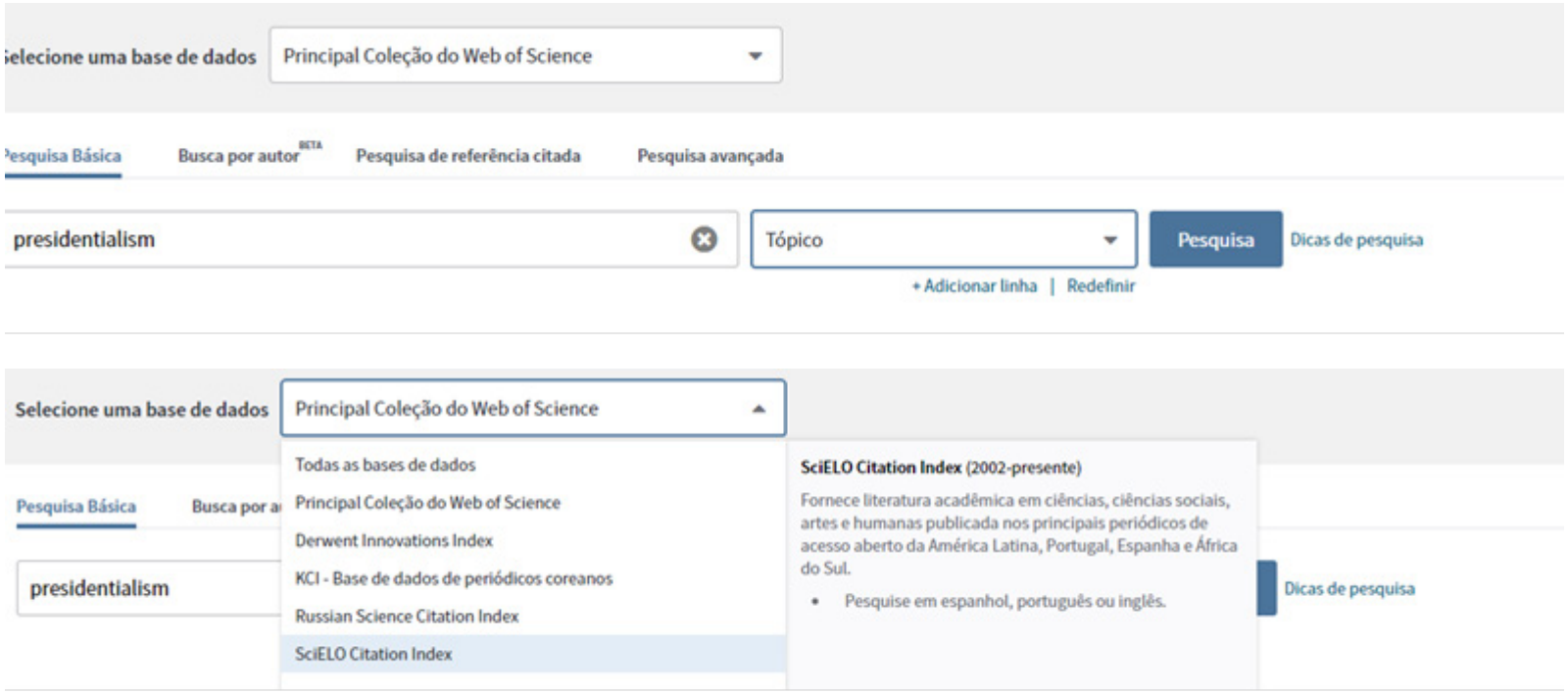
ANEXO 1 – CONSULTA SOBRE O TEMA PRESIDENCIALISMO NAS BASES BIBLIOGRÁFICAS E EXTRAÇÃO DOS DADOS[1]

Caminho para consulta e extração dos dados de bibliografia: acessar o Portal de Periódicos da CAPES/MEC e realizar acesso remoto via CAFe com login institucional, ir em "Buscar base" e buscar por título "Web of Science". Nesse momento encontra-se ambas as coleções que utilizamos.

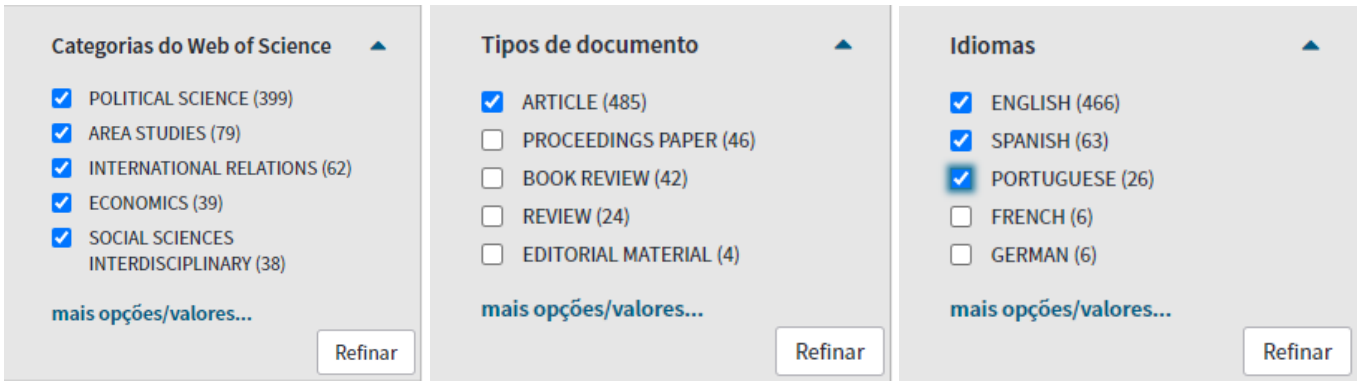


Para a consulta na coleção principal WoS escrevemos em Pesquisa Básica o termo de busca "presidentialism" e realizamos a busca do tipo "Tópico". Isso também foi feito para a coleção SCI.

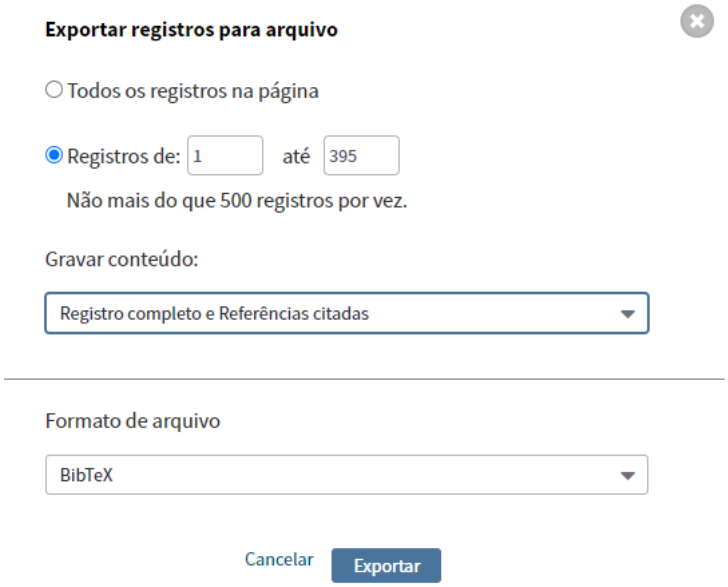
1 As imagens foram capturadas em 24 de março de 2021, após a realização do estudo, apenas para exemplificar nesse artigo introdutório, por isso, os números contidos nas imagens devem ser ignorados, pois não correspondem à busca realizada em março de 2019.



Após a busca, filtramos os resultados por "Categorias do Web of Science (ou da SciELO)", "Tipo de Documento" e "Idiomas".



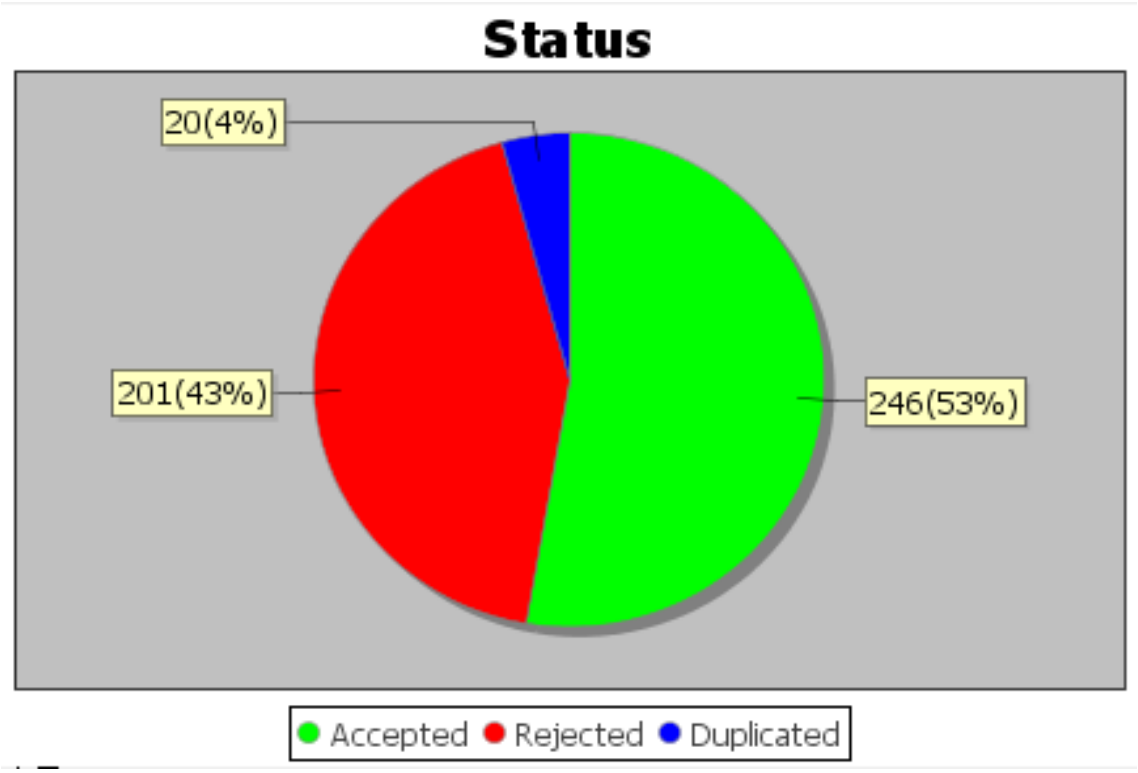
Por fim, foi extraído o "Registro completo e Referências citadas" de todos os documentos selecionados.



ANEXO 2 – PROTOCOLO START APLICADO À ANÁLISE DETALHADA DOS RESUMOS PARA SELEÇÃO DE ARTIGOS SOBRE O TEMA PRESIDENCIALISMO LATINO-AMERICANO

StArT Protocol

Title:	Uma introdução à bibliometria e cientometria: o caso do presidencialismo latino-americano
Researchers:	Amanda Sangalli; Samira Kauchakje.
Description:	Utilizar o StART para leitura detalhada dos resumos dos artigos selecionados em etapa anterior na coleção principal Web of Science e na coleção SciElo Citation Index. Os artigos coletados em ambas as bases têm como foco o tema presidencialismo, o StArt será utilizado apenas como um instrumento facilitador para a leitura dos resumos para refinar a nossa seleção, isto é, aplicar o recorte regional. Para um artigo seguir para a etapa de mapeamento, é necessário que cumpra os critérios de inclusão.
Objective:	Selecionar os artigos com enfoque analítico no(s) presidencialismo(s) latino-americano(s)
Main question:	Como mapear a produção científica sobre os países presidencialistas da América Latina?
Keywords and Synonyms:	presidentialism
Sources Selection Criteria Definition:	Principal base científica internacional e principal base regional (América Latina)
Studies Languages:	Português, espanhol e inglês.
Source Search Methods:	Busca manual na coleção principal Web of Science e na coleção SciELO Citation Index do tipo "tópico" utilizando o termo "presidentialism"; e, exportação dos metadados do tipo "registro completo e referências citadas" dos documentos em formato BibTex e TXT.
Source list:	Web of Science; SciELO Citation Index.
Study selection criteria: (Inclusion and exclusion criteria)	Inclusão: a) É um estudo sobre o presidencialismo; b) É um estudo que compara o presidencialismo com outros sistemas de governo. c) É um estudo de um ou mais países da América Latina; d) É uma análise comparada de países presidencialistas latino-americanos com outros países de outras regiões. Exclusão: a) Não é um estudo sobre presidencialismo. b) É sobre presidencialismo, mas não é sobre países da América Latina; c) É sobre países latino-americanos, mas não são estudos sobre o sistema de governo presidencialista.
Studies Types Definition:	Artigos publicados em revistas/periódicos científicos
Studies Initial Selection:	A partir dos títulos, palavras-chave e da leitura dos resumos
Studies Quality Evaluation:	Não se aplica
Data Extraction Form Fields:	Text



Como as bases bibliográficas não estavam padronizadas, foi fundamental o processo de exclusão de artigos duplicados (que constavam em ambas as coleções). Após selecionar via StArt realizamos outra checagem manual utilizando o Excel. Ordenamos os artigos por ano de publicação fazendo uma revisão se ainda havia duplicatas. Em alguns casos, o artigo estava indexado na coleção principal em inglês e na coleção SCI em espanhol ou português e o StArt não o identificou como duplicado, por isso foi necessária uma última checagem manual dos artigos selecionados. Assim, dos 246 artigos aceitos, 244 foram mantidos para a análise. Os 201 rejeitados não estudavam países presidencialistas da América Latina e sim de outras regiões.

ANEXO 3 – REVISTAS CONTEMPLADAS NA ANÁLISE E SUAS RESPECTIVAS ÁREAS, CLASSIFICADAS COM BASE NO ESCOPO DE SEUS WEBSITES (Busca realizada na WoS e SCI em 23 de março de 2019)

Ciência Política e Relacões Internacionais	**Ciencias Sociales (Interdisciplinar)**	**Direito**
American Journal of Political Science	América Latina Hoy	Revista Derecho del Estado
American Political Science Review	Archives Europeennes de Sociologie	Estudios constitucionales
Analecta Politica	Bulletin of Latin American Research	Revista de estudios histórico-jurídicos
Annual Review of Political Science	Co-herencia	Revista de Investigações Constitucionais
Brazilian Political Science Review	Dados - Revista de Ciencias Sociais	Cuestiones constitucionales
British Journal of Political Science	Daedalus, the Journal of the American Academy of Arts & Science	Revista chilena de derecho
Colombia Internacional	Desarrollo Economico	Revista de Derecho
Comparative Political Studies	E-LATINA – Revista Electronica de Estudios Latinoamericano	e-Pública: Revista Eletrónica de Direito Público
Comparative Politics	Espiral (Guadalajara)	Revista de Derecho Publico
Contemporary Politics	Iberoamericana	Estudios Socio-Jurídicos
Democratization	Journal of American Culture	**Administração Pública**
Electoral Studies	Journal of Interamerican Studies and World Affairs	Cadernos de Gestao Publica e Cidadania
Estudios Políticos	Journal of Latin American Studies	Governance: An International Journal of Policy, Admnistration and Institutions
European Journal of Political Research	Latin American Perspectives	Cadernos EBAPE.BR
Government and Opposition	Latin American Politics and Society	Saúde em Debate (Políticas Publicas)
International Political Science Review	Latin American Research Review	Revista de Administração Pública
International Studies Perspectives	Lua Nova: Revista de Cultura e Política	**Economía**
Journal of Democracy	Mexican Studies - Estudios Mexicanos	Revista de Economía Institucional
Journal of Politics	Mobilization: an international quarterly journal	Economic Systems
Journal of Politics in Latin America	Novos estudos CEBRAP	Journal of Economic Integration
Papel Politico	Opinião Pública	Journal of the European Economic Association
Parliamentary Affairs	Polis (Santiago)	The Economic Journal
Party Politics	Política y cultura	Problemas del Desarrollo

Ciência Política e Relacões Internacionais	Ciências Sociais (Interdisciplinar)	História
Policy Studies	Public Choice	HiSTOReLo. Revista de Historia Regional y Local
Politica y Gobierno	Regional and Federal Studies	Tzintzun Revista de Estudios Historicos
Politica y Sociedad	Religación - Revista de Ciencias Sociales y Humanidades	Historia Mexicana
Political Research Quarterly	Revista Brasileira de Ciências Sociais	Estudios de historia moderna y contemporánea de México
Politologický časopis - Czech Journal of Political Science	Revista Española de Investigaciones Sociológicas	Historia (Santiago)
Postdata	Revista Mexicana de Ciencias Políticas y Sociales	
Presidential Studies Quarterly	Social Science Journal	
Relações Internacionais (R:I)	Society	
Research and Politics	The Revista CS	
Revista Brasileira de Ciência Política		
Revista de Ciencia Politica (Santiago)		
Revista de Sociologia e Política		
Revista SAAP		
Revista Uruguaya de Ciencia Política		
Scandinavian Political Studies		
Southeast European and Black Sea Studies		
Temas y Debates		
The British Journal of Politics and International Relations		
The Journal of Politics		
World Politics		

Fonte: Sangalli (2020)

An Introduction to Bibliometrics and Scientometrics: the case of Latin American presidentialism

Amanda Sangalli - UFPR [1]
Samira Kauchakje - UDESC, UFPR [2]

Resume

Introduction: *What are they, what are they for and what is the applicability of bibliometrics and scientometrics in political science? In this article we present its differences, advantages and limitations from the literature and the application practice.* ***Method****: Through the Web of Science and SciELO Citation Index bibliographic collections, 244 scientific articles were selected with an analytical focus on the presidentialism of Latin American countries and built bibliometric networks with VOSviewer.* ***Results:*** *The applicability was demonstrated with the measurement of co-authorship, co-citation, bibliographic coupling and co-occurrence of words. They allowed to identify, respectively, the internationalization of the study agenda, the fundamental scientific influences of the theme, the research fronts and the concepts currently debated.* ***Discussion:*** *The main advantages of the methodology are: i) reliability: due to its transparency and ease of replication; and ii) breadth: due to the ability to analyze multiple dimensions of science. It is necessary to note that the method presents an approximation of reality, but it has limitations because it is restricted to the "bias" of indexing existing on the platforms, which results in the proportionally greater number of articles published in English-speaking countries with incentives for research and scientific publications. It is possible to manipulate other bibliographic collections or to insert texts manually for greater precision. One difficulty faced was the manipulation of metadata, which is not always standardized and can vary between different collections.*

Keywords: Bibliometrics; scientometrics; bibliometric indicators, presidentialism; political science.

1 Graduada em Ciências Sociais (PUCPR, UNSAM). Mestranda em Ciência Política pela Universidade Federal do Paraná (UFPR). Pesquisadora do Laboratório de análise do Campo Científico (LaCC/UFPR).

2 Cientista política. Graduada em Ciências Sociais/área Ciência Política (UFPR); Mestra em Ciência Política (UFPR); Doutora em Ciência Política (UFSCAR); Doutora em Educação (UNICAMP). Professora no departamento de Governança Pública da Universidade do Estado de Santa Catarina (UDESC) e professora colaboradora no Programa de Pós-Graduação em Ciência Política da Universidade Federal do Paraná (UFPR).

Resumo

Introdução: O que são, para que servem e qual a aplicabilidade da bibliometria e da cientometria na ciência política? Nesse artigo apresentamos suas diferenças, vantagens e limitações a partir da literatura e da prática de aplicação. **Método**: Por meio das coleções bibliográficas Web of Science e SciELO Citation Index, foram selecionados 244 artigos científicos com enfoque analítico no presidencialismo de países latino-americanos e construídas redes bibliométricas com o VOSviewer. **Resultados:** A aplicabilidade foi demonstrada com a medição de coautoria, cocitação, acoplamento bibliográfico e coocorrência de palavras. Permitiram identificar, respectivamente, a internacionalização da agenda de estudos, as influências científicas fundamentais da temática, as frentes de pesquisa e os conceitos debatidos atualmente. **Discussão:** As principais vantagens da metodologia são: i) confiabilidade: por sua transparência e facilidade de replicação; e ii) amplitude: pela capacidade de analisar múltiplas dimensões da ciência. É necessário atentar-se que o método apresenta uma aproximação da realidade, mas esta tem limitações por se restringir ao "viés" da indexação existente nas plataformas, o que acarreta o número proporcionalmente maior de artigos publicados em países de língua inglesa com incentivos à pesquisa e às publicações científicas. É cabível manipular outras coleções bibliográficas ou inserir textos manualmente para maior precisão. Uma dificuldade enfrentada foi a manipulação de metadado, que nem sempre estão padronizados e podem variar entre diferentes coleções.

Palavras-chave: Bibliometria; cientometria; indicadores bibliométricos; presidencialismo; ciência política.

1. Introdution

The exponential growth of world scientific production made it impossible for researchers to know it in its entirety. Although scientific articles have the role of presenting innovations in a particular field of knowledge and are a quick vehicle for communication between scientists, the volume of recorded information makes it difficult to systematize the literature and understand scientific and technological advances (GRANOVSKI, 2001). The relationship between working time and amount of information requires the researcher to prioritize automated procedures for compiling bibliographies. In this way, it makes it possible to map knowledge and direct reading by ranking documents according to their relevance in the community or their actuality.

Bibliometrics and scientometrics are procedures dedicated to measuring the development of science – whether in an area of knowledge, discipline or theme. They are especially used by public and private research institutions as a means of evaluating and financing scientific production. This is not the only use, researchers interested in the evolution or structuring of science, that is, who are dedicated to understanding its socio-historical trajectory, also operate such methodological strategies (SILVA, BIANCHI, 2001). We also punctuated the usability for designing careers, as it allows us to identify which are the main production centers on a given research object, facilitating the finding of partnerships and the planning of which paths to follow throughout the career. Both make it possible to know, not only what has been produced on a subject in past decades, but also what are the emerging themes in the literature and the research gaps.

The article is divided into three sections plus the conclusion. The first section is a brief methodological review. We present what they are, what they are for and what are the differences between bibliometrics and scientometrics, as well as the definitions and applications of some bibliometric indicators. In the second section, we apply the methodology to the study of a topic of interest to political science: Latin American presidentialism. We explain the procedures for building a reliable and replicable database, as well as the protocols that a search of this type needs to follow. Thirdly, we discuss the results obtained with this study, emphasizing the applicability of the method, so we present a brief analysis of the topic. In conclusion, we discuss the obstacles faced in this research, with recommendations and warnings for future studies.

2. Bibliometrics and scientometrics: what are they and what is the difference?

The field of Information Science (CI) began in the 19th century, from an approximation of the areas of Sociology of Science and Social History of Science, but it was only in the 20th century that the studies of information metrics were systematized. From the demand to understand scientific activity as a social phenomenon (HAYASHI, 2012), the classic concepts of bibliometrics and scientometrics originated. There is a vast theoretical discussion about these concepts and their applications, making it difficult to delimit the spaces of these subfields. Not being ideal for this article to exhaust this historical clash [3], we will make a theoretical discussion based on the definitions of Tague-Sutcliffe (1992) and Van Raan (1997).

3 For further details on the historical debate, see the bibliographic review by Sancho (1990) and the conceptual debate by Vanti (2002) and Hayashi (2012)

Table 1. Bibliometrics and scientometrics: differentiation of purpose, objects of study, variables, measures and objectives.

field of study	theoretical basis	Study objects	Variables	measurements	Goals
Bibliometrics	It is a field of study that develops mathematical models and measures to quantitatively measure the production, dissemination and use of recorded information. Its application provides a solid basis for decision making and forecasting (TAGUE-SUTCLIFFE, 1992).	Printed records, books, documents, magazines, scientific articles, authors and users (VANTI, 2005).	Number of hits (circulation) and citations, frequency of words, length of sentences (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998).	Ranking, frequency, distribution (MACIAS-CHAPULA, 1998)	Allocate resources: people, time, money; understand the production and dissemination of knowledge (adapted from VANTI, 2002)
Scientometrics	Scientometric analysis relates quantitative aspects to social and economic contexts in order to interpret advances in scientific development. It is based not only on mathematical methods, but also on sociological and behavioral ones (VAN RAAN, 1997).	Areas of knowledge, scientific and technological fields, professionals in the same field, disciplines, subjects, patents, dissertations and theses (VANTI, 2005).	Factors that differentiate disciplines and disciplinary subfields; how scientists communicate (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998).	Analysis of sets; correspondence analysis; co-occurrence of terms, expressions, keywords; citation and co-citation of documents, authors, institutions, countries; bibliographic coupling (adapted from VANTI, 2002)	Identify domains of interest, where studies on a given subject are concentrated; understand how and how much scientists communicate (MCGRATH, 1989 apud MACIAS-CHAPULA, 1998)

Source: elaboration of the authors from the literature.

In a theoretical sense, we can define bibliometrics as a field of study dedicated to the development of techniques to systematize bibliographic information. Confusion in the interpretation of the concept is common, being considered by the literature, sometimes methodology, sometimes technique. According to Gänzel (2003), there are three groups interested in the application of bibliometrics: bibliometricians, those interested in scientific information and those responsible for the management and elaboration of scientific policies. The first group carries out methodological development research in the field of CI. That is, they consider it a fundamental part in the execution of a research. The second group is multidisciplinary, composed of researchers from different areas, and seeks to employ bibliometrics techniques – bibliometric indicators – to obtain information on a given subject. It is in this group that this article fits when dealing with knowing and mapping the theme of Latin American presidentialism. The third group also uses indicators, but it differs from the previous group in terms of objective, as it has the purpose of evaluating science and thinking about development policies. According to this differentiation between groups, we can assume that, in a practical sense, bibliometrics is a methodology for information scientists, at the same time it is a set of techniques for researchers from other areas of knowledge.

While bibliometrics is fundamentally quantitative, adhering to scientometrics as a method is to adopt methodological strategies in both a quantitative and qualitative style for the collection and interpretation of data. In addition to bibliometric indicators, other methods applied in scientometric analysis are: the use of statistics, sociological network models, psychological surveys and other interview methods (VAN RAAN, 1997). Because they are concerned with the dynamics of science, its social activity, circulation and consumption (SANTOS, KOBASHI, 2009), philosophy and sociology of science and linguistics are often also present in this type of analysis (VAN RAAN, 1997).

2.1 bibliometric indicators

In light of the theory presented, we assume that bibliometrics is the quantitative analysis of bibliographic data and scientometrics the analysis of relations in the scientific field, with a point of convergence between them, which is the study of bibliographic documents and the use of bibliometric indicators.

Indicators are applied to cases with a large volume of bibliographic metadata. They can be used for performance analysis, measuring the amount of production of a group of researchers, journals, institutions or countries, as well as for scientific mapping and understanding of the dynamics of communication in a given field of knowledge (ZUPIC, ČATER, 2015). There are several indicators that can be applied in bibliographic analysis (SANCHO, 1990) [4], the ones selected for this methodological exercise were the techniques of co-citation, bibliographic coupling, co-authorship and co-word – presented in table 2 with their functions. If their limitations are understood and applied according to their specific functionalities, they are useful to understand the process of creation, reproduction and diffusion of science (MUGNAINI, JANNUZZI, QUONIAM, 2004).

4 Sancho (1990) describes the usability of the following indicators: number and distribution of publications; authors' productivity; collaboration in publications; connections between documents and authors; joint references or citations between documents; number of citations (impact of works); journal impact indicator; immediacy index (speed with which documents are cited); influence of magazines; thematic associations and co-occurrence of terms.

Table 2. Description, advantages and disadvantages of the bibliometric indicators selected for the case study.

Indicator	Unit of analysis	observation unit	objective	measure	Explanation	Advantage	Disadvantage
quotation	Documents	joint citation	Connect documents based on the joint appearance of references in a given document.	Relevance	It serves to identify the scientific foundations. It finds the fundamental references of a certain field of knowledge. It is a sophistication of the citation technique, because in addition to estimating the influence in the field (number of citations), it connects the units according to their similarities.	Filters the most important jobs in a given area.	Not ideal for mapping recent searches. Citations take time to accumulate, so new publications cannot be connected directly, but through clusters and knowledge base. In addition, when analyzing the co-citation of authors, only the information of the first author of the publication appears, making the analysis of author networks difficult.
bibliographic coupling	Documents	shared references	Connect documents based on the number of references they share with each other.	Resemblance	IT'S similar to co-citation, but inversely proportional. It serves to identify, based on the fundamental bibliography, which are the current references that cite this bibliography, that is, it presents the research fronts in a given field. With this technique it is possible to measure the approximation between institutions and countries, which can show whether the research agendas are local, regional or global.	No cumulative citations required. Can be used for new publications not yet cited, emerging fields and smaller subfields.	It can only be used for a limited time (up to an interval of five years). It does not inherently identify the most important works by citation count as co-citation; it is difficult to know whether the mapped publications are important or not.
co-authorship	countries	Collaboration	connects countries based on the number of co-authorships in publications.	internationalization	Along the same lines of identifying the breadth of agendas, the co-authorship technique can effectively show the internationalization of agendas, collaborations between institutions and the invisible colleges of authors. The connection occurs through explicit collaborations such as co-authorship on documents.	It allows to identify the academic collaborations and the social structure of the field.	Collaboration on works is not always recognized as co-authorship.
co-word	terms/words	Frequency of terms/words in titles and abstracts	connects terms by the number of occurrence simultaneously in titles and abstracts.	themes, theoretical or methodological affinities	Also known as co-occurrence of terms, this indicator detects the central terms of the works contained in the bibliographic database and relates them according to the frequency in which they appear together in the documents. Its main utility is to identify the dynamics of relationship or organization of literature into thematic groups and to find central concepts. When applied longitudinally, this technique is able to demonstrate agenda transformations, showing which concepts were frequent in past years and which are frequent today.	Uses the actual content of documents for analysis (other methods only use bibliographic metadata).	Words can appear in different ways and can have different meanings.

Source: adapted from Zupic and Čater (2015).

3. The application of the methodology: the theme of Latin American presidentialism

3.1 Building a bibliographic database[5]

The process of compiling bibliographies requires caution in choosing the bibliographic bases for consultation and in selecting documents for analysis. We will divide this topic into two steps. In the first one, we justify the choice of bases, the consultation of documents and the search filters. In the second we expose the criteria for inclusion and exclusion of documents.

STAGE 1 – We chose to analyze a topic that the discipline of political science – specifically, but not exclusively – is dedicated to understanding: presidentialism. It is essential that the consultation bases are contemplative for the discipline, so it is advisable to browse through several bases before choosing which one to use. It is also possible, and sometimes essential, to merge bases, as will be the case presented in this article. However, it is necessary to take into account that this will require more time from the researcher, because the bibliographic data will not always be standardized and with the same ordering of the variables. This will require adjustments, therefore, longer dedication time.

For our case, we compiled the scientific articles in the main *Web of Science* (WoS) [6]collection and in the SciELO *Citation Index (SCI) collection* [7]. Both are indexed on the WoS platform, but are separate for consultation . The first, produced by the United States, is a reference base for searching bibliography, known for having a set of documents with high scientific impact. Its collection contains articles with the highest number of citations, the best rated journals and presents a great diversity of research areas, including political science and related areas, so it is suitable for this study. SciELO, developed in Brazil, is a cooperative network aimed mainly at Latin American, Hispanic and Lusitanian scientific dissemination, in Portuguese and Spanish. We selected this collection because the research focus is of regional interest (Latin America), so documents that were not included in the main collection could be included in the SCI collection.

As it is a thematic area with terminological consensus, the search term on the platforms was " *presidentialism* ". Those who write about presidentialism usually use this term in the title, abstract or keywords (TRP) of the scientific article, which makes it easier to locate them in the bibliographic databases. Generally, at this stage it is necessary to formulate a search string. This must be comprehensive, in order to include the largest number of documents related to the topic of study, but objective, to the point of excluding those that are outside the theme. That is why it is recommended to study the Boolean search logic, a technique that makes it possible to combine terms using operators such as AND, OR, NOT, among others, to limit the search in documents. The choice of search term(s) is fundamental for this type of research and requires a reasonable prior knowledge of key words or concepts of the topic or disci-

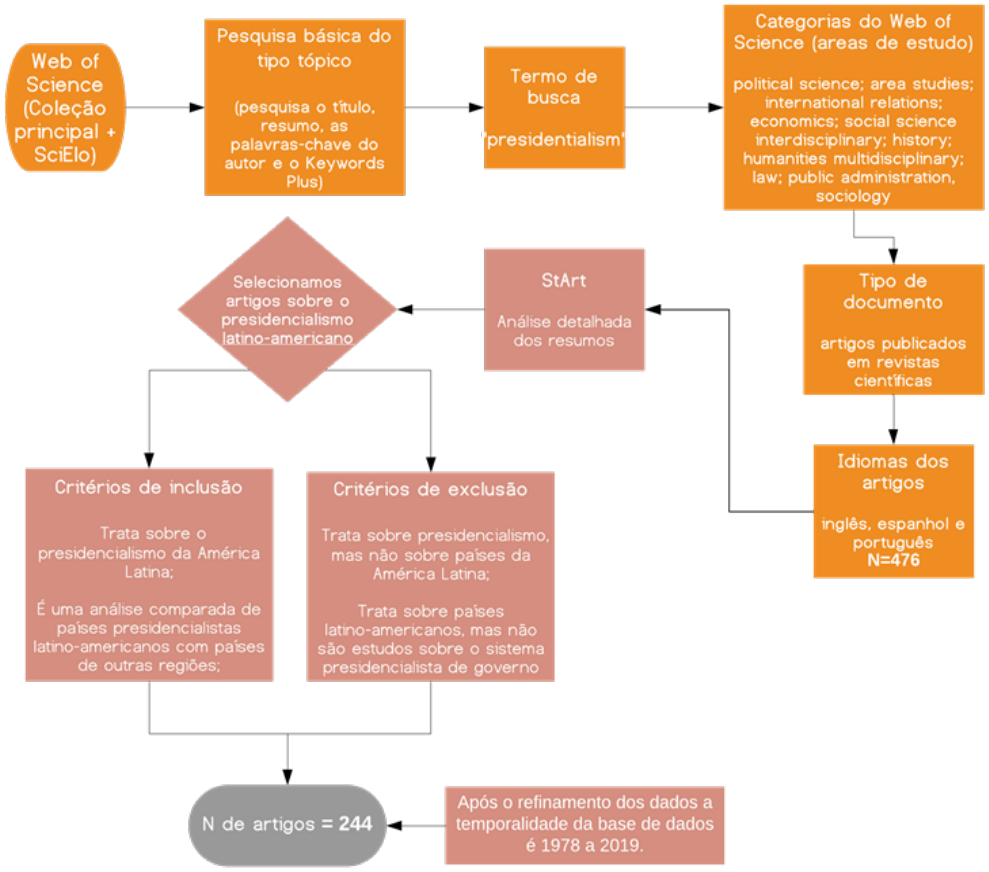
5 The entire methodological part presented here is also included in the course conclusion work by Sangalli (2019), as well as some results, which are present in this article only as an example of application of the methodology and not for an effective discussion of the subject in question.

6 "The Web of Science Core Collection is the world's leading citation database. It contains records of articles from the most impactful journals around the world—including open access journals—conference proceedings and books. Coverage of some securities dates back to 1900." Available at: https://images.webofknowledge.com/WOKRS517B4/help/en_BR/WOS/hp_whatsnew_wos.html

7 "The SciELO Citation Index (Scientific Electronic Library Online) was designed to facilitate the discovery, use and attribution of academic literature in the sciences, social sciences, arts and humanities published in the main open access journals in Latin America and the Caribbean as well as Spain, Portugal and South Africa." Available at: http://images.webofknowledge.com/WOKRS5132R4.2/help/pt_BR/SCIELO/hp_whatsnew_scielo.html

pline to be studied [8]. The fact that the term chosen is in English is because, in addition to the original language of the article, the TRPs are also indexed in English on the platforms, making it possible to find texts in different languages.

Figure 1. Database construction flowchart – presidential theme. Platform searches performed on March 23, 2019 .



Source: Sangalli (2019)

In figure 3 we present a flowchart of each step of selection of articles from WoS and SCI. It was selected from the first publication on the subject of presidentialism recorded in the collections. First we collected the main WoS collection and then applied the same procedure to the SCI collection (see Annex 1).

We filtered the search by selecting areas that could contain publications of interest to this research, which are areas related to political science. The applied base filters were: *political science; area studies; international relations; economics; interdisciplinary social science; history; multidisciplinary humanities; law; public administration; and, sociology.* Another filter applied was language. We only selected articles in English, Spanish and Portuguese, as they are the most frequent and the researchers' domain [9]. At this time of filtering, we obtained 476 articles, adding the two collections (figure 3).

8 "Preparing a search strategy, as well as selecting a database or database to be consulted to answer a specific question, may require certain knowledge from the user (languages, database search engines, tools to help identification of terminology, Boolean logic and other available resources). The implementation of the search strategy requires, therefore, technical knowledge and specific content for its operationalization" (SAKS, 2005, p.2).

9 There were seven articles in other languages, which we read their abstracts on the platform itself and identified that they did not have presidentialism in Latin America as their object of study - our focus of interest - so they were manually excluded, leaving only articles written in the three previously mentioned languages.

The first article on presidentialism found is indexed in WoS, dated in 1973, and the most recent articles in the database were published in March 2019 – the month in which the data were collected. It is always important to mention when the data was collected for greater transparency of the research, after all, the bases are constantly updated, which can directly influence the number of documents found by future researchers who are dedicated to replicating or proceeding with the research.

STAGE 2 – After downloading the database on presidentialism, the next procedure was to refine the selection by applying the cut of studies on Latin America. We performed a detailed reading of the abstracts of the articles selected in step 1 to select only those that analyzed Latin American countries. We used the *State of the Art through Systematic Review* (StArt) tool in this process. It is useful for designing systematic reviews, because it requires a protocol for planning, carrying out and summarizing the research (see annex 2). However, we only use it to read summaries and exclude duplicate documents. When it was not possible to identify in the abstracts whether Latin America was of interest, we performed a dynamic reading of the article to classify it. In this analysis, we used the following criteria for inclusion and exclusion of articles (Table 1):

Table 3. Criteria for inclusion and exclusion of texts on presidentialism to refine the database.

Nature of the criterion	Inclusion	Exclusion
Thematic	a) It is a study of presidentialism; b)It is a study that compares presidentialism with other systems of government.	a) It is not a study of presidentialism.
Regional	c) It is a study of one or more Latin American countries; d) It is a comparative analysis of Latin American presidential countries with other countries in other regions.	b) It's about presidentialism, but it's not about Latin American countries; c) It's about Latin American countries, but it's not about the presidential system of government.

Source: Sangalli (2019)

Only those whose central discussion were the presidential countries of Latin America were chosen, leaving 244 articles for bibliometric analysis. The initial temporality changed to 1978 when the regional cut was applied and remained until March 2019.

We found that the journals that published them belong to six different areas (Table 1). For this classification, we read the scope of each journal on their websites. They are divided between specific journals, such as: Political Science and International Relations; Right; Economy; Public administration; It is history; and interdisciplinary journals considered as Social Sciences – they are journals that allow publications of Sociology, Anthropology, Political Science and History – in their scope.

The articles are distributed in 101 scientific journals (table 1 and annex 3), 58% of them are in journals with a scope directed to the area of Political Science and International Relations, 28% in interdisciplinary journals of Social Sciences, 7% in journals of Law and 7% in other areas (Public Administration, Economics and History).

Table 1. Area of scientific journals in the database, number of journals in each area and total number of articles published in each area on the topic of Latin American presidentialism. Wos and SCI database, search performed on March 23, 2019.

Areas	Magazines		articles	
	No	%	No	%
Political Science and International Relations	43	43	141	58
Social Sciences (Interdisciplinary)	32	32	67	28
Right	10	10	18	7
Public administration	5	5	7	3
Economy	6	5	6	two
History	5	5	5	two
TOTAL	101	100	244	100

Source: Sangalli (2020)

As expected, discussions on presidentialism are mostly published in Political Science and International Relations (n=141 articles) and Social Sciences (n=67) journals, given their relevance to the field as presented by Basabe-Serrano and Huertas Hernández (2018). But we also find publications in Law, specifically in the analysis of Latin American constitutions (GARCÍA, 2014; DAZA, 2015; BENENTE, 2017; ORTIZ ORTIZ, 2018); in Public Administration, for the interest in discussing state capabilities and bureaucracy (BERSCH, PRAÇA, TAYLOR, 2017; LIMA, 2018); in Economics, especially on *accountability* (NAKAGUMA, 2015); and in History, with discussions in a historical perspective about the government system (SAN FRANCISCO, 2003; TIMMERMANN, 2021).

3.2 Two-dimensional visualization of data

The *VOSviewer software* enables the creation, visualization and exploration of data in network map format. It allows viewing a large amount of items or metadata, being able to generate two-dimensional maps. We consider this tool advantageous because it is free and because it is easy to use. However, it is more suitable for those who are starting to contact this style of analysis, as it does not present detailed options about the distance measurements of network connections. For more metric details we recommend using the *Gephi software* . Despite this, *VOSviewer* is efficient in building proximity maps between authors, journals, research institutions and countries. Connections are made by the number of links [10]or mutual citations, which are represented by edges, where their thickness reflects the strength of the relationship between two items. The smaller the distance between items plotted in two-dimensional space, the greater the relationship between them. According to the intensity of these links, the conformation of *clusters occurs* , that is, groups of items with common aspects that are represented by the same color.

There are three types of visualizations provided by VOSviewer for maps: by network, by overlay and by density (*network visualization* , *overlay visualization* , and the *density visualization* , respectively). The first view presents items by label, color pattern and a "node", with the label and node size being the weight of each item. An item's color is determined by the *cluster* it belongs to. The edges between items represent links, so the closer the items are in

10 "A link is a connection or relationship between two items. Examples of links are bibliographic coupling links between publications, co-authorship links between researchers, and co-occurrence links between terms. A map usually includes only one type of link" (VAN ECK, WALTMAN, 2017, p. 4)

two-dimensional space, the stronger their relationships. The second view is identical to the previous one, except for its coloring. There are two ways of clustering, by citation (major/minor) or by year of publication (old/current). A color legend indicates how items are mapped. The third view has two possible options, the *item density* , establishes the density of the items at each point on the map, which is also indicated by colors, however, it is defined by the weight and number of neighboring items. And the density per cluster (*cluster density*), where the color of a point is defined by the mixture of neighboring clusters and the weight is assigned to the color, according to the number of items belonging to the cluster that are in the neighborhood margin (VAN ECK, WALTMAN, 2017) [11].

The software allows you to create maps based on bibliographic data from specific database formats, namely: Web of Science, Scopus, PubMed, RIS and Crossref JSON. Thus, the data collected from the SCI collection needed a treatment so that they could be recognized. The order of variables was standardized using a Microsoft Excel spreadsheet and data from the SCI were manually incorporated into the WoS collection base. After that, the file was saved in text document format for reading in the software.

11 Another visualization function is the map layout and clustering techniques. Among the techniques we have the methods of normalization, strength of association, fractionation and LinLog/modularity, the three correct the strengths of the links between the items for a better two-dimensional visualization. To understand their differences, see Van Eck and Waltman (2009, 2017). It is also possible to modify the layout using the attraction and repulsion mechanisms, which only serves to improve the visualization of the two-dimensional map and does not change its result, as the repulsion value must always be lower than the attraction value. More information on handling the VOSviewer is available in the manual prepared by Van Eck and Waltman (2017).

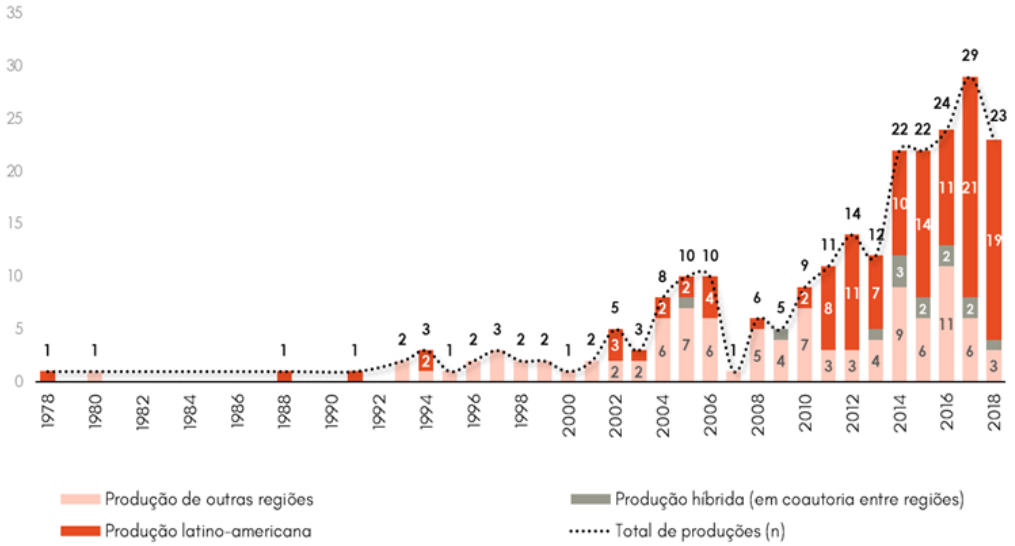
4. Results

4.1 Production and collaborations: countries, gender and co-authorship

The number of articles published on the topic under analysis has been increasing over the years, even though it fluctuates (Graph 1). From 1978 to 1991 we found only 4 articles on the subject on the platforms. From 1992 to 2003, an average of 2.16 articles were published per year, totaling 26 articles. In 2004 the number of publications increased, from this year to 2013 the average is 8.6 articles per year. From 2014 to 2018 is when we have the largest number of publications (n=120), in recent years the average has increased to 24 articles.

When we examined the institutional origin of the authors of the articles, we noticed that, in the bases in question, there was a predominance of studies originating from countries outside Latin America until 2010 (graph 1), which we will call Latin American studies. There are few Latin American productions so far. This can be explained by the fact that political science in the region is young, since the structuring of graduate studies took place in the 1980s (BARRIENTOS DEL MONTE, 2013) and only after its institutionalization and autonomy, around the 2000s (LEITE , CODATO, 2013), is that the discipline provided incentives and investments for the publication of scientific articles. In 2011 there is a significant increase in articles, in which the Latin American production becomes predominant for the theme. Due to the context of the discipline, this growth also occurs for other themes of political science, as demonstrated by Basabe-Serrano and Huertas Hernández (2018), such as parties, democracy, public policies, among others.

Graph 1. Number of scientific articles on Latin American presidentialism by year of publication and institutional origin of authors in the database (1978-2018). Wos and SCI database, search performed on March 23, 2019.



Source: Sangalli (2020)

The proportion of each type of production in the dataset is 51% from Latin American production and 43% from other regions. Only 6% is hybrid production (co-authored between a Latin American author and an author from another region), concentrated in the last decade. When we check the platforms separately, most of the articles are indexed in WoS (66%) and the majority of productions in Latin America are in the SCI platform (29%), this is explained by the very reach of the platforms (table 2).

Table 2. Proportion of types of production separated by bibliographic collection and the distribution of articles in number and percentage. Articles on Latin American presidentialism (1978-2019) in the WoS and SCI databases. Search carried out on March 23, 2019.

	production Latin-American		hybrid production		production from other regions		TOTAL	
	No	%	No	%	No	%	No	%
Web of Science	53	22	11	5	97	40	161	66
SciELO Citation Index	72	29	3	1	8	3	83	34
TOTAL	125	51	14	6	105	43	244	100

Source: Sangalli (2020)

From the division between type of production (graph 1 and table 2), we show which are the research centers on the subject in question (tables 3 and 4). The United States and Brazil are the countries that most produced articles, making up 25% and 22% of the articles in the database, respectively. Mexico's production (9%) is in third place, followed by Argentina, Chile and the United Kingdom (6% each country). Although Latin American productions are the majority, there is a predominance of American studies distributed throughout the analyzed period, although there is a predominance in the first decades (1980-1990).

In addition to being the second research center on the subject, Brazil is also the most studied country, with 24% of the articles contributing about the country. This is due to the study agenda on coalition presidentialism, inaugurated by Abranches in 1988 (SANGALLI,

2019). Production on Mexico is in second place (8%) contributing to the structure of Mexican presidentialism and the country's constitutional reforms (HERNANDEZ RODRIGUEZ, 1978, 1994; HERNANDEZ CHAVES, 1994; PANIZZA, PHILIP, 2005; NEGRETTO, 2006; MAYER -SERRA, 2017). In third place is the production on Argentina (7%) focused on the discussion about democracy, populism, authoritarianism and their effects on the presidential system of government (PHILIP, 1998; LARKINS, 1998; QUINCHE RAMÍREZ, 2013; TEDESCO, DIAMINT, 2014; TARTAKOFF, 2019), there are also discussions about political and economic crisis in the Argentine case (MUSTAPIC, 2005; MALAMUD, 2015; PEÑA, DAVIES, 2017). While studies that analyze Chilean presidentialism focus on the debate on democracy (RABKIN, 1996; SAN FRANCISCO, 2003; CAMERON, 2010) and on political coalitions RENIU, ALBALA, 2012; HUNEEUS 2012; SIAVELIS, 2016; BORGES, TURGEON, 2019; CHAISTY, POWER, 2019).

The presence of productions from thirteen countries from other regions of the world on a specific theme such as the one analyzed demonstrates that this is an internationalized study agenda. As certified in Figure 4, the co-authorship with the highest production intensity is between the United States and Brazil, with a link strength equal to 6. Then there is the link between Brazil and England (3) and Argentina and Spain (2). It is particularly interesting to visualize Brazil's distance from other Latin American countries on the two-dimensional map, which demonstrates greater collaboration with countries in the global north [12]. There is also a group of Latin American countries that are closer to each other on the map – they form the yellow cluster –, indicating scientific ties in the study of the theme between Chile, Colombia, Venezuela and Ecuador.

Table 3 and 4. Countries that produced the most and countries that were most studied in the Latin American presidential system. Proportion and number of articles (WoS and SCI, 1978-2019). Search carried out on March 23, 2019.

country of production	articles	
	No	%
U.S	69	25
Brazil	61	22
Mexico	24	9
Argentina	17	6
Chile	17	6
United Kingdom	17	6
Colombia	13	5
Spain	8	3
Canada	7	3
Uruguay	6	two
Venezuela	4	1
Portugal	5	two
Norway	4	1
Germany	4	1
Ecuador	3	1
France	3	1
Sweden	two	1
Turkey	two	1
Australia	two	1
South Korea	two	1
Finland	1	0
Ireland	1	0
China	1	0
Total	273*	100

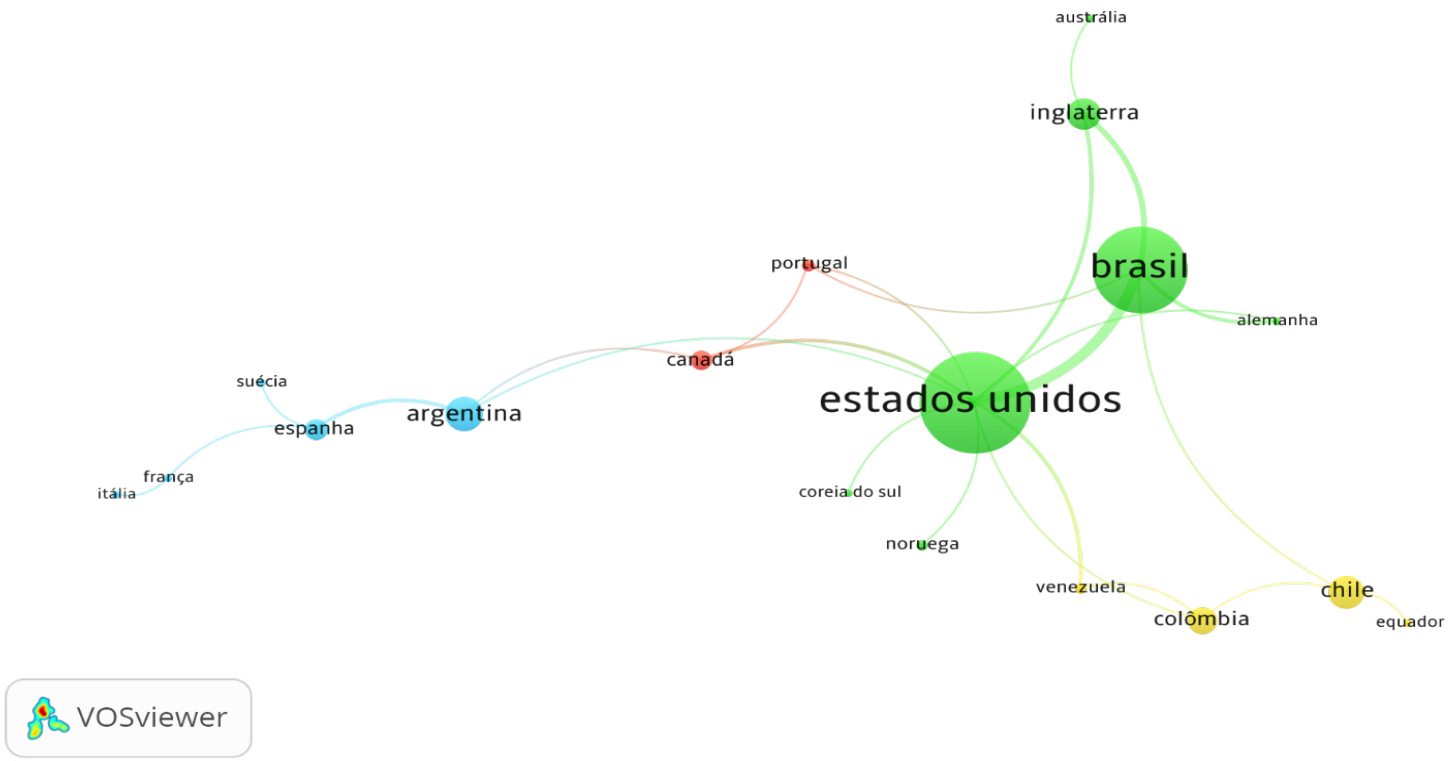
Country studied	articles	
	No	%
Brazil	73	24
Mexico	24	8
Argentina	22	7
Chile	18	6
Venezuela	14	5
Ecuador	13	4
Uruguay	11	4
Peru	9	3
Colombia	8	3
Paraguay	6	two
Rep. Dominican	4	1
Costa Rica	two	1
Nicaragua	1	0
Bolivia	1	0
Grouped region*	96	32
Total	302**	100

Source: Sangalli (2020)

*this total represents the number of authors, as the measure is their institutional origin, thus co-authored articles counted for both countries.

12 The country has only one liaison force with Chile.

Figure 2. Country co-authored network map: 18 items* (minimum of 1 document and 0 citations. **Counting method:** complete. **Normalization method:** LinLog/modularity (attraction 1, repulsion -1); **Visualization method:** network visualization, weight by number of documents and coloring by cluster. The greater the number of documents published by the country, the greater the label and node of the item. The closer the items in two-dimensional space and the thicker the connecting edge, the stronger the co-authorship relationship between them. (Databases: WoS and SCI, search carried out on March 23, 2019. Database on Latin American presidentialism).



Source: Sangalli (2020)

* to obtain this map it was necessary to build a thesaurus.

With regard to the profile of authors and authors,

> Bibliometric research has shown that much of the knowledge in Political Science is produced by groups of people with similar characteristics, which in itself makes other actors and, consequently, their discourses and visions invisible. There is a deficit of representation of gender, race and regional origin, just to mention some of the most recurrent cleavages in bibliometric studies. The expansion of Political Science across the globe was not accompanied by a democratization of knowledge dissemination, impacting not only on the invisibility of themes and epistemological approaches of historically minority groups in the discipline; as well as careers and their incentive structure (LENINE, MÖRSCHBÄCHER, 2020).

Our data on the subject also reveal disparities (table 5 and graph 2). We divided the literature into three groups to measure the intensity of production by gender and co-authorship: i) male production: without the presence of women in any authorship position; ii) mixed production: with the presence of women in some instance of authorship (1st, 2nd or 3rd) and men in the other positions; and, iii) production of women: those with the presence of women in all instances of authorship (table 5).

Table 5. Proportion of productions by genre in bibliographic collections and distribution of articles on Latin American presidentialism in number and percentage. Search carried out on March 23, 2019.

	production of men		mixed production		production of women		TOTAL	
	No	**%**	**No**	**%**	**No**	**%**	**No**	**%**
Web of Science	113	46	21	8	27	11	161	66
SciELO Citation Index	62	26	9	4	12	5	83	34
TOTAL	175	72	30	12	39	16	244	100

Source: Sangalli (2020)

As argued by Freidenberg (2019), political science still faces resistance to gender equality. The production on the theme also reveals this gender disparity (Table 5). In the dataset, 72% of the articles are written by men, while the production of women represents 16% of the articles, with the percentage of collaboration between genders being even lower, 12% of mixed production. Even when we look at each database, both have a higher proportion of articles written by men, not showing a bias between the platforms.

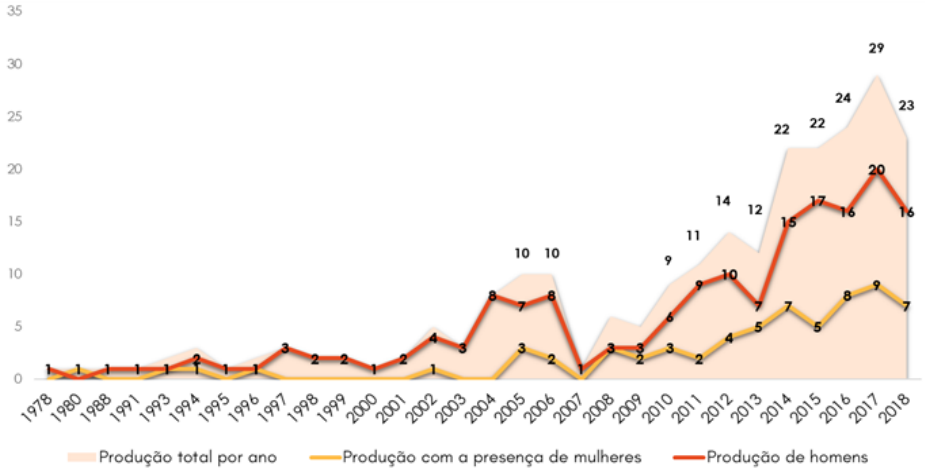
Among the three pillars of the social sciences (sociology, anthropology and political science), political science in Latin America, and specifically in Brazil, is still the most unequal in terms of the distribution of teaching positions, both in terms of gender and race (CANDIDO , FERES JÚNIOR, CAMPOS, 2019).

For Rocha-Carpiuc (2016, apud Rocha-Carpiuc, Madeira, 2019, p.552),

> women are underrepresented in the most important academic posts (universities and national research and evaluation systems). The data indicate that throughout their academic career, women have greater difficulties in reaching such positions. As for academic production, women are more present as presenters at conferences than as authors of articles in the most prestigious academic journals.

According to the author, there are still greater obstacles for women to publish scientific articles. Indeed, as we verified in our case study, the production of women is lower. Although, in Graph 2, it is noticeable that in the last decade there has been an increase in the presence of women in the authorship of articles in relation to previous years.

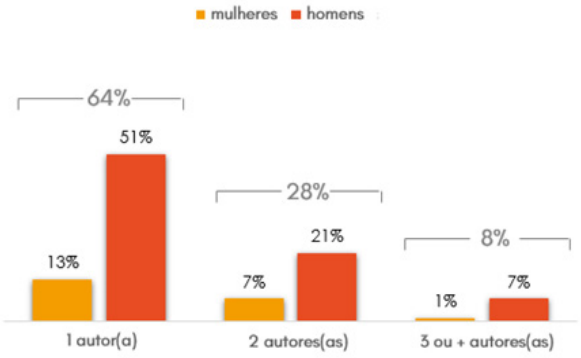
Graph 2. Presence of gender in the annual production on Latin American presidentialism (WoS and SCI, 1978-2018, search carried out on March 23, 2019).



Source: Sangalli (2020)

These studies that investigate gender inequality in science have shown that there are gaps in evaluation policies that disregard the different socializations of gender [13]. In maternity contexts, for example, the scientific production of research mothers is affected by the double or triple working day, making it difficult to produce and publish articles. It is essential that bibliometric studies take into account that productivity is not always synonymous with quality, so even if production is low, an impact analysis is necessary, which combines multiple factors - not just citation - and considers the social context that the researcher is inserted. Thus, a more democratic evaluation of researchers would be possible, promoting financial or symbolic incentives for different conditions of professional performance.

Graph 3. Instances of authorship by gender in articles on Latin American presidentialism (WoS and SCI, 1978-2019, search carried out on March 23, 2019).



Source: Sangalli (2020).

13 ANDRADE, Rodrigo de Oliveira. Motherhood in the curriculum: researchers' project estimates the impact of activities as a mother on the scientific career. Pesquisa Fapesp, São Paulo, ed. 269, 2018. Available at: https://revistapesquisa.fapesp.br/2018/07/19/maternity-no-curriculo/ . Accessed on: December 05, 2020.

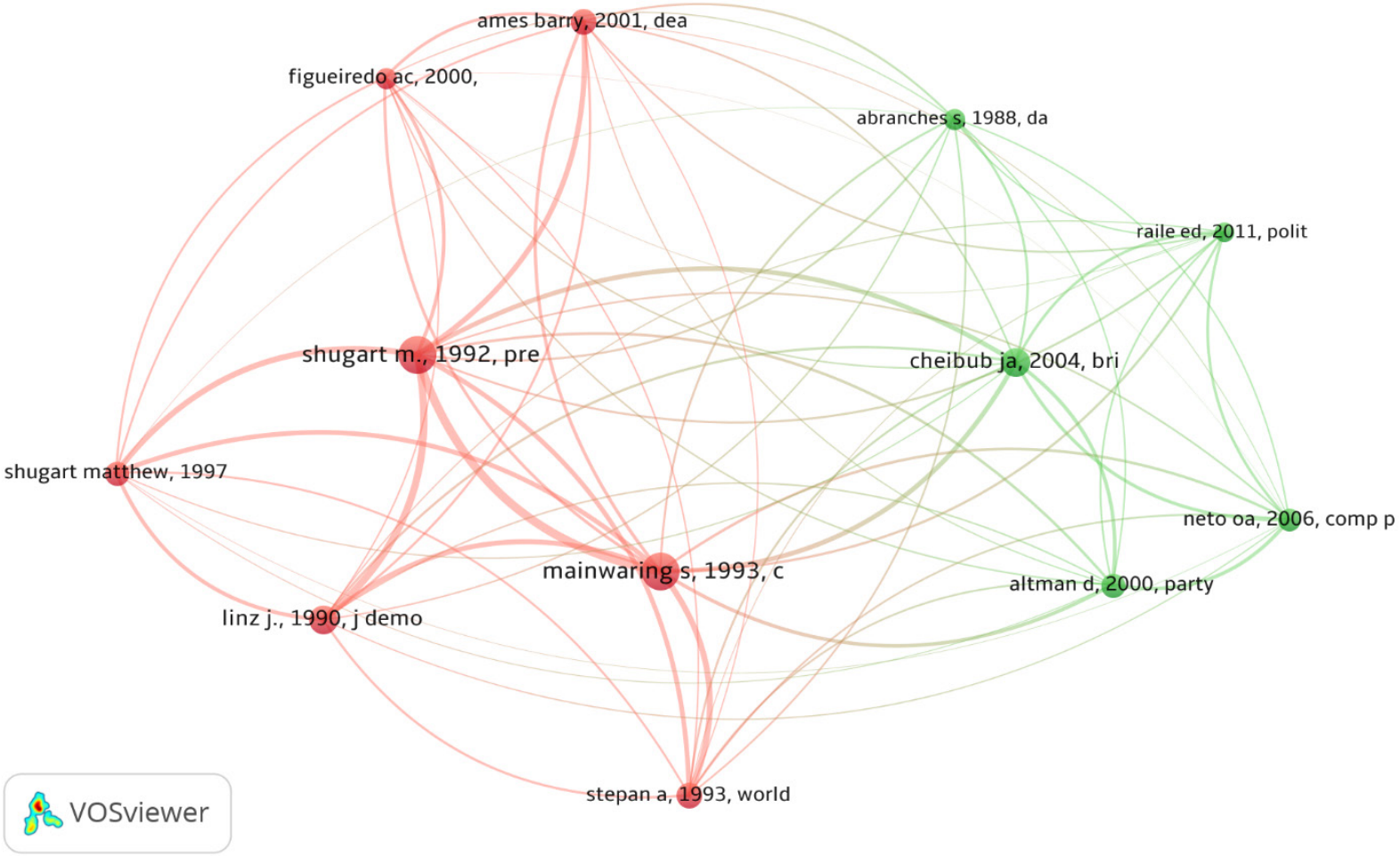
Most of the production is single authored (64% of the articles), 28% of them are co-authored between two authors and 8% written in collaboration between 3 or more authors. Even though numerically smaller, the production of women is concentrated in the same instance of authorship as the production of men, not appearing to be uneven or indicating marginalization in this sense.

4.2 Co-citation: knowing the fundamental documents of the subject

Co-citation is more advantageous than just citing the article alone, as it presents a chain of articles that are cited together, showing discussion affinities beyond their relevance to the scientific community. In the co-citation map, the texts that appear are not necessarily those that are in our database, but the common references between the articles in our database, and articles of discussions parallel to the central theme of Latin American presidentialism may appear.

We identified two intellectual traditions that shape the structure of this area of study (figure 4). In the red cluster we find classic texts on government systems, which discuss the advantages and disadvantages of presidentialism and compare it with parliamentarism. Most of these studies were written in the 1990s and early 2000s. Based on the total strength of the link (Table 2), we can consider that the book by Shugart and Carey (1992) and the articles by Mainwaring (1993) and de Linz (1990), belonging to the red cluster, are the most mentioned references in studies on the subject, with a connection strength >100. While in the green cluster, we find specific studies on coalition presidentialism, with studies by Cheibub, Pzeworski and Saiegh (2004), Altman (2000) and Abranches (1988) being the most influential in this cluster, with a connection strength >60.

Figure 3. Document co-citation network map with the 12 most cited articles on Latin American presidentialism together with the database and their relationships. Documents are linked according to the number of reciprocal citations . **Counting method:** complete. **Normalization method:** LinLog/modularity (attraction 2, repulsion 1); **Visualization method: n**etwork visualization, weight by total link strength. The closer the items are in two-dimensional space, the stronger their relationships. Staining occurs by conformation of clusters. The greater the number of links, the greater the item's label and node (Search carried out in WoS and SCI on March 23, 2019).



Source: Sangalli (2020)

Table 4. Ranking of the 12 fundamental documents on Latin American presidentialism according to the total number of joint appearances (links) in the references of the articles in the database (Search carried out in WoS and SCI on March 23, 2019).

ranking	Reference	Species	quotes	links	full strength of the link	cluster
1	Shugart, MS, & Carey, JM (1992). *Presidents and assemblies: Constitutional design and electoral dynamics* . Cambridge University Press.	book	65	11	166	**Red**
two	Mainwaring, S. (1993). Presidentialism, multipartism, and democracy: the difficult combination. *Comparative political studies* , 26(2), 198-228.	article	54	11	156	**Red**
3	Linz, JJ (1990). The profiles of presidentialism. *Journal of Democracy* , 1(1), 51-69.	article	41	11	102	**Red**
4	Shugart, MS, & Mainwaring, S. (1997). Presidentialism and democracy in Latin America: Rethinking the terms of the debate. In *Presidentialism and Democracy in Latin America* (pp. 12-54). Cambridge University Press.	book chapter	34	10	81	**Red**
5	Cheibub, JA, Przeworski, A., & Saiegh, SM (2004). Government coalitions and legislative success under presidentialism and parliamentarism. *British Journal of Political Science* , 565-587.	article	33	11	102	**green**
6	Stepan, A., & Skach, C. (1993). Constitutional frameworks and democratic consolidation: Parliamentarianism versus presidentialism . *World Politics* , 1-22.	article	32	11	88	**Red**
7	Ames, B. (2001). *The deadlock of democracy in Brazil.* University of Michigan Press.	book	26	11	84	**Red**
8	Altman, D. (2000). The politics of coalition formation and survival in multiparty presidential democracies: The case of Uruguay, 1989-1999. *Party Politics* , 6(3), 259-283.	article	25	11	79	**green**
9	Abranches, S. (1988). Coalition presidentialism: the Brazilian institutional dilemma. *Data* , 31(1), 5-38.	article	23	11	61	**green**
10	Figueiredo, AC, & Limongi, F. (2000). Presidential power, legislative organization, and party behavior in Brazil. *Comparative Politics* , 151-170.	article	21	11	66	**Red**
11	Raile, ED, Pereira, C., & Power, TJ (2011). The executive toolbox: Building legislative support in a multiparty presidential regime. *Political Research Quarterly* , 64(2), 323-334.	article	21	10	59	**green**
12	Neto, OA (2006). The presidential calculus: Executive policy making and cabinet formation in the Americas. *Comparative Political Studies* , 39(4), 415-440.	article	20	11	68	green

Source: Sangalli (2020).

4.3 Bibliographic and co-word coupling: identifying research fronts and emerging concepts of the theme

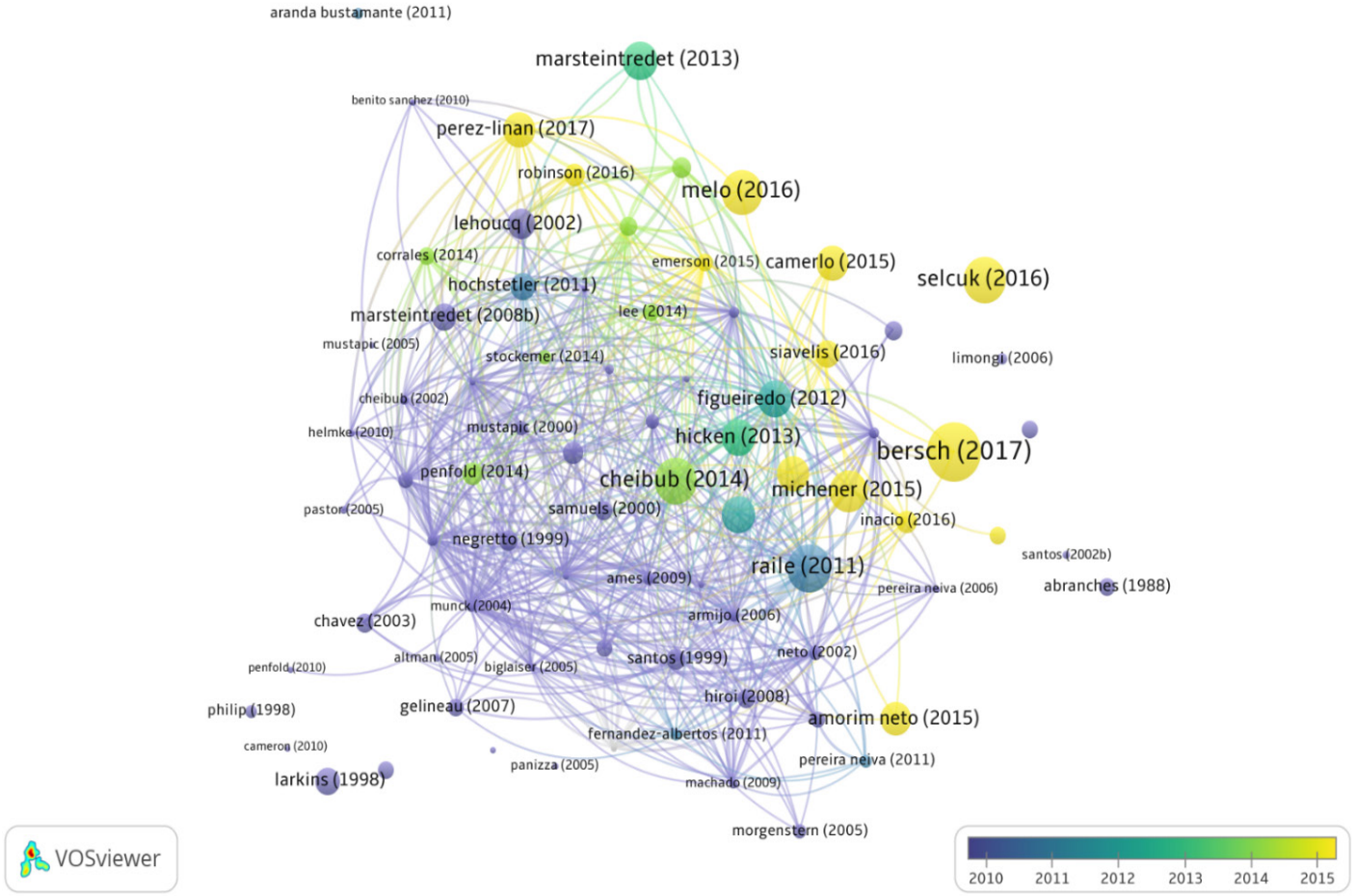
Bibliographic coupling is the most recommended technique to identify research fronts. Its applicability is for short publication periods (up to 5 years), so it is not advisable to base your interpretation on the number of citations. However, an alternative is to use normalized citation [14]. Normalization corrects the fact that older documents had more time to receive citations than newer documents (VAN ECK, WALTMAN, 2017), assigning a weighted citation weight given the article's publication time.

We created an overlay map (figure 5) where the focus of analysis is the articles with yellow color. The more references the articles share, the stronger the connection between them. Thus, the closer they are in the two-dimensional space, the greater their theoretical and/or methodological affinities, indicating an implicit relationship between them (GRÁCIO, 2016). The most recent studies, that is, the research fronts, are shown in Table 3 in a ranking format by the weight of normalized citation. It was possible to identify two thematic clusters, cluster 1 is about political crisis and cluster 2 is about public administration and bureaucracy.

Complementarily, the co-word map (figure 6) shows the main concepts of these research fronts. Connecting words that appear together in the texts, we identified the emerging concepts (yellow color) which are: a) Impeachment: it is connected with words such as crisis, democracy, participation and stability, pointing out the descendant of the impeachment debate. The emergence of this concept is related to the Latin American situation, since recent impeachment cases have occurred in the region, in Ecuador, Paraguay, Guatemala and Brazil; b) Ministry, cabinet and agency: these terms are connected to each other and to other terms such as coalition government and governability. It is presumable, therefore, that they are central in studies of bureaucratic capacities and in studies that analyze the relationship between the Legislative and Executive powers.

14 "The normalized number of citations of a document is equal to the number of citations of the document divided by the average number of citations of all documents published in the same year and included in the data that is provided to VOSviewer" (VAN ECK, WALTMAN, 2017, p. 32, our translation).

Figure 4. Bibliographic coupling overlay map with 79 items (minimum of 3 citations and maximum of 50 citations). **Counting method**: complete. **Normalization method**: strength of association (attraction 2, repulsion 0); **Visualization method**: overlay visualization, weight per normalized citation and coloring per year of publication. The higher the number of normalized citations, the greater the label and node of the item, the more recent the publication, the more yellowish the color of the item (Database on Latin American presidentialism. Search carried out in WoS and SCI on 23 March 2019).



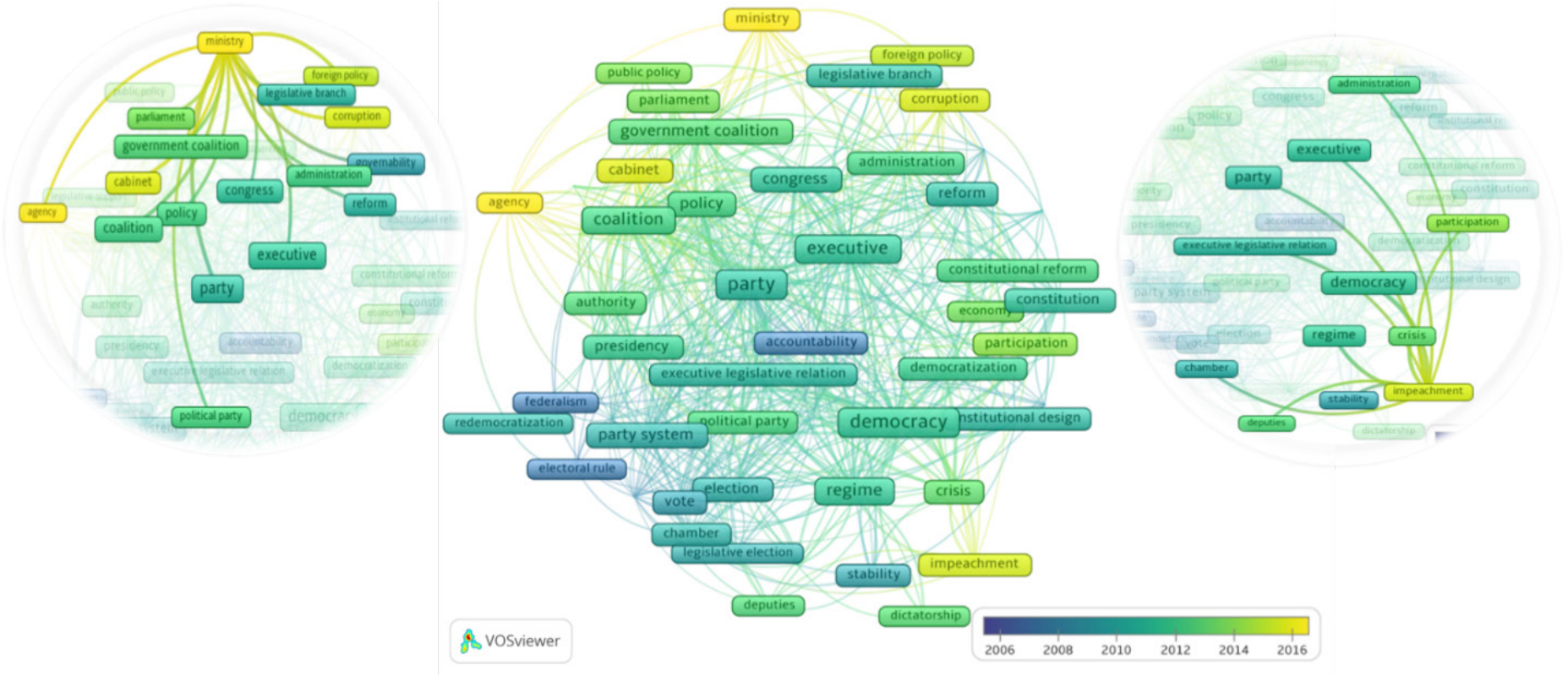
Source: Sangalli (2020).

Table 5. Ranking of search fronts: the 10 most recent articles on Latin American presidentialism, with the highest weight of normalized citation (Search carried out in WoS and SCI on March 23, 2019).

ranking	authorship	Article	Score: year of publication	Cientific magazine	quotes	Weights: normalized citation	links	full strength of the link	cluster
1	Bersch K. Square S. Taylor MM	State Capacity, Bureaucratic Politicization, and Corruption in the Brazilian State	2017	governance	57	10.80	34	55	two
two	Selçuk, O.	Strong presidents and weak institutions: populism in Turkey, Venezuela and Ecuador	2016	Southeast European and Black Sea Studies	3	6.70	5	5	1
3	Melo, MA	Latin America's New Turbulence: Crisis and Integrity in Brazil	2016	Journal of Democracy	5	6.14	49	75	1
4	Michener, G.	How cabinet size and legislative control shape the strength of laws transparency	2015	governance	18	5.33	50	87	two
5	Pérez-Liñán, A., Polga-Hecimovich, J.	Explaining military coups and impeachments in Latin America	2017	democratization	10	4.05	42	116	1
6	Martínez-Gallardo, C., Schleiter, P.	Choosing Whom to Trust: Agency Risks and Cabinet Partisanship in Presidential Democracies	2015	Comparative Political Studies	48	3.76	40	72	two
7	Camerlo, M., Pérez-Liñán	Minister Turnover, Critical Events, and the Electoral Calendar in Presidential Democracies	2015	The Journal of Politics	14	3.76	38	60	two
8	Amorim Neto, O., Malamud, A.	What Determines Foreign Policy in Latin America? Systemic versus Domestic Factors in Argentina, Brazil, and Mexico, 1946–2008	2015	Latin American Politics and Society	11	3.45	28	42	two
9	Sivelis, PM	Cabinets and Informal Advisory Networks in Multiparty Presidential Systems	2016	Presidential Studies Quarterly	9	2.23	51	94	two
10	Inácio, M., Llanos, M.	The Institutional Presidency in Latin America: A Comparative Analysis	2016	Presidential Studies Quarterly	20	1.67	52	110	two

Source: Sangalli (2020)

Figure 5. Co-occurrence network map with the 56 most frequent terms* (minimum of 5 appearances) in the titles and abstracts of the 244 articles on Latin American presidentialism and their reciprocal relationships. The terms are connected according to the number of joint appearances in text. **Counting method:** binary; **Normalization method:** LinLog/modularity (attraction 3, repulsion 0); **Visualization method:** overlay visualization: weight per occurrence. The greater the frequency of the term, the greater its weight, therefore, the greater its label. The thicker the edge, the stronger the connection between the terms. The proximity between terms in two-dimensional space represents greater strength in their relationships. The color is associated with the year of publication of the article that contains the term, the more recent, the more yellow the color of the item. Next to the map are two zoom images on emerging concepts with their connections. (Search conducted on WoS and SCI on March 23, 2019)



Source: Sangalli (2020).

*To obtain this map it was necessary to build a thesaurus.

5. Conclusion

This article differentiated the concepts and applications of bibliometrics and scientometrics methods and proposed to demonstrate how to apply them to the study of a topic of interest in political science. The main advantages of using information metric methods are their reliability and breadth. When well-executed data collection saves the researcher's time in organizing the data and facilitates its transposition to bibliometric network software (HAYASHI, 2013), in addition to ensuring replication through the transparency of the selection steps. Bibliometric contact with an area of knowledge has a wide range of uses. At first, it allows the analysis of multiple dimensions of science, such as: who writes, what writes, where they disclose, how the community communicates, allowing the understanding of the profile and behavior of the scientific community.

In the background, we consider it valuable that a student-researcher at the beginning of his career knows how to use such a methodology, because in addition to measuring and measuring scientific activity, learning scientometrics can serve as a means of career planning. Bibliometric indicators such as co-authorship can help to identify which countries and institutions are most dedicated to the study of a topic, making it possible, for example, for this student to organize themselves to apply for postgraduate scholarships in institutions that can contribute more significantly to their training and future professional performance. Bibliographic coupling and co-word are also advantageous in this sense, as visualizing which are the research fronts and emerging concepts makes it easier to know the current debates and to be informed about the gaps in the studies already carried out, being an essential point to elaborate innovative projects. It is also possible to carry out a bibliometric analysis to find the journals that publish more frequently on a given topic, simplifying the process of choosing the best journal to submit an article. Furthermore, the bibliographic databases have the authors' electronic addresses, enabling contact with other scientists with theoretical-methodological affinities for future scientific partnerships.

Our methodological application contributed to exemplify the use of the VOSviewer software in creating networks with bibliographic data. The applied indicators proved to be advantageous for the knowledge of the subject. We identified that the profile of authors who write about Latin American presidentialism is mostly male and from the United States. However, in a grouped manner, it is possible to say that Latin American production has a greater number of articles in our database than that of other regions. The presence of co-authorship with several countries outside the region demonstrates that the theme is of international interest, as well as regional. It was evident that in the analyzed bases, the strongest relationship of co-authorship on the topic is between American and Brazilian researchers.

The main limitation of this study is that it measures only productions indexed to selected bibliographic collections, not comprehensively covering all productions on the subject. Among the limitations of the bibliographic search, we emphasize a "hidden bias" in the indexing, as the main scientific databases have a high percentage of journals from institutions in the northern hemisphere and ends up reproducing the existing inequality between the scientific production of economically more developed and less developed countries. (KAUCHAKJE, ROSA, 2020). Thus, it is necessary that future studies also explore other bibliographic platforms or manually add articles not covered by this research. Another limitation is that the document selection filter was not restricted to studies published in journals in the field of political science,

as we consider that the theme is of multidisciplinary interest, which can interfere with the interpretation of the transformations of the theme for the discipline of political science in isolation.

As we saw in this article, bibliometrics is concerned with evaluating productivity and scientometrics with field relations. On the one hand, this type of study is rich in detail and useful at an exploratory level. On the other hand, there was resistance to bibliometric studies due to their exclusively quantitative nature, which does not allow deeper analysis of the contents of the documents (HAYASHI, 2013). In light of this, a new approach to the concept, neo-bibliometrics, emerged in the theoretical-methodological debate. This approach presents a qualitative path to be followed together with the quantitative procedures and this combination provides both an approximation and a theoretical and methodological design of the studies. We emphasize, however, that even though the bibliometric indicators are of a quantitative nature, a qualitative approach between the researchers and the database was essential for a better interpretation of the findings. In this way, we conclude that the new guise of the term bibliometrics emerged more as a transparent improvement of the information knowledge process, than as a new concept and application.

We also recognize that to go beyond the descriptive data we have presented, it would be important to contextualize the results – both within the context of the discipline or field under analysis, as well as the development of science and funding policies – for better interpretation. That said, we reinforce the importance of the methodology for the analysis of other central themes and theoretical aspects in the area of political science such as State and Government, Political Institutions, Political Culture, Political Behavior and Public Policies for a comparative perspective. We also advise you to deepen this research in the search to understand the relationship between state capabilities and governability in presidential regimes to produce a systematic analysis of the literature. Or even, to continue the studies on gender in Latin American political science.

BIBLIOGRAPHICAL REFERENCES

ABRANCHES, Sérgio H. H. (1988), "Presidencialismo de coalizão: o dilema institucional brasileiro", *Dados–Revista de Ciências Sociais*, vol. 31, no. 1: 1-39.

ALTMAN, David. (2000), "The politics of coalition formation and survival in multiparty presidential democracies: The case of Uruguay, 1989-1999", *Party politics*, vol. 6 no.3: 259-283.

BARRIENTOS DEL MONTE, Fernando. (2013), "La ciencia política en América Latina: una breve introducción histórica", *Convergencia*, vol. 20, np. 61: 105-133.

BASABE-SERRANO, Santiago; HUERTAS-HERNANDEZ, Sergio. (2018), "El estado de la investigación en Ciencia Política sobre América Latina", *Revista Española de Ciencia Política-RECP*, no. 47: 153-170.

BENENTE, Mauro. (2017), "Presidencialismo y poder económico en la reforma constitucional Argentina de 1994", *Revista de Derecho*, no. 47: 73-104.

BERSCH, Katherine; PRAÇA, Sérgio; TAYLOR, Matthew M. (2017), "State capacity, bureaucratic politicization, and corruption in the Brazilian state", *Governance*, vol. 30, no. 1: 105-124.

BORGES, André; TURGEON, Mathieu. (2019), "Presidential coattails in coalitional presidentialism", *Party Politics*, vol. 25, no. 2: 192-202.

CAMERON, Maxwell A. (2010), "The State of Democracy in the Andes: Introduction to a thematic issue of Revista de Ciencia Política", *Revista de ciencia política (Santiago)*, vol. 30, no. 1: 5-20.

CANDIDO, Marcia Rangel; FERES JÚNIOR, João; CAMPOS, Luiz Augusto. (2019), "Desigualdades na elite da Ciência Política brasileira", *Civitas-Revista de Ciências Sociais*, vol. 19, no.3: 564-582.

CHAISTY, Paul; POWER, Timothy J. (2019), "Flying solo: Explaining single-party cabinets under minority presidentialism", *European Journal of Political Research*, vol. 58, no. 1: 163-183.

CHÁVEZ, Alicia Hernández. (1994), "Mexican presidentialism: A historical and institutional overview", *Estudios Mexicanos*, vol. 10, no.1: 217-225.

CHEIBUB, José Antonio; PRZEWORSKI, Adam; SAIEGH, Sebastian M. (2004), "Government coalitions and legislative success under presidentialism and parliamentarism", *British Journal of Political Science*, vol. 34, no. 4: 565-587.

DAZA, Javier Duque. (2015), "¿Presidentes legislando? Las facultades legislativas en el presidencialismo colombiano 1991-2014", *Revista Estudios Socio-Jurídicos*, vol. 17, no. 1: 79-123.

FREIDENBERG, Flavia. (2019), "La ceguera de género en la ciencia política latinoamericana", *La gaceta de ciencia política*, no. 1: 26-44. Disponível em: https://gacetadecpol.files.wordpress.com/2019/03/gaceta-2019-final-1.pdf . Acesso em: 22 ago. 2020.

GÄNZEL, Wolfang. (2003), *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. Course Handouts.

GARCÍA, Jean-René. (2014), "¿Presidencialismo o Poder Ejecutivo ambivalente?: De la pertinencia de la noción jurídica para el estudio de los sistemas políticos de América Latina", *Cuestiones constitucionales*, no. 31: 57-79.

GRANOVSKI, Yuri V. (2001), "Is it possible to measure science? V. V. Nalimov's research in scientometrics", *Scientometrics*, vol. 52, no. 2: 127-150

HAYASHI, Carlos. R. Massao. (2013), "Apontamentos sobre a coleta de dados em estudos bibliométricos e cientométricos", *Filosofia e Educação*, vol. 5, no. 2: 89–102. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rfe/article/view/8635396 . Acesso em: 15 nov. 2020.

HAYASHI, Maria Cristina. (2012), "Sociologia da Ciência, Bibliometria e Cientometria: Contribuições para a análise da produção científica", Anais Eletrônico – IV EPISTED – Seminário de Epistemologia e Teorias da Educação, Faculdade de Educação/Unicamp. Disponível em: https://www.marilia.unesp.br/Home/Graduacao/PETBiblioteconomia/soc-da-ciencia-pet.pdf . Acesso em: 20 jun. 2020

HERNÁNDEZ, Rogelio. (1978), "¿Presidencialismo u oligarquía? Algunas reflexiones sobre la estructura política mexicana", *Problemas del Desarrollo*, vol. 9, no. 35:53-85.

HUNEEUS, Carlos. (2012), "Variedades de governos de coalizão no presidencialismo. Chile, 1990-2010", *Dados-Revista de Ciências Sociais*, vol. 55, no. 4: 877-910.

KAUCHAKJE, Samira; ROSA, Maria Arlete. (2020), *Revisão sistemática da literatura: Como Fazer? vol. 1*, Curitiba, Editora UTP.

LARKINS, Christopher. (1998) "The judiciary and delegative democracy in Argentina", *Comparative Politics*, vol. 30, no. 4: 423-442.

LEITE, Fernando; CODATO, Adriano. (2013), "Autonomização e institucionalização da Ciência Política brasileira: o papel do sistema Qualis-Capes", *Agenda Política*, vol. 1, no.1: 1-21.

LENINE, Enzo; MÖRSCHBÄCHER, Melina. (2020), "Pesquisa bibliométrica e hierarquias do conhecimento em Ciência Política", *Revista Brasileira de Ciência Política*, no. 31: 123-160.

LIMA, Iana Alves de. (2018) "O impacto do arranjo institucional brasileiro no controle político sobre a burocracia", *Cadernos EBAPE. BR*, vol. 16, no. 4: 656-666.

LINZ, Juan J. (1990), "The perils of presidentialism", *Journal of democracy*, vol. 1, no.1: 51-69.

MAINWARING, Scott. (1993), "Presidentialism, multipartism, and democracy: The difficult combination", *Comparative political studies*, vol. 26, no. 2: 198-228.

MALAMUD, Andrés. (2015), "Social revolution or political takeover? The Argentine collapse of 2001 reassessed", *Latin American Perspectives*, vol. 42, no.1: 11-26.

MARCIAS-CHAPULA, Cesar A. (1998), "O papel da informetria e da ciênciometria e sua perspectiva nacional e internacional", *Ciência da Informação*, vol. 27, no. 2: 134-140.

MAYER-SERRA, Carlos Elizondo. (2017), "Reforma de la Constitución: la economía política del Pacto por México", *Revista mexicana de ciencias políticas y sociales*, vol. 62, no. 230: 21-49.

MUGNAINI, Rogério; JANNUZZI, Paulo de Martino; QUONIAM, Luc. (2004), "Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal", *Ciência da informação*, vol. 33, no. 2: 123-131.

MUSTAPIC, A. M. (2005), "Instability without collapse. The resignation of the presidents: Argentina in 2001", *Desarrollo Economico-Revista De Ciencias Sociales*, vol. 45, no. 178: 263-280.

NAKAGUMA, Marcos Yamada. (2015), "Choosing the form of government: Theory and evidence from Brazil", *The Economic Journal*, vol. 125, no. 589: 1986-2023.

NEGRETTO, Gabriel L. (2006), "Minority presidents and democratic performance in Latin America", *Latin American Politics and Society*, vol. 48, no. 3: 63-92.

ORTIZ ORTIZ, Richard. (2018), "Los problemas estructurales de la Constitución ecuatoriana de 2008 y el hiperpresidencialismo autoritario", *Estudios constitucionales*, vol. 16, no. 2: 527-566.

PANIZZA, Francisco; PHILIP, George. (2005), "Second generation reform in Latin America: Reforming the public sector in Uruguay and Mexico", *Journal of Latin American Studies*, vol. 37, no. 4: 667-691.

PEÑA, Alejandro Milcíades; DAVIES, Thomas Richard. (2017), "Responding to the street: government responses to mass protests in democracies", *Mobilization: An International Quarterly*, vol. 22, no. 2: 177-200.

PHILIP, George. (1998), "The New Populism, Presidentialism and Market-Orientated Reform in Spanish South America", *Government and Opposition*, vol. 33, no. 1: 81-97.

QUINCHE RAMÍREZ, Manuel Fernando. (2013), "Presidentialism, conventionality control and democracy in Andean countries", *Co-herencia*, vol. 10, no. 19: 167-206.

RABKIN, Rhoda. (1996), "Redemocratization, electoral engineering, and party strategies in Chile, 1989-1995", *Comparative Political Studies*, vol. 29, no. 3: 335-356.

RENIU, Josep Ma; ALBALA, Adrián. (2012), "Los gobiernos de coalición y su incidencia sobre los presidencialismos latinoamericanos: el caso del Cono Sur", *Estudios políticos (México)*, no. 26: 161-214.

RÍOS GÓMEZ, Claudia; HERRERO SOLANA, Víctor. (2005), "La producción científica latinoamericana y la ciencia mundial: una revisión bibliográfica (1989-2003)", *Revista interamericana de Bibliotecología*, vol. 28, no. 1: 43-61.

ROCHA-CARPIUC, Cecilia; MADEIRA, Rafael Machado. (2019), "Desigualdade de gênero, internacionalização e trajetórias acadêmicas na Ciência Política: Evidências no Brasil e no Uruguai", *Civitas-Revista de Ciências Sociais*, vol. 19, no. 3: 545-563.

SAN FRANCISCO, Alejandro. (2003), "La gran convención del partido liberal democrático en 1893: Un hito en la reorganización del balmacedismo después de la guerra civil chilena de 1891", *Historia (Santiago)*, no. 36: 333-377.

SANCHO, Rosa. (1990), "Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Revisión bibliográfica", *Revista española de documentación científica*, vol. 13, no. 3-4: 842-865.

SANGALLI, Amanda F. (2019), "Presidencialismo latino-americano: uma análise descritiva e cientométrica da temática (1978-2019)", Trabalho de Conclusão de Curso (Graduação em Ciências Sociais) – Pontifícia Universidade Católica do Paraná, Curitiba, p.1-71. Disponível em: https://www.academia.edu/43909793/PRESIDENCIALISMO_LATINO_AMERICANO_UMA_AN%C3%81LISE_DESCRITIVA_E_CIENTOM%C3%89TRICA_DA_TEM%C3%81TICA_1978_2019_

SANGALLI, Amanda F. (2020), "Mapeando las producciones sobre el presidencialismo latinoamericano (1978-2019)", Jornadas de Ciencia Política, área temática: política latinoamericana, Facultad de Ciencias Sociales de la Universidad de Buenos Aires, Argentina. Disponível em: https://www.academia.edu/44239452/Mapeando_las_producciones_sobre_el_presidencialismo_latinoamericano_1978_2019_ . Acesso em: 20 set. 2020.

SANTOS, Raimundo Nonato Macedo dos; KOBASHI, Nair Yumiko. (2009), "Bibliometria, cientometria, infometria: conceitos e aplicações", Grupo de Trabalho 7: Produção e Comunicação da Informação em CT&I, Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, vol. 2, no. 1: 155-172.

SHUGART, Matthew Soberg; CAREY, John M. (1992), *Presidents and assemblies: Constitutional design and electoral dynamics*, Cambridge University Press.

SIAVELIS, Peter M. (2016), "Cabinets and Informal Advisory Networks in Multiparty Presidential Systems", *Presidential Studies Quarterly*, vol. 46, no. 3: 569-591.

SILVA, José Aparecido da; BIANCHI, Maria de Lourdes Pires. (2001), "Cientometria: a métrica da ciência", *Paidéia (Ribeirão Preto)*, vol. 11, no. 21: 5-10.

TAGUE-SUTCLIFFE, Jean. (1992), "An introduction to informetrics", *Information Processing & Management*, vol. 28, no. 1: 1-3.

TARTAKOFF, Laura Y. (2019), "Ecuador, Argentina, and Paraguay: At Last Immunity to Authoritarian Rule?", *Society*, vol. 56, no. 1: 15-22.

TEDESCO, Laura; DIAMINT, Rut. (2014), "Latin American Democracy. What to Do with the Leaders?", *Bulletin of Latin American Research*, vol. 33, no. 1: 31-45.

TIMMERMANN, Andreas. (2012), "El concepto de" gobierno moderado" como hilo conductor en el constitucionalismo temprano de Hispanoamérica", *Estudios de historia moderna y contemporánea de México*, no. 44: 3-48.

VAN RAAN, A. F. J. (1997), "Scientometrics: state-of-the-art", *Scientometrics*, vol. 38, no. 1: 205-2018

VANTI, Nadia Aurora Peres. (2002), "Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento", *Ciência da informação*, vol. 31, no. 2: 152-162.

VANTI, Nadia Aurora Peres. (2005), "Os links e os estudos webométricos", *Ciência da informação*, vol. 34, no. 1: 78-88.

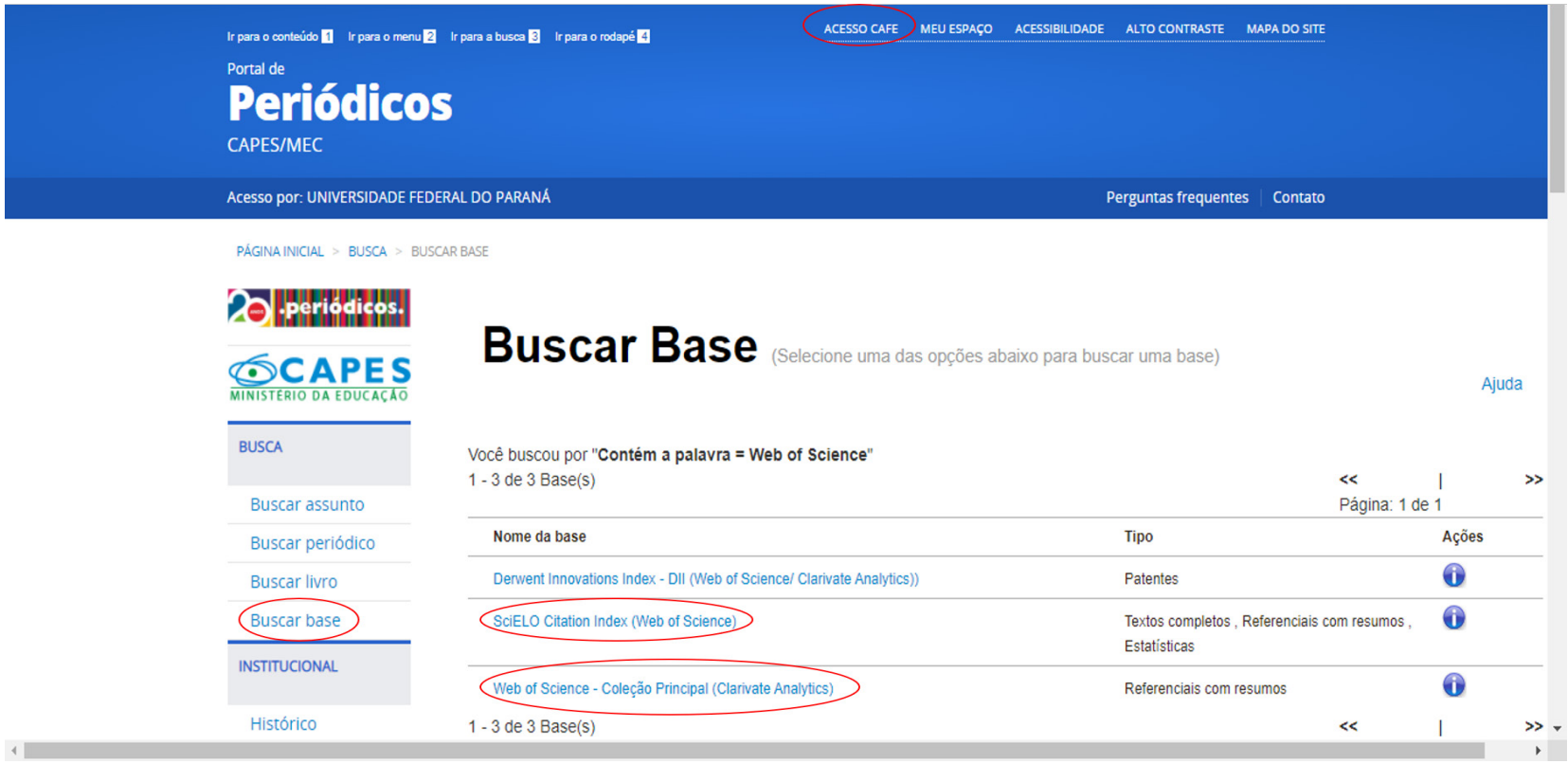
VANTI, Nadia Aurora Peres. (2011), "A cientometria revisitada à luz da expansão da ciência, da tecnologia e da inovação", *Ponto de Acesso*, vol. 5, no. 3: 5-31.

WALTMAN, L.; VAN ECK, N. VOSviewer manual: manual for VOSviewer version 1.6.6, Centre for Science and Technology Studies, Leiden University, The Netherlands, p. 1-48. Disponível em: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.6.pdf . Acesso em: 10 jun. 2019.

ZUPIC, Ivan; ČATER, Tomaž. (2015), "Bibliometric methods in management and organization", *Organizational Research*

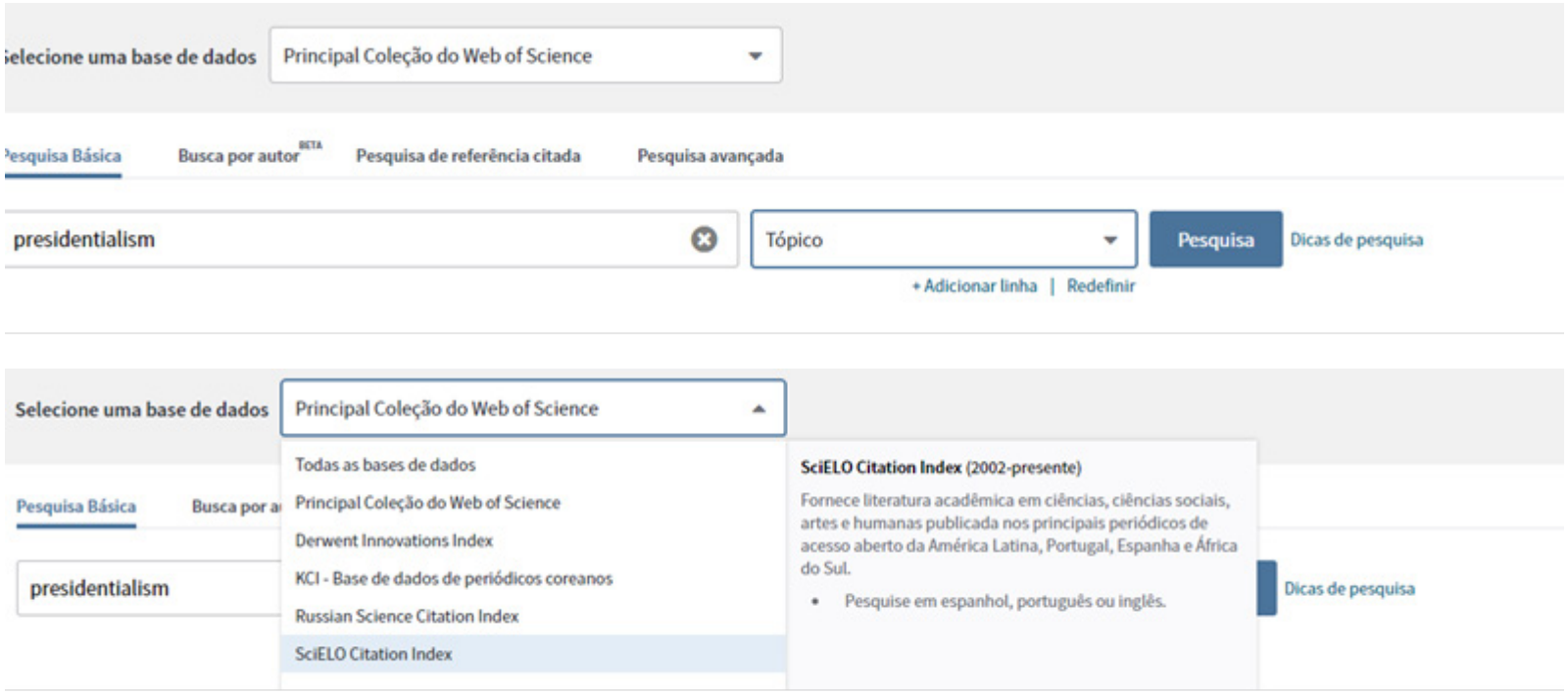
ANNEX 1 - CONSULTATION ON THE TOPIC OF PRESIDENTIALISM IN THE BIBLIOGRAPHIC BASES AND DATA EXTRACTION[1]

Path for consulting and extracting bibliographic data: access the CAPES/MEC Journals Portal and perform remote access via CAFe with institutional login, go to "Search base" and search by title "Web of Science". At this moment, you can find both collections that we use.

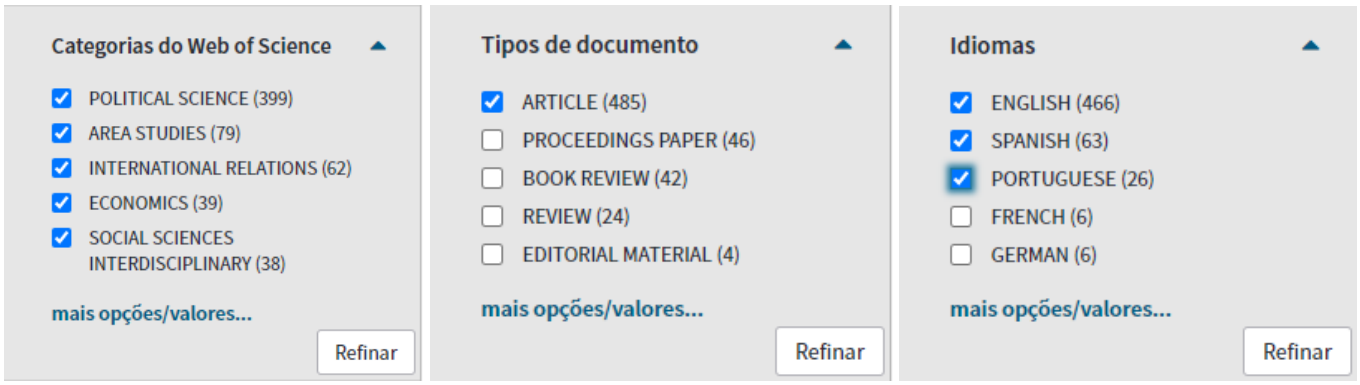


To search the main WoS collection, we wrote the search term "presidentialism" in Basic Search and performed a search of the type "Topic". This was also done for the SCI collection.

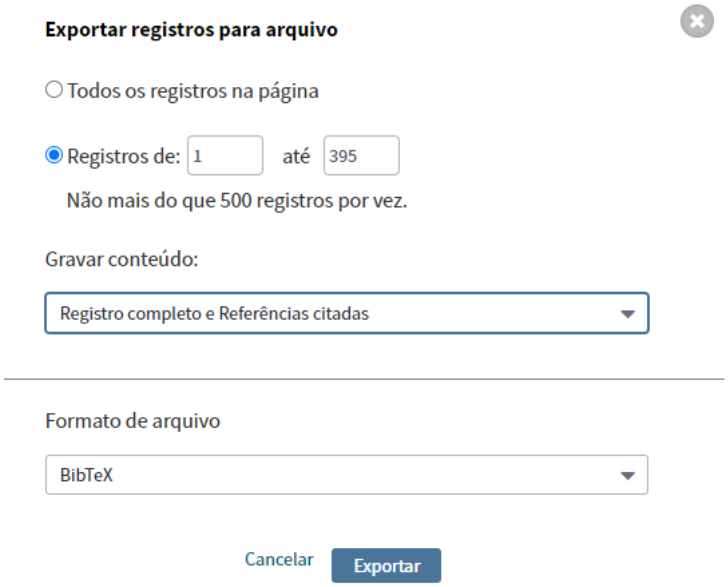
1 The images were captured on March 24, 2021, after the study was carried out, just to exemplify this introductory article, therefore, the numbers contained in the images should be ignored, as they do not correspond to the search carried out in March 2019.



After the search, we filtered the results by "Web of Science (or SciELO) Categories", "Document Type" and "Languages".



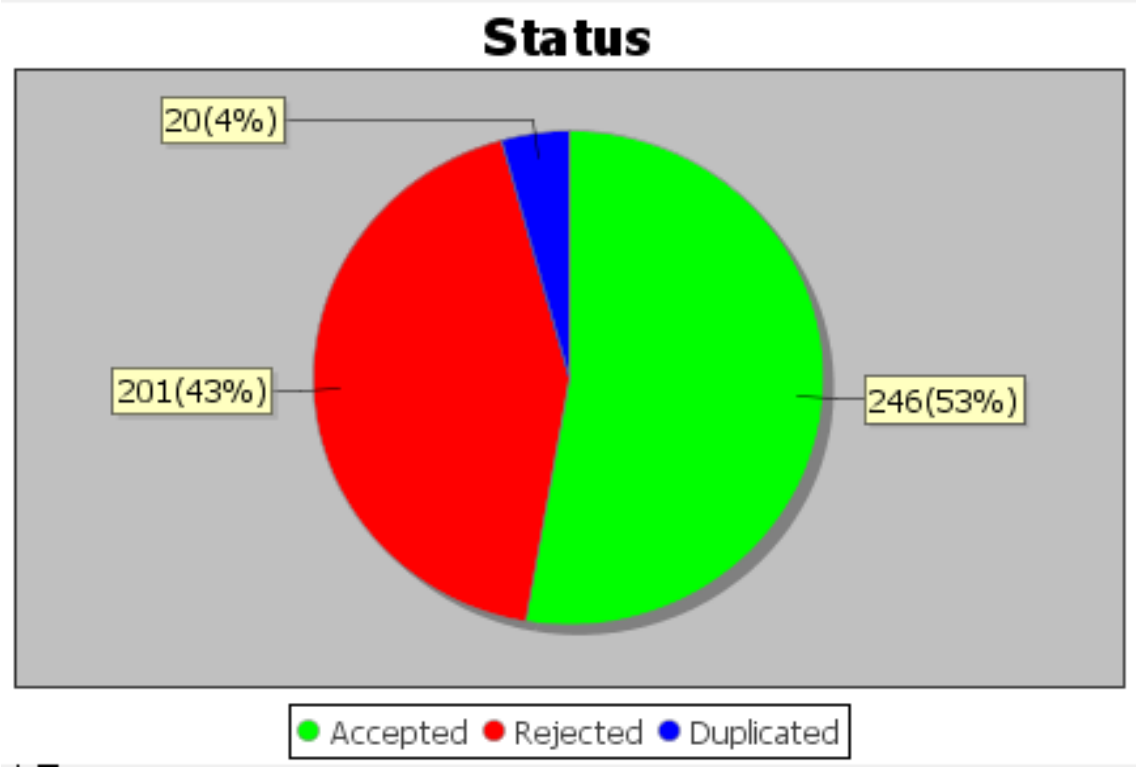
Finally, the "Complete Record and Cited References" of all selected documents was extracted.



ANNEX 2 – START PROTOCOL APPLIED TO THE DETAILED ANALYSIS OF ABSTRACTS FOR THE SELECTION OF ARTICLES ON THE SUBJECT OF LATIN AMERICAN PRESIDENTIALISM

START Protocol

Title:	An introduction to bibliometrics and scientometrics: the case of Latin American presidentialism
Researchers:	Amanda Sangalli; Samira Kauchakje.
Description:	Use StART for detailed reading of abstracts of articles selected in a previous step in the main Web of Science collection and in the SciElo Citation Index collection. The articles collected in both databases focus on the theme of presidentialism, StArt will only be used as a facilitating instrument for reading the abstracts to refine our selection, that is, apply the regional cut. For an article to proceed to the mapping step, it must meet the inclusion criteria.
Objective:	Select articles with an analytical focus on Latin American presidentialism(ies)
Main question:	How to map scientific production on presidentialist countries in Latin America?
Keywords and Synonyms:	presidentialism
Sources Selection Criteria Definition:	Main international scientific base and main regional base (Latin America)
Studies Languages:	Portuguese, Spanish and English.
Source Search Methods:	Manual search in the main Web of Science collection and in the SciELO Citation Index collection of the "topic" type using the term "presidentialism"; and, export of "complete record and cited references" metadata of documents in BibTex and TXT format.
Source list:	Web of Science; SciELO Citation Index.
Study selection criteria: (Inclusion and exclusion criteria)	Inclusion: a) It is a study on presidentialism; b) It is a study that compares presidentialism with other systems of government. c) It is a study of one or more Latin American countries; d) It is a comparative analysis of Latin American presidential countries with other countries in other regions. Exclusion: a) It is not a study on presidentialism. b) It's about presidentialism, but it's not about Latin American countries; c) It is about Latin American countries, but they are not studies about the presidential system of government.
Studies Types Definition:	Articles published in scientific journals/periodicals
Studies Initial Selection:	From the titles, keywords and reading the abstracts
Studies Quality Evaluation:	Not applicable
Data Extraction Form Fields:	Text



As the bibliographic bases were not standardized, the process of excluding duplicate articles (which appeared in both collections) was essential. After selecting via StArt, we performed another manual check using Excel. We sorted the articles by year of publication by reviewing whether there were still duplicates. In some cases, the article was indexed in the main collection in English and in the SCI collection in Spanish or Portuguese and StArt did not identify it as a duplicate, so a final manual check of the selected articles was necessary. Thus, of the 246 articles accepted, 244 were kept for analysis. The 201 rejected did not study presidential countries in Latin America, but in other regions.

ANNEX 3 - JOURNALS CONTEMPLATED IN THE ANALYSIS AND THEIR RESPECTIVE AREAS, CLASSIFIED BASED ON THE SCOPE OF THEIR WEBSITES (Search carried out on WoS and SCI on March 23, 2019)

Political Science and International Relations	**Social Sciences (Interdisciplinary)**	**Right**
American Journal of Political Science	Latin America Today	Derecho del Estado Magazine
American Political Science Review	Archives Europeennes de Sociologie	constitutional studies
Political Analect	Bulletin of Latin American Research	Journal of historical-legal studies
Annual Review of Political Science	co-inheritance	Journal of Constitutional Investigations
Brazilian Political Science Review	Data - Journal of Social Sciences	constitutional questions
British Journal of Political Science	Daedalus, the Journal of the American Academy of Arts & Science	Chilean Derecho Magazine
Colombia International	Economic Development	Derecho Magazine
Comparative Political Studies	E-LATINA – Electronic Journal of Latin American Studies	e-Pública: Electronic Journal of Public Law
Comparative Politics	Spiral (Guadalajara)	Public Law Magazine
Contemporary Politics	iberoamerican	Socio-Legal Studies
democratization	Journal of American Culture	**Public administration**
Electoral Studies	Journal of Interamerican Studies and World Affairs	Public Management and Citizenship Notebooks
Political Studies	Journal of Latin American Studies	Governance: An International Journal of Policy, Administration and Institutions
European Journal of Political Research	Latin American Perspectives	EBAPE.BR notebooks
Government and Opposition	Latin American Politics and Society	Health in Debate (Public Policies)
International Political Science Review	Latin American Research Review	Public Administration Journal
International Studies Perspectives	New Moon: Journal of Culture and Politics	**Economy**
Journal of Democracy	Mexican Studies - Mexican Studies	Institutional Economics Magazine
Journal of Politics	Mobilization: an international quarterly journal	Economic Systems
Journal of Politics in Latin America	New CEBRAP studies	Journal of Economic Integration
Political Role	Public opinion	Journal of the European Economic Association
Parliamentary Affairs	Polis (Santiago)	The Economic Journal
Party Politics	politics and culture	Problems of Development

Political Science and International Relations	Social Sciences (Interdisciplinary)	History
Policy Studies	Public Choice	HISTOReLo. Regional and Local History Magazine
Politics and Government	Regional and Federal Studies	Tzintzun Journal of Historical Studies
Politics and Society	Religación - Journal of Social Sciences and Humanities	Mexican history
Political Research Quarterly	Brazilian Journal of Social Sciences	Studies of modern and contemporary history of Mexico
Politologický časopis - Czech Journal of Political Science	Spanish Journal of Sociological Investigations	History (Santiago)
postdate	Mexican Journal of Political and Social Sciences	
Presidential Studies Quarterly	Social Science Journal	
International Relations (R:I)	society	
Research and Politics	The CS Magazine	
Brazilian Journal of Political Science		
Political Science (Santiago)		
Journal of Sociology and Politics		
SAAP Magazine		
Uruguayan Journal of Political Science		
Scandinavian Political Studies		
Southeast European and Black Sea Studies		
Themes and Debates		
The British Journal of Politics and International Relations		
The Journal of Politics		
World Politics		

Source: Sangalli (2020)

A abordagem da grounded theory como alternativa metodológica qualitativa para pesquisas em Ciência Política

José Alberto de Siqueira Brandão - Instituto de Gestão Pública de Pernambuco - IGPE

Resumo

A já proeminente produção científica no campo de ciência política no Brasil pode ser ampliada com a incorporação de importantes oportunidades metodológicas relacionadas a estudos de natureza qualitativa. Com esta perspectiva em mente, arvora-se a grounded theory (GT) como uma abordagem alternativa, direcionada para proporcionar insumos para a geração de teorias baseadas em dados qualitativos. Neste sentido, o presente artigo tem como objetivo discutir possibilidades de aplicação da abordagem de grounded theory no campo da ciência política. Para tanto, realizou-se um estudo bibliográfico destinado a conhecer suas origens e sua trajetória, enquanto abordagem metodológica, bem como descrever as etapas demandadas para sua aplicação, em consonância com a orientação da corrente social construtivista (CHARMAZ, 2009). Como conclusão, identificou-se que sua inserção no campo da ciência política configura-se como significativo potencial a ser desenvolvido pela academia, sobretudo no âmbito de estudos qualitativos e/ou mistos relacionados à cultura política, às instituições políticas e às políticas públicas.

Palavras-chave: teoria fundamentada; métodos qualitativos; ciência política.

Abstract

The already prominent scientific production in the field of political science in Brazil can be expanded with the incorporation of important methodological opportunities related to studies of a qualitative nature. With this perspective in mind, grounded theory (GT) is raised as an alternative approach, aimed at providing inputs for the generation of theories based on qualitative data. In this sense, this article aims to discuss possibilities for applying the grounded theory approach in the field of political science. To this end, a bibliographic study was carried out aimed at knowing its origins and its trajectory, as a methodological approach, as well as describing the steps required for its application, in line with the orientation of the constructivist social current (CHARMAZ, 2009). As a conclusion, it was identified that its insertion in the field of political science is configured as significant potential to be developed by the academy, especially in the scope of qualitative and / or mixed studies related to political culture, political institutions and public policies.

Keywords: grounded theory; qualitative methods; political science.

1. Introdução

O caminho da construção científica da ciência política tem sido objeto de reflexões de alguns pesquisadores, que passam a questionar os rumos que esta opção pode proporcionar para este campo científico. Nesta perspectiva, há que se destacar a contundência do pensamento do cientista político italiano Giovanni Sartori (2004), ao apresentar seus questionamentos acerca dos caminhos percorridos pela ciência política, como se observa nesta passagem:

> De modo geral, então, eu considero que a ciência política dominante adotou um modelo inadequado de ciência (extraído das ciências exatas e duras) e falhou em estabelecer sua própria identidade (como uma ciência suave) ao falhar em estabelecer sua própria e distinta metodologia. Certamente, minhas estantes estão inundadas de livros cujo título é "metodologia das ciências sociais"; mas esses trabalhos tratam simplesmente de técnicas de pesquisa e processamento estatístico. Quase nada têm a ver com o "método do logos", com o método de pensamento. Portanto, agora temos uma ciência sombria que carece de método lógico e, de fato, ignora a lógica pura e simples. (SARTORI, 2004, p. 785, tradução nossa)

As reflexões de Giovanni Sartori encontraram eco em outros pesquisadores da ciência política. Para Cansino (2007, p.15), o campo da ciência política está em crise, especialmente em função da ausência de reflexão e autocrítica de pesquisadores, que precisariam reconhecer as fragilidades do conhecimento produzido, baseado prioritariamente em métodos quantitativos. Menciona o direcionamento da pesquisa em ciência política para estudos embasados nas chamadas ciências naturais, com instrumentos e técnicas que procuram compreender a realidade a partir de conceitos definitivos e universais. Soares (2005) destaca que "o repúdio aos métodos qualitativos não foi feito após o exame detalhado desses métodos, mas aprioristicamente – tudo o que não fosse quantitativo era classificado, automaticamente, como não-científico".

A existência de diferentes polos metodológicos, situados entre o objetivismo e o subjetivismo, possibilita uma equivocada interpretação dualista, como se tratassem de polos antagônicos. A construção do pensamento de separação entre os estudos de natureza quantitativa e os estudos de natureza qualitativa influenciou alguns autores, como Goertz e Mahoney (2012), a reforçarem os limites entre estas tradições, concebendo-as como "duas culturas" distintas no ensino e na pesquisa das ciências sociais, indicando traços que reforçam esta separação, tais como normas, práticas e ferramentas adotadas por pesquisadores atuantes na área de ciências sociais. Apesar disso, reconhecem que estas culturas não podem ser concebidas como "hermeticamente seladas uma da outra, mas são permeáveis e permitem o cruzamento de fronteiras" (GOERTZ; MAHONEY, 2012, p. 2).

Em outra perspectiva, Burrell e Morgan (1979) argumentam que as dimensões metodológicas existentes entre estes dois polos formam um *continuum* entre os contextos da Sociologia enquanto mudança radical e enquanto regulação, possibilitando a observância de quatro paradigmas para análise da teoria social, sendo estes o humanismo radical, o interpretativismo, o estruturalismo radical e o funcionalismo. Nesta visão, as limitações presentes em cada polo indicam possibilidades de complementaridade entre ambas, buscando-se o alcance de melhores resultados nas pesquisas a serem desenvolvidas.

Destacando o processo de globalização e as mudanças tecnológicas como fatores de transformações metodológicas na ciência política contemporânea, Rezende (2015) ressalta a preocupação com a integração entre estudos de natureza quantitativa e de natureza qualitativa. Enxerga na conexão entre estes diferentes métodos

de pesquisa uma oportunidade para o desenvolvimento de uma consciência metodológica que possibilita a promoção de conhecimento relevante gerado pelas qualidades inerentes a cada um dos tipos.

Alguns estudos, em face de suas especificidades, indicam a necessidade de uma orientação multifacetada sobre o fenômeno que se pretende compreender, demandando a realização de pesquisas que assumam tanto o caráter quantitativo, quanto o caráter qualitativo. Assim, entende-se que a valorização desta dicotomia prejudica o desenvolvimento científico, tendo em vista restringir possibilidades ao se buscar o reconhecimento e a valorização das áreas limítrofes entre estas duas vertentes (PECI, 2004).

A opção por uma abordagem interpretativista do fenômeno social, por exemplo, desperta a busca pelo entendimento de uma rede de representações complexas e subjetivas, em um processo emergente, uma extensão da experiência vivida pelas pessoas encarnadas no mundo "real" (VERGARA; CALDAS, 2005), visando à compreensão dos significados vivenciados pelos sujeitos (GEPHART, 2004). Neste sentido, sob o ponto de vista do relativismo ontológico, a realidade é interpretada como socialmente construída, de modo que indivíduos e mundos são mutuamente concebidos (CLARKE, 1991; STRAUSS, 1993), a partir das experiências vivenciadas (GUBA; LINCOLN, 1994).

Reconhece-se ainda que o contexto histórico, social, político e econômico apresenta características que reforçam as particularidades do fenômeno a ser estudado, demandando multidimensionalidade no olhar do pesquisador. Neste sentido, é importante ressaltar as contribuições metodológicas oriundas de diferentes vertentes da ciência e, em especial, da antropologia social, cujos compreensão dos alicerces metodológicos resgatam a importância da realização de estudos de natureza qualitativa para a compreensão da realidade, por vezes denotando a preocupação com a compreensão do indivíduo (MALINOWSKI, 1992), por outras buscando um entendimento sistêmico (RADCLIFF-BROWN, 1958) e ainda buscando uma integração entre estas correntes, com a construção da teoria da ação, em que são considerados aspectos históricos e artefatos sociais, combinados com observação e inquérito aos envolvidos para a reconstrução de processos sociais, mediante o entendimento da sequência de experiências vivenciadas pelos sujeitos (FELDMAN-BIANCO, 1987).

Nesta perspectiva, destaca-se a relevância da compreensão do contexto, especialmente no sentido cultural, como modelador do comportamento humano (GEERTZ, 2008). Marcus (2001) destaca ainda a importância da consideração de conexões existentes entre indivíduos, acentuada em função do processo globalizante, e que demanda o desenvolvimento de uma etnografia que considere uma abordagem multilocal. As relações contextuais decorrentes dessas interações realizadas por indivíduos, seus semelhantes, natureza e artefatos possibilitam o entendimento das ações desenvolvidas no âmbito social de forma múltipla. Nesta seara, é necessário compreender a possibilidade de criação de conhecimento a partir da interpretação dos significados atribuídos pelos sujeitos às suas experiências por intermédio de suas interações sociais (CHARMAZ, 2014).

O desenvolvimento de pesquisas no campo da ciência política no Brasil tem sido ampliado significativamente nos últimos anos. No entanto, pelo que foi exposto até aqui, observa-se a importância da ampliação de estudos baseados em métodos qualitativos no âmbito da ciência política, especialmente como meio para geração de hipóteses a partir de dados advindos do campo. Uma das possibilidades de promover esta ampliação reside na adoção da abordagem de *grounded theory*, "que tem sido usada com sucesso em diferen-

tes circunstâncias, disciplinas e culturas" (McNABB, 2015, p. 251). Com este enfoque, o presente artigo se propõe a discutir possibilidades de aplicação da abordagem de *grounded theory* no campo da ciência política, a partir de sua caracterização, descrição de suas etapas, apresentação de possibilidades de aplicação e indicação de temas potenciais para a realização de pesquisas que podem se beneficiar com sua inserção no rol de alternativas metodológicas.

2. Surgimento e evolução da grounded theory

A *grouded theory* emerge como abordagem alternativa de pesquisa a partir da necessidade da realização de um estudo sobre a morte e o processo de morte em pacientes com comorbidades graves (CHARMAZ, 2009) em meados dos anos 60. O lançamento do livro "*The Discovery of grounded theory*" em 1967, escrito pelos sociólogos Barney Glaser e Anselm Strauss, é o marco que representa o início da aplicação das estratégias de pesquisa que proporcionam a construção de teorias baseadas em dados, contrapondo-se às estratégias vigentes de dedução de hipóteses a partir de teorias pré-concebidas (McNABB, 2015). A abordagem da *grounded theory* apresenta fundamentos que coincidem com a perspectiva do interacionismo simbólico (BLUMER, 1969), procurando compreender ações coletivas e interações entre indivíduos e a sociedade, "em função do significado atribuído às definições sociais coletivas, formadas por meio do processo de socialização" (BANDEIRA-DE-MELO; CUNHA, 2010, p. 246).

Curiosamente, a *grounded theory* surge em um momento crítico dos estudos e pesquisas sociológicas, em que a aplicação de métodos e técnicas qualitativas encontrava-se em declínio frente ao crescente investimento em recursos quantitativos de pesquisa (CHARMAZ, 2009). O cenário indicava o enfraquecimento da tradição de pesquisa qualitativa na Sociologia e o surgimento de uma nova abordagem qualitativa proporcionou renovação àqueles pesquisadores que se dedicavam aos estudos desta natureza e contribuiu para o desenvolvimento científico na área.

A preocupação desta abordagem reside em explicar as razões da diversidade e variedade de comportamentos ao longo do tempo e em relação a determinado fenômeno (MERRIAM; TISDELL, 2016). Neste diapasão, a *grounded theory* é predominantemente caracterizada como qualitativa e descritiva, visto que a preocupação central reside na interpretação dos dados obtidos com os diversos atores envolvidos, além da observação e análise de documentos relacionados ao objeto de estudo. Portanto, difere-se diametralmente do método hipotético-dedutivo, visto que este se inicia com uma hipótese que virá ou não a ser confirmada para gerar a compreensão da teoria explicativa da realidade objetiva, enquanto a *grounded theory* envolve a própria construção teórica para explicar uma realidade subjetiva, a partir da interpretação do pesquisador e dos participantes da pesquisa.

Apesar de terem formulado conjuntamente a abordagem da *grounded theory*, Glaser e Strauss cindiram o pensamento em duas correntes. Glaser (2002) defende a total separação entre a descrição e a teorização, o que proporcionaria uma clara distinção com outros métodos de caráter qualitativo, como a fenomenologia e a etnografia, que apresentam uma perspectiva com ênfase descritiva, na tentativa de "descobrir explicações causais e realizar previsões sobre um mundo externo e conhecível" (CHARMAZ, 2009, p. 18). Neste sentido, Glaser se aproxima mais da abordagem positivista, ao acreditar que uma realidade objetiva irá emergir dos dados, desde que seja respeitada a neutralidade na participação do pesquisador no processo de construção teórica (CHARMAZ, 2014).

Por outro lado, Strauss apresenta a teoria fundamentada em sentido subjetivista/interpretativista, a partir do desenvolvimento de procedimentos de coleta e análise de dados baseados na interpretação do pesquisador acerca dos resultados alcançados. Nesta corrente, seguida por outros pesquisadores como Juliet Corbin, busca-se com a abordagem da *grounded theory* a construção de teorias substantivas a partir da interpretação do pesquisador acerca da multiplicidade das realidades percebidas e de sua contextualização (STRAUSS; CORBIN, 1994; STRAUSS; CORBIN, 2008; CHARMAZ, 2014).

A abordagem da *grounded theory* tem como foco o processo de geração de teoria a partir de conteúdos particulares (PATTON, 2002; MERRIAM; TISDELL, 2016). Existem divergências entre o paradigma positivista e o paradigma interpretativista no entendimento sobre o que se constitui uma teoria. Para Charmaz (2014, p. 230), enquanto para os positivistas a teoria "busca as causas, favorece as explicações deterministas e enfatiza a generalidade e a universalidade", os interpretativistas defendem que a teoria "visa compreender significados e ações, e como as pessoas os constroem", pressupondo a existência de realidades múltiplas e emergentes, e expondo a verdade como algo provisório e a vida social como um processo. Sob esta ótica, as teorias originadas conforme a grounded theory apresentam caráter empírico, baseadas em dados representativos da realidade. (GOULDING, 2002; STRAUSS; CORBIN, 2008; CHARMAZ, 2014).

A preocupação da abordagem da *grounded theory* reside em se tentar explicar as razões da diversidade e variedade de comportamentos ao longo do tempo e em relação a determinado fenômeno (MERRIAM; TISDELL, 2016). Neste diapasão, a teoria fundamentada é predominantemente caracterizada como qualitativa e descritiva, visto que a preocupação central reside na interpretação dos dados obtidos com os diversos atores envolvidos, além da observação e análise de documentos relacionados ao objeto de estudo. Portanto, difere-se diametralmente do método hipotético-dedutivo, visto que este parte de uma hipótese que virá ou não a ser confirmada para gerar a compreensão da teoria explicativa da realidade objetiva, enquanto a abordagem da *grounded theory* envolve a própria construção teórica para explicar uma realidade subjetiva, a partir da interpretação do pesquisador e dos participantes da pesquisa.

Além da corrente clássica construída por Glaser e Strauss (1967), alguns autores destacam ainda a existência de duas perspectivas interpretativistas mais populares nos estudos da *grounded theory*: a corrente dos procedimentos sistemáticos, desenvolvida por Strauss e Corbin (1994), e a corrente social construtivista, desenvolvida por Charmaz (2014), dentre outros (CRESWELL, 2007; ESLAVA, 2014; KENNY; FOURIE, 2015). Para uma melhor compreensão acerca das três correntes constituídas, alguns autores (ESLAVA, 2014; KENNY; FOURIE, 2015) construíram quadros para apresentar as principais diferenças entre ambas, inspirando a construção do quadro 1. Nele, destacam-se como elementos de distinção entre as correntes estudadas a base filosófica, o momento de uso da literatura, o sistema de codificação, os métodos de codificação e de análise, a seleção e identificação do problema, o nível de operacionalização, a natureza da teoria emergente e os elementos de prova da teoria concebida, conforme se pode observar a seguir:

Quadro 1 – Diferenças entre as correntes da *grounded theory*

	Clássica (Glaser)	Procedimentos sistemáticos (Strauss e Corbin)	Social construtivista (Charmaz)
Base filosófica	Positivismo moderado	Pós-positivismo e interacionismo simbólico	Construtivismo e interacionismo simbólico
Uso da literatura	Somente ao final	Em todas as etapas	Previamente (conceitos sensibilizantes), em todas as etapas e compilada ao final
Sistema de codificação	Original, para descobrir a teoria	Rigoroso, para criar a teoria	Em aberto, para construir a teoria
Método de codificação e análise	Quadro geral baseado em comparação constante: categorias emergem e não podem ser forçadas pela metodologia	Codificação aberta, axial e seletiva; método analítico com etapas estruturadas; procedimentos e técnicas precisam ser detalhados com precisão	Codificação inicial, focalizada, axial e teórica; método analítico com etapas semiestruturadas; procedimentos e técnicas podem sofrer adaptações
Seleção e identificação do problema	Problema emerge do estudo e das percepções conjunta dos atores e dos pesquisadores	Determinado pelo pesquisador antes de ir a campo	
Graus de operacionalização	Baixo devido às dificuldades do amplo quadro de referência.	Alto em função da especificidade das etapas analíticas	
Natureza da teoria emergente	Gera conceitos e suas relações para explicar e interpretar as variações de comportamentos na área substantiva estudada	Gera uma teoria derivada indutivamente sobre um fenômeno, graças a conceitos inter-relacionados	Criação de teorias substantivas por meio da construção da análise indutiva a partir dos dados específicos, da experiência compartilhada e da relação com participantes
Prova da teoria	Análises quantitativas subsequentes	Teste provisório para validar conceitos e suas inter-relações	Esquemas conceituais validados por participantes da pesquisa

Fonte: adaptado de Eslava (2014) e de Kenny e Fourie (2015)

Além da identificação de diferenças entre as três correntes, é importante identificar os pontos de convergência entre elas. Neste sentido, Kenny e Fourie (2015) destacam como princípios de unidade da *grounded theory* o método analítico de comparação constante dos dados, a amostragem teórica visando à identificação do ponto de saturação, a escrita de memorandos ao longo do processo de análise, e a construção de teoria substantiva como fechamento do processo.

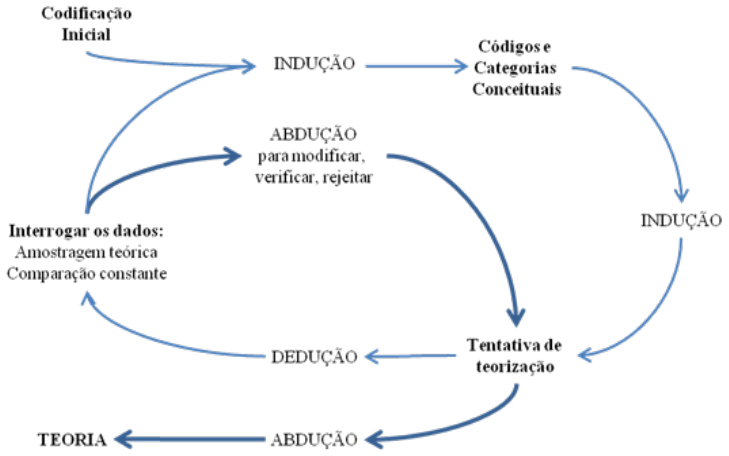
Tomando como ponto de partida esta breve trajetória das correntes da *grounded theory* enquanto abordagem de pesquisa, pretende-se apresentar as etapas requeridas para sua aplicação em estudos e pesquisas científicas. Para tanto, optou-se por considerar a perspectiva social construtivista como alternativa a ser explorada na ciência política, especialmente por representar a modalidade mais dissonante com relação ao modelo positivista.

Uma característica típica deste método é a intensa interação entre o pesquisador e realidades vivenciadas pelos sujeitos, bem como o envolvimento simultâneo na coleta e análise de dados. Outro ponto

que se destaca é a importância do pesquisador nesta abordagem que, ao adotar abordagem social construtivista, precisar estar vigilante em relação às condições nas quais ocorrem diferenças e distinções entre pessoas, de modo a identificar como elas surgem e se mantém, a partir de uma análise "contextualmente situada no tempo, no espaço, na cultura e na situação" (CHARMAZ, 2009, p.179). É nítida a preocupação com o contexto, especialmente no sentido de compreender as contribuições para a formação da realidade observada.

Outra característica da perspectiva social construtivista é a adoção da lógica abdutiva no desenvolvimento das ideias, em que o raciocínio se inicia a partir dos dados e, a partir da análise destes, são formuladas códigos e categorias em um processo iterativo, até o alcance de uma interpretação plausível acerca dos resultados observados (CHARMAZ, 2014). A figura 1, elaborada a partir do caso desenvolvido por Ward, Gott e Hoare (2017), representa esta relação iterativa no processo de raciocínio do pesquisador entre as etapas indutiva, dedutiva e abdutiva na construção teórica conforme a *grounded theory*.

Figura 1 – raciocínio em pesquisa com grounded theory



Fonte: traduzido de Ward, Gott e Hoare (2017)

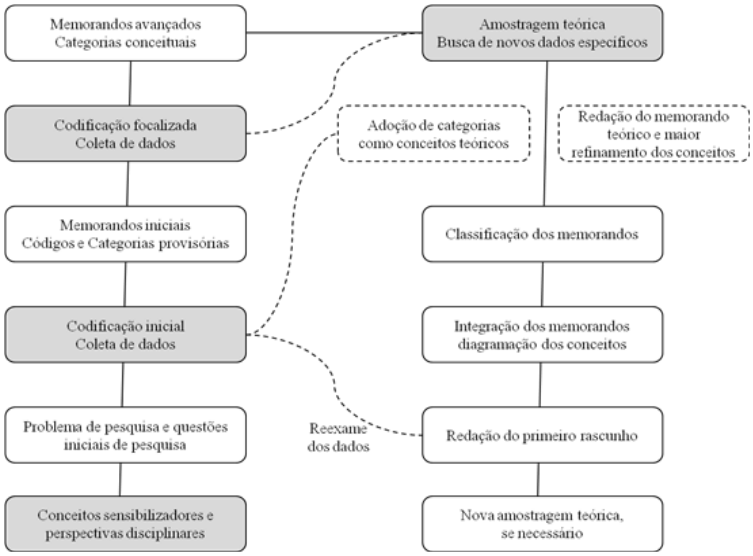
Considerando o que se apresentou na figura acima, ressalta-se a importância do reconhecimento do contexto como um elemento essencial na construção teórica. Torna-se primordial a compreensão dos atores, no exercício do desempenho de suas atividades, residindo o interesse sobre como estes atuam, despertando a curiosidade peculiar às origens dos achados científicos.

Para Charmaz (2014, p.239) a abordagem construtivista "estabelece a prioridade nos fenômenos do estudo e concebe ambos, dados e análise, como tendo sido gerados a partir de experiências compartilhadas e das relações com os participantes e com outras fontes de dados". Decorre desta ideia a proposição da aplicação do método das comparações constantes com os resultados obtidos no campo. As interpretações das realidades observáveis emergem de um processo baseado em interações sucessivas entre a coleta e análise de dados, partindo de comparações, codificações e categorizações dos fenômenos estudados, propiciados por comportamentos, ações e palavras extraídas dos atores pesquisados (STRAUSS; CORBIN, 1994; GOULDING, 2002; STRAUSS; CORBIN, 2008; CHARMAZ, 2014). A próxima seção se destina a explicitar etapas necessárias ao processo de *grounded theory*, conforme a corrente social construtivista.

3. O processo da grounded theory

Nesta seção do artigo serão apresentadas as principais etapas para a elaboração de uma pesquisa mediante a abordagem da *grounded theory*. De início, é importante ressaltar que, diante das perspectivas apresentadas anteriormente, para efeito da construção do presente artigo foi considerada a perspectiva social construtivista, desenvolvida por Kathy Charmaz (2009, 2014). A figura 2 explicita o processo da *grounded theory* conforme a perspectiva adotada neste artigo, demonstrando a sequência de etapas a serem desenvolvidas em uma pesquisa.

Figura 2. O processo da grounded theory



Fonte: adaptado de Charmaz (2009, p. 26)

A distinção entre as perspectivas se verifica desde a primeira etapa do trabalho a ser desenvolvido, haja vista que na perspectiva escolhida existe a necessidade de escolha da literatura para a construção de conceitos sensibilizantes, e que não se verifica nos modelos propostos por Glaser e por Strauss.

3.1. Conceitos sensibilizantes

A iniciativa de conceber conceitos sensibilizantes e se origina no interacionismo simbólico de Blumer (1969) ao considerar que os pesquisadores "iniciam os seus estudos com determinados com determinados interesses de pesquisa e com um conjunto de conceitos gerais" (CHARMAZ, 2009, p. 34). Considerando as características da pesquisa de caráter interpretativista, Blumer (1969) propõe que a lógica da pesquisa social seja alterada e, ao invés de se almejar o encontro com conceitos definitivos, rompa-se com este primado, estabelecendo o que denomina como "conceitos sensibilizantes". Em contraste com os conceitos definitivos, um conceito sensibilizante é um termo amplo, sem características definitivas, funcionando como uma fagulha para que aspectos de uma questão sejam refletidos, servindo como uma instância geral de referência e orientação na abordagem empírica. Deste modo, proporcionam ao pesquisador uma tentativa inicial de perseguir e questionar o surgimento de códigos e categorias, visando ao desenvolvendo de ideias sobre os processos que os dados irão definir (CHARMAZ, 2014).

Os conceitos sensibilizantes representam um direcionamento preliminar sobre aquilo que se espera observar no campo, configurando-se, neste sentido, como efêmeros e passíveis de confrontação com as peculiaridades do contexto estudado e revisão a partir dos dados obtidos no campo (BLUMER, 1969). Com fulcro nesta concepção, os conceitos sensibilizantes se caracterizam como ferramentas provisórias que impulsionarão o pesquisador para, ao confrontarem tais conceitos com os dados oriundos da coleta, começarem a construir os códigos iniciais do estudo. Configuram-se a partir de perspectivas disciplinares adotadas pelos pesquisadores e funcionam, portanto, como gatilhos, como pontos de partida para o desenvolvimento do estudo, oriundos de levantamento prévio de literatura relacionada aos principais termos que pautam o interesse de pesquisa do investigador.

Atualmente, diante dos recursos tecnológicos disponíveis, é possível a realização deste levantamento em bases de pesquisa nacionais e internacionais, de modo a possibilitar a ampliação do escopo inicial do estudo a ser desenvolvido. Entretanto, os conceitos sensibilizantes não precisam ser necessariamente oriundos de referências formais, mas precisam representar o olhar do pesquisador sobre o aspecto a ser abordado, visto que contribuirão para a construção do problema de pesquisa e das questões iniciais de pesquisa, elementos fundamentais para a concepção do projeto de pesquisa a ser elaborado.

No entanto, cabe ressaltar que uma contribuição importante da *grounded theory* reside exatamente nos subsídios oriundos dos dados coletados para a formulação de novas hipóteses que venham a confrontar ou reforçar os conhecimentos existentes. Deste modo, os conceitos sensibilizantes precisam ser confrontados com dados de campo desde o contato inicial do pesquisador, haja vista ter como característica a intensa interação entre pesquisador e as realidades vivenciadas pelos sujeitos, além do envolvimento simultâneo na coleta e análise de dados.

3.2. Definição da amostra inicial

A circularidade entre as etapas de coleta e análise de dados indica que estes processos ocorrem de modo simultâneo e iterativo, até que se alcance a saturação teórica, ou seja, quando não se consegue alcançar nenhuma nova propriedade ou novos insights teóricos a partir da coleta de mais dados (CHARMAZ, 2009; BANDEIRA-DE-MELO; CUNHA, 2010). No entanto, para que seja possível iniciar a coleta de dados, a seleção da amostra inicial é necessária, devendo ser "dirigida à construção da teoria, e não visando à representatividade populacional" (CHARMAZ, 2014, p.8). Deste modo, a seleção amostral é caracterizada como não probabilística, distinguindo-se em relação à tradicional perspectiva da amostragem estatística ou probabilística.

Para que a seleção da amostra ocorra é necessária uma etapa preliminar, que é a definição do objeto de pesquisa, que precisa necessariamente ser "um fenômeno social do qual se busca aprofundar a análise teórica" (LAPERRIÈRE, 2010, p. 357). O fenômeno social escolhido precisa refletir um processo e que se busque explorar sua evolução em um contexto ou situação delimitada. Quando a pesquisa envolve um estudo de caso, a seleção da amostra demanda duas etapas, sendo a primeira a própria seleção do caso e, a segunda, a definição dos indivíduos a serem pesquisados (MERRIAM; TISDELL, 2016).

A escolha do caso a ser estudado depende de diversos fatores, podendo um estudo de caso "ser escolhido porque representa uma instância de um processo, problema ou preocupação" (MERRIAM, 2009, p. 41). Os estudos sob a égide da *grounded theory* envolvem intencionalidade, mediante estabelecimento de critérios que possibilitem a identificação de participantes que tenham experiências com o fenômeno estudado e

possam contribuir para a elaboração teórica, podendo ser iniciado com uma amostra homogênea de indivíduos e, posteriormente, englobar grupos heterogêneos (CRESWELL, 2007; STRAUSS; CORBIN, 2008).

É importante realizar uma importante distinção entre amostra inicial e amostragem teórica, componente do processo de *grounded theory*. A amostra inicial, constituída pela definição não probabilística de caso(s) e de sujeito(s) a participarem da pesquisa, representa apenas o ponto de partida, pelo qual serão obtidos os dados. A amostragem teórica representa o refinamento dos dados em códigos, o desenvolvimento das propriedades das categorias que possibilitem a explicação do fenômeno estudado. Assim, "ao empregar a amostragem teórica, o pesquisador busca pessoas, eventos ou informações para elucidar e definir os limites e relevâncias das categorias", obtendo amostras para a construção das categorias teóricas (CHARMAZ, 2009, 249).

3.3. Coleta e análise de dados

Definida a amostra inicial, o investigador poderá encetar a fase de coleta de dados, estabelecendo as técnicas mais adequadas à amostra definida. Apesar de Charmaz (2009) enfatizar as possibilidades de coletar dados mediante processos de observação etnográfica, de entrevista intensiva e de análise textual de textos extraídos ou de textos existentes, resta evidente que outras possibilidades de coleta de dados podem ser adotadas no processo de *grounded theory*. A possibilidade de realização de entrevistas semiestruturadas em pesquisas baseadas na *grounded theory* no âmbito da ciência política foi inclusive ressaltada por Rocha (2020). A adoção de diferentes fontes de dados proporciona enriquecimento do estudo, possibilitando a realização de comparações, confrontos e questionamentos que conduzem à geração de maior refinamento das categorias formuladas. Independente de qual seja a técnica adotada para coletar dados, é importante que seja considerado "o modo como os participantes, invocam as ideias, os costumes e os relatos tanto no que diz respeito à cultura de modo geral quanto em relação à cultura local da qual eles fazem parte", pois isto contribuirá para o entendimento do contexto social, facilitando o processo de codificação e categorização (CHARMAZ, 2009, p. 65).

Concomitantemente ao processo de coleta de dados se inicia o processo de análise de dados, que juntos se configuram, no âmbito da *grounded theory*, como etapas indissociáveis (BANDEIRA-DE-MELO; CUNHA, 2010). Em consonância com Charmaz (2009, p. 69), a análise de dados começa com o processo de codificação, que "significa categorizar segmentos de dados com uma denominação concisa que, simultaneamente, resume e representa cada parte dos dados". Neste sentido, o procedimento central do processo analítico de uma pesquisa realizada conforme a *grounded theory* é o pensamento comparativo sistemático entre os dados obtidos ao longo de cada etapa de coleta e os códigos e categorias construídos neste processo, visando à identificação de convergências e divergências que contribuam para a saturação teórica.

A codificação, descrita por Charmaz (2014) como o processo de seleção, separação, classificação e denominação que resume e representa os dados, objetiva a construção da estrutura analítica do estudo, funcionando como uma fase intermediária entre os dados e a teoria. A codificação é realizada em etapas, sendo a primeira denominada como codificação inicial, que corresponde à fase em que ocorrem as denominações sobre os dados obtidos, de modo que é desenvolvida uma definição e classificação inicial na forma de uma síntese. A codificação inicial necessita estar centrada nos dados, permitindo a geração de códigos preliminares que sejam temporários, mas que possibilitem comparações a partir dos dados disponíveis. Esta etapa é compreendida

como a que "gera os ossos" da análise, ou seja, permite a ocorrência do processo de integração teórica, que tem o papel de "agregar esses ossos para formar um esqueleto do trabalho" (CHARMAZ, 2009, p.70).

A codificação inicial pode ser desenvolvida de diferentes modos, como a codificação palavra por palavra, que permite observar a estrutura e o fluxo das palavras, a codificação linha a linha, que possibilita a identificação de códigos emergidos a partir das palavras expressadas neste processo, e também codificação incidente por incidente, em que são comparados os incidentes observados na fala do entrevistado (CHARMAZ, 2009). Existe a possibilidade de adoção de programas computacionais para dar suporte ao processo de codificação, em função de possuírem "grande capacidade de organizar enormes quantidades de dados, facilitando a análise e auxiliando a comunicação entre os membros de uma equipe de pesquisa" (MERRIAM; TISDELL, 2016, p. 221). Nesta fase, pode ser observada ainda a emergência de códigos *in vivo*, que se constituem em termos amplamente utilizados e emitidos pelos entrevistados e que podem representar o significado dado por estes às suas opiniões e atitudes. A flexibilidade desta modalidade de codificação possibilita a captura de tais aspectos da linguagem, sem que estes códigos *in vivo* sejam perdidos pela interpretação do pesquisador.

A realização da codificação inicial possibilita o ajuste do estudo ao mundo empírico no processo de construção dos códigos e, posteriormente, o desenvolvimento das categorias. Possibilita ainda a percepção de relevância do estudo desenvolvido, visto que serve como "esqueleto" para a interpretação do pesquisador acerca do fenômeno estudado, permitindo que outras pessoas possam compreender o modo com que ele "estabelece relações entre os processos implícitos e as estruturas visíveis" (CHARMAZ, 2009, p. 83). O meio pelo qual isto se torna possível é através da elaboração de memorandos iniciais, que se configuram como registros de pensamentos do pesquisador sobre os dados, de modo a gerar novas ideias, cristalizar questões e direções assumidas ao longo do estudo, mediante comparação entre dados, códigos, categorias e conceitos sensibilizantes (CHARMAZ, 2014).

Uma segunda etapa de análise é denominada codificação focalizada, que "significa utilizar os códigos anteriores mais significativos e/ou frequentes para analisar minuciosamente grandes montantes de dados" (CHARMAZ, 2009, p. 87), com o intuito de gerar e validar propriedades e categorias provisórias por meio de comparações constantes com novos dados empíricos. Os códigos originalmente encontrados são agrupados, de modo a possibilitar a descoberta de possíveis associações, desconexas ou despercebidas na etapa anterior, de modo que estes possam proporcionar compreensão ao fenômeno estudado, a partir da emergência de novas ideias.

O refinamento dos dados, códigos e categorias provisórias é realizado através da elaboração de memorandos avançados, onde são descritas relações entre categorias e subcategorias, em um processo de reagrupamento dos códigos gerados anteriormente. Este processo é denominado por Strauss e Corbin (2008) como codificação axial e procura explorar questões sobre quando, onde, por que, quem e como, de modo a gerar explicações mais precisas sobre o fenômeno e proporcionar coerência à análise (CEPELLOS; TONELLI, 2020). O processo sistemático de comparações constantes dos dados gera confrontos entre as proposições criadas de modo a se verificar a consistência da análise realizada, proporcionando maior robustez às categorias estabelecidas e suas propriedades (CHARMAZ, 2014).

Como sinalizado anteriormente, este processo de refinamento de categorias com a saturação de dados é denominado como amostragem teórica. O objetivo da amostragem teórica é elaborar e refinar as categorias

que constituem a proposição teórica, utilizando dados novos, que podem ser obtidos mediante entrevistas de *follow-up*, para preencher possíveis lacunas existentes na formulação das categorias, bem como possibilitar seu desenvolvimento analítico, reforçando suas propriedades até que estas sejam consolidadas com dados disponíveis (CHARMAZ, 2009).

A amostragem teórica segue em um processo iterativo de coleta e análise que se encerra com a classificação e integração de memorandos, com a formulação das categorias definitivas e com a proposição de diagrama elaborado mediante a utilização de um processo de inferência abdutiva, na qual o raciocínio sobre a experiência possibilita a elaboração de uma teoria substantiva, desenvolvida para o campo imediato de estudo (GOULDING, 2002; CHARMAZ, 2009). Este processo se desenvolve até que não se apresentem novas contribuições significativas oriundas dos dados obtidos que auxiliem a capacidade explicativa da formulação elaborada.

3.4. Validade e confiabilidade da pesquisa

Em estudos de natureza qualitativa um tópico que demanda esforço adicional do pesquisador se refere à definição dos critérios de validade e confiabilidade do estudo. Esta preocupação transcende a estrutura de etapas representativas do processo de *grounded theory* construído sob a ótica social construtivista, anteriormente apresentado.

Antes de tudo, é importante destacar o aspecto ético inerente a este tipo de pesquisa. A adoção de conduta ética é essencial, visto que é a partir desta que se possibilita a construção da credibilidade do pesquisador. Merriam e Tisdell (2016) destacam que o comportamento ético precisa estar presente nas principais formas de coleta de dados em pesquisas qualitativas. Em entrevistas, os indivíduos podem sentir que sua privacidade está sendo invadida, ou podem ficar embaraçados com certas questões formuladas, ou ainda podem afirmar coisas que não poderiam revelar. Os procedimentos de observação, de igual modo, podem representar armadilhas como o envolvimento excessivo do pesquisador em questões relativas ao campo de pesquisa, de modo a ter sua capacidade de julgamento influenciada pela situação observada. Semelhante risco também está presente nos dados disponibilizados através de documentos, visto que estes podem introduzir informações de caráter sigiloso para a organização envolvida.

Ultrapassada esta preliminar, cabe o entendimento de que os estudos qualitativos se diferenciam dos estudos quantitativos no tocante aos critérios de validade e confiabilidade. Tal lógica é nítida quando considerado o modo como a realidade é interpretada. Desta maneira, os estudos que assumem caráter construtivista assumem que a realidade é tida como múltipla, oriunda de construções sociais baseadas nas experiências vivenciadas pelos indivíduos. Busca-se interpretar os significados construídos pelos sujeitos para que seja possível a formulação da concepção própria do pesquisador sobre o objeto investigado. Em função de tal peculiaridade, são apresentados critérios que asseguram qualidade a uma pesquisa de cunho qualitativo, em especial destaque àqueles que proporcionam o reconhecimento de validade e de confiabilidade.

Considerando especificamente a abordagem da *grounded theory*, Strauss e Corbin (2008) recomendam alguns critérios para a verificação da qualidade da pesquisa, apresentados em forma de questionamentos. Para estes autores é preciso que haja rigor no processo de codificação e pesquisa, qualidade na elaboração dos conceitos, relacionamentos sistemáticos entre os conceitos formulados, densidade conceitual, faixas de variação e especificidade, significância dos achados e sensitividade teórica.

De modo mais amplo, Creswell (2010) aponta que questionamentos sobre a validade da pesquisa permeiam todas as fases do trabalho de pesquisa, visto que representam um pesquisador que adota procedimentos que verificam a precisão dos resultados. Merriam e Tisdell (2016) defendem que a validade de uma pesquisa qualitativa assume características internas e externas. A validade interna (ou credibilidade) está relacionada à congruência dos dados obtidos com a realidade pesquisada. A validade externa (ou transferibilidade) corresponde à condição em que os resultados alcançados no estudo podem ser reproduzidos em outros contextos ou situações, delimitando a amplitude de aplicação dos achados obtidos no estudo.

Merriam e Tisdell (2016) destacam como principais possibilidades de validação interna: a triangulação, a validação pelos respondentes, a participação adequada do pesquisador na coleta de dados, a reflexividade do pesquisador, a revisão por pares, a auditoria, a descrição densa e rica do fenômeno e a variação máxima na aplicação dos achados. Cho e Lee (2014) apontam que, nos estudos sob a abordagem da teoria fundamentada, além das possibilidades apresentadas, cabe ainda a apresentação de informações discrepantes ou negativas identificadas ao longo do estudo.

No tocante à confiabilidade (ou consistência), Merriam e Tisdell (2016) apontam que esta se refere à condição da obtenção dos mesmos achados (ou semelhantes) a partir da replicação da pesquisa, mediante circunstâncias semelhantes. No entanto, as autoras afirmam que a confiabilidade não se aplica em pesquisas qualitativas, visto que a concepção de múltiplas realidades socialmente construídas não é condizente com a possibilidade de sua replicação. "Os pesquisadores [qualitativos] buscam descrever e explicar o mundo como este é experimentado" (MERRIAM; TISDELL, 2016, p. 230). Deste modo, avalia-se a consistência e segurança do estudo, associando-a as questões de validade interna, esperando que outros pesquisadores possam, a partir dos mesmos dados coletados, ter entendimentos semelhantes.

4. Possibilidades de aplicação da grounded theory em Ciência Política

Após a apresentação da trajetória da abordagem de *grounded theory* enquanto alternativa de pesquisa qualitativa e das etapas do processo metodológico vinculado à corrente social construtivista, o desafio que se apresenta é verificar as possibilidades de sua aplicação no campo da ciência política. Para que fosse possível identificar, ainda que em uma perspectiva preliminar, a produção científica com utilização da *grounded theory* no campo da ciência política, procedeu-se levantamento de artigos publicados e registrados no portal Periódicos da CAPES. Para tanto, considerou-se como termos de busca "grounded theory" e "ciência política". O resultado desta busca indicou um total de 62 artigos, sendo que apenas 23 escritos na língua portuguesa.

No entanto, ao se aprofundar na leitura destes artigos selecionados, verificou-se que, de fato, apenas 3 artigos se apresentavam com o perfil desejado. Os demais artigos, apesar de contarem com os termos escolhidos, ou usaram alguma referência que continha os termos, ou se vinculavam predominantemente a outras áreas ou ainda foram escritos originalmente em outra língua. Desta maneira, cabe o registro dos artigos: "Derrotas eleitorais e carreiras políticas: o caso do PT no Rio Grande do Sul", de Frédéric Louault (2011), publicado na Revista de Sociologia e Política; "Jogando luzes sobre os determinantes burocráticos e de gestão da política de assistência social nos municípios brasileiros", de Natália Sátyro e Eleonora Cunha (2019), publicado na revista Opinião Pública; e "A demanda como determinante na

execução dos serviços socioassistenciais nos municípios brasileiros", de Raquel D'Albuquerque e Natália Sátyro (2020), também publicado na revista Opinião Pública.

A realidade explicitada pelos dados deste levantamento preliminar pode levantar ao menos um questionamento: os pesquisadores do campo de ciência política realmente não apresentam interesse em desenvolver estudos baseados na *grounded theory* ou não a reconhecem como uma alternativa viável para realização de pesquisas qualitativas ou mistas? Para que seja possível responder tal questão é preciso compreender melhor o que os dados obtidos no levantamento querem representar.

Para tanto, alterou-se o escopo do levantamento anterior, substituindo apenas o termo "*grounded theory*" pelo termo "qualitativ*". O número de artigos selecionados sobe para 1.201, sendo 625 escritos em português, o que conduz ao entendimento de que, diante do volume de trabalhos qualitativos identificados, ainda são parcos os investimentos acadêmicos em *grounded theory* no campo da ciência política no Brasil, possivelmente pela ausência de seu reconhecimento como uma alternativa qualitativa viável. Seria esta mais uma evidência de deficiências na formação metodológica de cientistas políticos, apontadas por Soares (2005)? De fato, o trabalho desenvolvido por Barberia, Godoy e Barboza (2014) acerca da evolução do ensino de métodos e técnicas de pesquisas em ciência política no Brasil ratifica a carência de formação em análise de modelos formais e em técnicas de pesquisa qualitativa.

Complementarmente, percebe-se ainda que críticas foram formuladas sobre a aplicação da *grounded theory* no campo da ciência política. Sintetizando-as, Becker (2012) questiona sobretudo a corrente clássica por não facultar a possibilidade de referências prévias, o que se encontra superado pelas demais correntes. Também questiona a limitação de sua aplicação no âmbito de estudos voltados para relações internacionais, cuja unidade social seja um Estado-nação ou uma região do mundo, o que é confrontado por Glaser e por Strauss, ao defenderem esta possibilidade com a abordagem da *grounded theory*. McNabb (2015), ao relatar estudo desenvolvido por Cook e Barry (1995), indica como desvantagem da aplicação da *grounded theory* uma superabundância de dados brutos gerados pelas pesquisas, algo que atualmente poderia ser superado com a adoção de programas computacionais de apoio ao processo de codificação, conforme relatado na seção anterior.

Apesar de estas críticas serem embasadas, estes mesmos pesquisadores identificam potencial na aplicação da *grounded theory* no campo da ciência política. Além da pesquisa de Cook e Barry (1995), sobre as interações de políticas públicas entre os setores público e privado, McNabb (2015) dedica um capítulo de seu livro "Research methods for political science: quantitative and qualitative approaches" para discorrer sobre o potencial da *grounded theory*, referenciando como exemplo o trabalho realizado por King, Feltey e Susel (1998) sobre as causas subjacentes da antipatia pública no processo político e as formas de melhorar a participação nas decisões de formulação de políticas da administração pública.

De um modo mais amplo, Becker (2012) enxerga um potencial mais da aplicação de *grounded theory* em estudos de sociologia política e de políticas públicas, bem como estudar temas voltados para a cultura política, como o comportamento do eleitor ou respostas relativas às campanhas políticas. Ademais, reconhece que a *grounded theory* poderia suprir algumas carências no campo da ciência política, especialmente em processos de geração de teorias, de modo a se aproximar dos fenômenos que procura explicar. Besana (2018) sinaliza ainda a possibilidade de aplicação da *grounded theory* em estudos sobre participação política em nível micropolítico, que dificilmente seriam alcançadas mediante a adoção de outras perspectivas de análise.

Outro campo em que se vislumbra um elevado potencial de exploração de métodos qualitativos em ciência política é o da avaliação de políticas públicas. Neste sentido, Rossi, Lipsey e Henry (2019, p. 576) entendem que os pesquisadores sociais que versam por esta perspectiva de atuação são orientados "para a avaliação formativa, ou seja, fazer um programa funcionar melhor alimentando informações para seus gerentes e patrocinadores (...) eles tendem a confiar em informações sobre as experiências vividas por aqueles que estão sendo atendidos pelo programa". Neste diapasão, estudos realizados mediante adoção da *grounded theory* poderiam contribuir para o processo de avaliação de políticas públicas, tanto no sentido da análise *ex ante*, visando identificar problemas e necessidades vivenciados por indivíduos e assim contribuir para a formulação de políticas, como também na análise *ex post*, na linha de se verificar se os efeitos proporcionados pelas políticas efetivamente contribuíram para a solução dos problemas identificados, com fundamento nas experiências vivenciadas por participantes beneficiários das políticas.

Observa-se, portanto, que apesar de existirem críticas ao uso de métodos qualitativos nas pesquisas em ciência política e, especificamente, direcionados à *grounded theory*, estas críticas tem sido continuamente combatidas com a ampliação de sua aplicação no campo em discussão, como também pela incorporação de novas ferramentas computacionais de apoio à pesquisa que incorporam técnicas associadas à abordagem em debate. Para além da tradição no desenvolvimento de pesquisas de natureza quantitativa, é preciso que haja também um crescente investimento na formação de pesquisadores sobre métodos qualitativos e mistos. Apenas com este investimento e com estímulo à produção científica oriunda de trabalhos desenvolvidos com métodos qualitativos (incluindo-se a *grounded theory*) ou com métodos mistos é que se pode esperar a superação de lacunas que ainda insistem em se manter no campo da ciência política, por não se aproveitar possibilidades de aplicação que foram apresentadas nesta seção.

5. Considerações finais

O principal intuito observado na construção do presente artigo foi o de estimular a realização de pesquisas de natureza qualitativa em ciência política, especialmente mediante o uso da abordagem social construtivista da *grounded theory*. A preocupação com esta lacuna se fez presente em trabalhos de Giovanni Sartori e de outros cientistas políticos, que ressaltam a valorização de pesquisas de natureza quantitativa neste campo de estudos.

Para que este debate fosse possível, foi apresentada a origem e a evolução da *grounded theory* enquanto abordagem metodológica. Ficou demonstrada que suas raízes estão fincadas na área de ciências sociais, mediante o trabalho realizado pelos sociólogos Barney Glaser e Anselm Strauss, sob influência interacionismo simbólico de Herbert Blumer. O rompimento entre seus idealizadores implicou na criação de três principais correntes: a corrente clássica, seguida por Glaser, a corrente dos procedimentos sistemáticos, desenvolvida por Strauss e Corbin (1994), e a corrente social construtivista, proposta por Kathy Charmaz. Considerando ser esta última corrente a que mais se coaduna com pesquisas em ciência política, em função de melhor responder às críticas oriundas de pesquisadores deste campo, optou-se por utilizá-la como referência neste artigo.

Deste modo, procedeu-se a apresentação das principais etapas do processo de *grounded theory* segundo Charmaz (2009). Destacou-se a importância da construção dos conceitos sensibilizantes

como ponto de partida para o desenvolvimento da pesquisa, bem como os critérios de constituição da amostra inicial. Tratou-se ainda da natureza indissociável do processo de coleta e análise de dados, enfatizando-se o caráter iterativo existente entre estas etapas até se alcançar a saturação teórica e a elaboração da teoria substantiva. Além destas etapas, discutiu-se ainda a validade e a confiabilidade de pesquisas realizadas mediante a abordagem de *grounded theory*, ressaltando-se a importância da aplicação de técnicas para reforçar estes critérios de qualidade da produção científica.

Por fim, foi apresentado um levantamento que indica a limitada produção de conhecimentos mediante a adoção da abordagem de *grounded theory* em ciência política no Brasil. Apesar de existirem muitas críticas a esta abordagem metodológica, observa-se que tais críticas começam a ser superadas, quer pela evolução de suas técnicas e ferramentas, quer pela própria adoção em pesquisas desenvolvidas neste campo. Deste modo, foram evidenciadas algumas possibilidades de aplicação da *grounded theory*, especialmente no âmbito de cultura política, de instituições políticas e de políticas públicas.

Com isso, espera-se ter contribuído para que pesquisadores vinculados à ciência política possam desenvolver estudos qualitativos e/ou mistos baseados na abordagem da *grounded theory*. De igual modo, é preciso um esforço significativo no sentido de incorporar esta perspectiva na formação de novos pesquisadores vinculados ao campo. Isto se configura como um importante desafio para a academia, de modo que seja aproveitado o potencial de novos conhecimentos que podem ser gerados mediante iniciativas que venham a convergir com esta possibilidade.

REFERÊNCIAS BIBLIOGRÁFICAS

Bandeira-de-Mello, R.; Cunha, C. (2010), "Grounded theory", in C. Godoi; R. Bandeira-de-Mello; A. Silva, A. (org.). *Pesquisa qualitativa em estudos organizacionais: paradigmas, estratégias e métodos*. São Paulo, Saraiva.

Barberia, L.; Godoy, S.; Barboza, D. (2014), "Novas Perspectivas sobre o 'Calcanhar Metodológico': O Ensino de Métodos de Pesquisa em ciência política no Brasil", *Teoria & Sociedade*, vol. 22, no. 2: 156-186.

Becker, Brittney. (2012), "The Grounded Theory Method and its uses for Political Science", in B. Becker; G. Kaufmann (org.), Series of Paper: Methods of Field Research.

Besana, P. (2018), "Notas sobre el uso de la etnografía y la teoría fundamentada en ciencia política. Un análisis amplio de la participación política y el Estado en asentamientos informales de la periferia de Buenos Aires, Argentina", *Universitas humanística*, vol. 86, nº. 2:107-137.

Blumer, H. (1969), *Symbolic interactionism: perspective and method*. Englewood Cliffs, Prentice-Hall.

Burrell, G.; Morgan, G. (1979), *Sociological paradigms and organisational analysis.* Aldershot, Ashgate.

Cansino, C. (2007), "Adiós a la ciencia política: crónica de uma morte anunciada", *Temas y debates*, vol. 14, nº 2:13-29.

Cepellos, V.; Tonelli, M. (2020), "Grounded theory: passo a passo e questões metodológicas na prática", *Revista de Administração Mackenzie*, vol. 21, nº 5:1–29.

Charmaz, K. (2009), *A construção da teoria fundamentada: guia prático para análise qualitativa.* Porto Alegre, Artmed.

Charmaz, K. (2014), *Constructing grounded theory*. London, Sage.

Cho, J.; Lee, E. (2014), "Reducing confusion about grounded theory and qualitative content analysis: similarities and differences", *The Qualitative Report*, v. 19, nº 32:1-21.

Clarke, A. (1991), "Social words/arenas theory as organizational theory", in D. Maines (org.) *Social organization and social process: essays in honor of Anselm Strauss*. New York, Aldine de Gruyter.

Creswell, J. W. (2007), *Qualitative inquiry and research design*. Thousand Oaks, Sage.

Creswell, J. W. (2010), *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. Porto Alegre, Artmed.

D'albuquerque, R.; Sátyro, N. (2020), "A demanda como determinante na execução dos serviços socioassistenciais nos municípios brasileiros", *Opinião Pública*, vol. 26, nº 2:246-282.

Eslava, A. (2014), "Análisis cualitativo y cuantitativo para los estudios políticos: trilateralismo, metodología experimental y grounded theory", *Cinta de moebio*, vol. 3, nº. 51:111-126.

Feldman-Bianco, B. (1987), *A Antropologia das sociedades contemporâneas*, Global, São Paulo.

Geertz, C. (2008), *A interpretação das culturas*. Rio de Janeiro, LTC.

Gephart, R. (2004), "From the editors: qualitative research and the academy of management journal". *Academy of Management Journal*, v. 47, nº. 4:454-462.

Glaser, B. (2002), "Constructivist grounded theory?", *Forum: qualitative social research*, vol. 3, nº. 3:1-14.

Glaser, B.; Strauss, A. (1967), *The discovery of grounded theory: strategies for qualitative research*. New Brunswick, Aldine.

Goertz, G.; Mahoney, J. (2012), *A tale of two cultures: qualitative and quantitative research in the social sciences*. Oxfordshire, Princeton University Press.

Goulding, C. (2002), *Grounded theory: a practical guide for management, business and market researchers*. Thousand Oaks, Sage.

Guba, E.; Lincoln, Y. (1994), "Competing paradigms in qualitative research", in N. Denzin; Y. Lincoln, (ed.). *Handbook of qualitative research*. Thousand Oaks, Sage.

Kenny, M.; Fourie, R. (2015), "Contrasting classic, straussian, and constructivist grounded theory: methodological and philosophical conflicts", *The qualitative report*, vol. 20, nº 8:1270-1289.

Laperrière, A. (2010), "A teorização enraizada (grounded theory): procedimento analítico e comparação com outras abordagens similares", in J. Poupart; J. Deslauriers; L. Groulx; A. Laperrière; R. Mayer; A. Pires. *A pesquisa qualitativa: enfoques epistemológicos e metodológicos*. Petrópolis, Vozes.

Louault, F. (2011), "Derrotas eleitorais e carreiras políticas: o caso do PT no Rio Grande do Sul", *Revista de Sociologia e Política*, vol. 19, nº 40:233-253.

Malinowski, B. (1992), *Los argonautas del Pacífico Occidental*, Península, Barcelona.

Marcus, G. (2001), Etnografía en/del sistema mundo: el surgimiento de la etnografía multilocal, *Alteridades*, vol. 11, nº 22: 111-127.

McNabb, D. (2015), *Research methods for political science: quantitative and qualitative approaches*. New York, Routledge.

Merriam, S. (2009), *Qualitative research: a guide to design and implementation*. San Francisco, Jossey-Bass.

Merriam, S.; Tisdell, E. (2016), *Qualitative research: a guide to design and implementation.* San Francisco, Jossey-Bass.

Peci, A. (2004), "Além da dicotomia objetividade-subjetividade", in M. Vieira; D. Zouain (org.), *Pesquisa qualitativa em Administração*. Rio de Janeiro, FGV.

Patton, M. (2002), *Qualitative research and evaluation methods*. Thousand Oaks, Sage.

Radcliff-Brown, A. (1958), *El método de la Antropología Social*, Editorial Anagrama, Barcelona.

Rezende, F. (2015), "Transformações metodológicas na ciência política contemporânea", *Política Hoje*, vol. 1, nº. 24:13-45.

Rocha, V. (2020), "Da teoria à análise: uma introdução ao uso de entrevistas individuais semiestruturadas na ciência política", *Política Hoje*, vol. 1, nº. 29:1-26.

Rossi, P.; Lipsey, M.; Henry, G. (2019), *Evaluation: a systematic approach*. Thousand Oaks, Sage.

Sartori, G. (2004), "Where is political science going?", *Political science and politics*, vol. 37, nº. 4: 785-789.

Sátyro, N.; Cunha, E. (2019), "Jogando luzes sobre os determinantes burocráticos e de gestão da política de assistência social nos municípios brasileiros", *Opinião Pública*, vol. 25, nº 2:401-432.

Soares, G. (2005), "O calcanhar metodológico da ciência política no Brasil", *Sociologia, problemas e práticas*, vol. 1, nº. 48:27-52.

Strauss, A. (1993), *Continuous permutations of action*. New York, Aldine de Gruyter.

Strauss, A.; Corbin, J. (1994), "Grounded theory methodology: an overview", in N. Denzin; Y. Lincoln (ed.). *Handbook of qualitative research*. Thousand Oaks, Sage.

Strauss, A.; Corbin, J. (2008), *Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento de teoria fundamentada*. Porto Alegre, Artmed.

Vergara, S.; Caldas, M. (2005), "Paradigma interpretativista: a busca da superação do objetivismo funcionalista nos anos 1980 e 1990", *Revista de Administração de Empresas*, vol. 45, nº. 4:66-72.

Ward, K.; Gott, M.; Hoare, K. (2017), "Analysis in grounded theory: how is it done?", *SAGE Research Methods Cases*. Thousand Oaks, Sage.

Grounded theory approach as a qualitative methodological alternative for Political Science research

José Alberto de Siqueira Brandão, Instituto de Gestão Pública de Pernambuco - IGPE[1]

Resumo

A já proeminente produção científica no campo de ciência política no Brasil pode ser ampliada com a incorporação de importantes oportunidades metodológicas relacionadas a estudos de natureza qualitativa. Com esta perspectiva em mente, arvora-se a grounded theory (GT) como uma abordagem alternativa, direcionada para proporcionar insumos para a geração de teorias baseadas em dados qualitativos. Neste sentido, o presente artigo tem como objetivo discutir possibilidades de aplicação da abordagem de grounded theory no campo da ciência política. Para tanto, realizou-se um estudo bibliográfico destinado a conhecer suas origens e sua trajetória, enquanto abordagem metodológica, bem como descrever as etapas demandadas para sua aplicação, em consonância com a orientação da corrente social construtivista (CHARMAZ, 2009). Como conclusão, identificou-se que sua inserção no campo da ciência política configura-se como significativo potencial a ser desenvolvido pela academia, sobretudo no âmbito de estudos qualitativos e/ou mistos relacionados à cultura política, às instituições políticas e às políticas públicas.

Palavras-chave: teoria fundamentada; métodos qualitativos; ciência política

Abstract

The already prominent scientific production in Brazilian political science can be broadened by incorporating important methodological opportunities related to studies of a qualitative nature. With this perspective in mind, grounded theory (GT) is raised as an alternative approach, aimed at providing inputs to generate theories based on qualitative data. Consequently, this article aims to discuss possibilities for a grounded theory approach in political science. To this end, a bibliographic study was carried out aimed at knowing the origins and trajectory of grounded theory as a methodological approach, as well as describing the steps required for its application, in line with the orientation of the constructivist social current (CHARMAZ, 2009). We conclude by identifying that its insertion in the field of political science is a significant opportunity to be developed by academia, especially within the scope of qualitative and/or mixed-methods related to political culture, political institutions, and public policy.

Keywords: grounded theory; qualitative methods; political science.

1 Public Management Institute of Pernambuco.

1. Introdution

Political science's path of scientific construction has been the object of some researchers' reflections, who have begun questioning the directions this option may provide this scientific field. Under this perspective, the relevance of Italian political scientist Giovanni Sartori's (2004) thought must be highlighted, when questioning the paths taken by political science, as observed in this section:

> Overall, then, I take the view that mainstream political science has adopted an unsuited model of science (drawn from the hard, exact sciences) and has failed to establish its own identity (as a soft science) by failing to establish its own, distinctive methodology. To be sure, my shelves are inundated by books whose title is "methodology of the social sciences"; but these works simply address research techniques and statistical processing. They have almost nothing to do with the "method of logos," with the method of thinking. So we now have a dismal science that lacks logical method and indeed ignores pure and simple logic. (SARTORI, 2004, p. 785).

Giovanni Sartori's reflections were echoed by other political science researchers. For Cansino (2007, p. 15), the political science field is in crisis, especially given the absence of reflection and self-criticism by researchers, who would need to recognize the fragilities of the knowledge produced, based primarily on quantitative methods. He mentions the direction of research in political science towards studies based on the so-called natural sciences, with tools and techniques that seek to understand reality from definitive and universal concepts. Soares (2005) highlights that the "rejection of qualitative methods was not done after a detailed analysis of those methods, but a priori – all that was not quantitative was automatically classified as non-scientific".[2]

2 In the original: o repúdio aos métodos qualitativos não foi feito após o exame detalhado desses métodos, mas aprioristicamente – tudo o que não fosse quantitativo era classificado, automaticamente, como não-científico".

The existence of different methodological extremes, situated between objectivism and subjectivism creates a mistaken dualist interpretation, as if they were antagonistic poles. The construction of the separation line of thinking among studies of quantitative and qualitative nature influenced some authors, such as Goertz and Mahoney (2012) to reinforce the boundaries between these traditions, conceiving them as "two cultures", distinct in teaching and research in the social sciences, indicating the characteristics that reinforce this separation, such as rules, practices, and tools adopted by social sciences researchers. Despite this, they acknowledge that these cultures cannot be thought of as "hermetically sealed from one another but rather are permeable and permit boundary crossing" (GOERTZ; MAHONEY, 2012, p. 2).

In another perspective, Burrell and Morgan (1979) argue that the methodological dimensions present in between these two poles form a continuum within the contexts of Sociology as radical change and as regulation, enabling the observance of four paradigms in social theory analysis: radical humanism, interpretivism, radical structuralism, and functionalism. In this view, the limitations present in each pole indicate the possibilities of complementary among the two, seeking better results in the studies being developed.

Highlighting the globalization process and the technological changes as factors in the methodological transformations in contemporary political science, Rezende (2015) highlights the concern with the integration between the studies of quantitative and qualitative nature. He sees in the connection between these different research methods an opportunity to develop a methodological awareness that enables the promotion of relevant knowledge generated by the inherent qualities of each type.

Some studies, given their specificities, indicate the need for a multifaceted view on the phenomenon under analysis, requiring the conduction of research

that takes on both quantitative and qualitative aspects. Thus, it is understood that the value given to this dichotomy harms scientific development, in that it restricts possibilities when seeking recognition and value in the border areas between these two streams (PECI, 2004).

The option for an interpretivist approach of the social phenomenon, for example, awakens the search for understanding a network of complex and subjective representations, in an emergent process, an extension of the experience lived by flesh and blood people in the "real" world (VERGARA; CALDAS, 2005), aiming to comprehend the meaning experienced by subjects (GEPHART, 2004). Therefore, under the perspective of ontological relativism, reality is interpreted as socially constructed, such that individuals and worlds are mutually constructed (CLARKE, 1991; STRAUSS, 1993) from lived experiences (GUBA; LINCOLN, 1994).

It also acknowledged that the historical, social, political, and economic context has characteristics that reinforce particularities of the phenomenon being studied, demanding a multidimensional view by the researcher. Thus, it is important to highlight the methodological contributions from different scientific areas, especially social anthropology since its understanding of methodological foundations retrieve the importance of conducting qualitative studies to comprehend reality. This is sometimes through demonstrating a concern with understanding the individual (MALINOWSKI, 1992), and other times seeking a systemic understanding (RADCLIFF-BROWN, 1958) as well as seeking to integrate these traditions, by constructing a theory of action, in which historical aspects and social artefacts are considered, combined with observation and inquiry of those involved to reconstruct social processes, through comprehending the sequence of experiences lived by the subjects (FELDMAN-BIANCO, 1987).

In this view, the relevance of comprehending the context is highlighted, especially in the cultural sense, as shaping of human behavior (GEERTZ, 2008). Marcus (2001) also points out the importance of considering existing connections between individuals, accentuated by the globalizing process, and which demands the development of an ethnography that considers a multilocal approach. The contextual relationships arising from these interactions by individuals, their equivalents, nature and artefacts enable the understanding of actions developed in the social environment in a multiple form. In this context, it is necessary to comprehend the possibility of creating knowledge from interpreting the meanings given by subjects to their experiences through their social interactions (CHARMAZ, 2014).

The development of field research in Brazilian political science has significantly increased in the last few years. However, given what has been discussed here, the importance of expanding studies based on qualitative methods in political science is noted, especially as a way of generating hypotheses from field data. One of the possibilities to promote this expansion resides in adopting the grounded theory approach which "has been used successfully in many circumstances, disciplines and cultures" (McNABB, 2015, p. 251). With this focus, this paper proposes to discuss possibilities to apply the grounded theory approach in political science, from its characterization, description of steps, presentation of application possibilities, and indication of potential topics for research that may benefit from its inclusion the list of methodological alternatives.

2. Emergence and evolution of grounded theory

Grounded theory appears as an alternative research approach from the need to conduct a study on the

death and dying of patients with severe comorbidities (CHARMAZ, 2009) in the mid-60s. The release of the book "*The Discovery of grounded theory*" in 1967, by sociologists Barney Glaser and Anselm Strauss is the event that represents the beginning of applying research strategies that enable the construction of theory based on data, in opposition to the current strategy of deduction of hypotheses from pre-conceived theories (McNABB, 2015). The grounded theory approach includes foundations that coincide with symbolic interactionism (BLUMER, 1969), seeking to understand collective actions and interactions between individuals and society "due to the meaning attributed to collective social definitions, formed through the socialization process" (BANDEIRA-DE-MELO; CUNHA, 2010, p. 246).[3]

Curiously, grounded theory arises at a critical stage in sociological studies and research, in which the application of qualitative methods and techniques was found to be in decline in comparison with the increased investment into quantitative research resources (CHARMAZ, 2009). The context indicated weakening in the qualitative research tradition in Sociology and the appearance of a new qualitative approach enabled renewal to those researchers who were dedicated towards studies of this kind and contributed to scientific development in the field.

This approach's concern lies in explaining the reasons for the diversity and variety of behaviors over time and in relation to a given phenomenon (MERRIAM; TISDELL, 2016). In this view, grounded theory is predominantly characterized as qualitative and descriptive, given that the central concern is the interpretation of the data obtained from the various actors involved, in addition to the observation and analysis of documents related to the object of study. Therefore, it differs completely from the hypothetical-deductive method, since this begins with a hypothesis that will or will not be confirmed to generate understanding on the explanatory theory of objective reality. Meanwhile, grounded theory involves theoretical construction itself to explain a subjective reality, from the researcher's and the research participants' interpretation.

Despite having jointly formulated the grounded theory approach, Glaser and Strauss divided it into two streams of thought. Glaser (2002) defends total separation between description and theorization, which would enable a clear distinction from other methods of a qualitative nature, such as phenomenology and ethnography, which present a perspective with a descriptive emphasis, aiming to "discover cover causal explanations and to make predictions about an external, knowable world" (CHARMAZ, 2006, pp. 4-5). Therefore, Glaser is closer to the positivistic approach, by believing that an objective reality will emerge from the data, so long as the neutrality in the researcher's participation in the process of theoretical construction is respected (CHARMAZ, 2014).

On the other hand, Strauss presents the theory founded on a subjectivist/interpretivist view, from the development of data collection and analysis procedures based on the researcher's interpretation regarding the results achieved. In this tradition, followed by other researchers such as Juliet Corbin, the aim with the grounded theory approach is the construction of substantive theories from the researcher's interpretation on the multiplicity of perceived realities and their contextualization (STRAUSS; CORBIN, 1994; STRAUSS; CORBIN, 2008; CHARMAZ, 2014).

The grounded theory approach focuses on the process of generating theory from particular contents (PATTON, 2002; MERRIAM; TISDELL, 2016). There are divergences between the positivist and interpretivist paradigms in understanding what constitutes a theory.

3 In the original: "em função do significado atribuído às definições sociais coletivas, formadas por meio do processo de socialização".

For Charmaz (2014, p. 229) while for positivists theory "seeks causes, looks for explanations, and emphasizes generality and universality", interpretivists assert that theory "aim[s] to understand meanings and actions and how people construct them" (p. 231), presupposing the existence of multiple and emergent realities and exposing the truth as something temporary and social life as a process. Under this perspective, theories originated according to grounded theory have an empirical nature, based on data representative of reality (GOULDING, 2002; STRAUSS; CORBIN, 2008; CHARMAZ, 2014).

In addition to the classic tradition constructed by Glaser and Strauss (1967), some authors also highlight the existence of two more popular interpretivist perspectives in grounded theory studies: the systematic procedures tradition, developed by Strauss and Corbin (1994), and the social constructivist tradition, developed by Charmaz (2014), among others (CRESWELL, 2007; ESLAVA, 2014; KENNY; FOURIE, 2015). To better understand the three traditions developed, some authors (ESLAVA, 2014; KENNY; FOURIE, 2015) built boxes to present the main differences between them, inspiring the construction of box 1. In it, are highlighted as elements to differentiate the traditions studied the philosophical basis, when to use the literature, the coding system, coding and analysis methods, problem selection and identification, degree of operationalization, nature of emergent theory, and the elements of evidence of the conceived theory, as observed below:

Box 1 – Differences among grounded theory traditions

	Classic (Glaser)	Systematic procedures (Strauss and Corbin)	Social constructivist (Charmaz)
Philosophical basis	Moderate positivism	Post-positivism and symbolic interactionism	Constructivism and symbolic interactionism
Use of literature	Only at the end	At all stages	Prior (sensitizing concepts), at all stages, and compiled at the end
Coding system	Original, to uncover theory	Rigorous, to create theory	Open-ended, to construct theory
Coding and analysis method	General picture based on constant comparison: categories emerge and cannot be forced by methodology	Open, axial, and selective coding: analytical method with structured stages; procedures and techniques need to be precisely detailed	Initial, focused, axial, and theoretical coding; analytical method with semi-structured steps; techniques and procedures can be adapted
Problem selection and identification	Problem emerges from the study and the joint perceptions of the actors and researchers	Determined by the researcher before fieldwork	
Degrees of operationalization	Low due to difficulties given the broad frame of reference	High due to the specificity of the analytical stages	
Nature of the emergent theory	Generates concepts and their relationships to explain and interpret the variation in behavior in the area substantively studied	Generates a theory inductively derived about a phenomenon given interrelated concepts	Creation of substantive theories through construction of inductive analysis from specific data, shared experience, and relationship with participants
Evidence of theory	Subsequent quantitative analyses	Provisional test to validate concepts and their interrelationships	Conceptual schemes validated by research participants

Source: adapted from Eslava (2014) and Kenny and Fourie (2015)

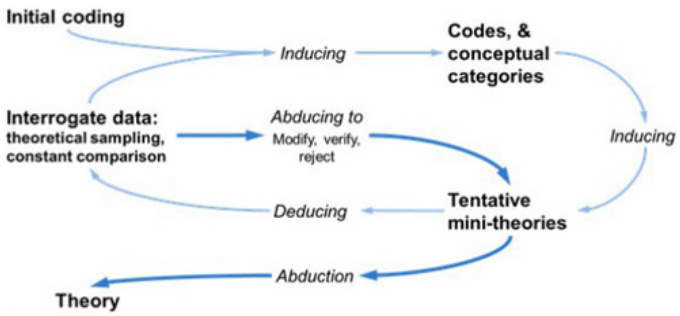
Beyond identifying differences between the three traditions, it is important to identify their points of convergence. Here, Kenny and Fourie (2015) highlight as unifying principles of grounded theory the analytical method of constant data comparison, the theoretical sampling aiming to identify the saturation point, writing memos throughout the analysis process, and the construction of substantive theory to close the process.

Taking as a starting this brief trajectory of grounded theory traditions as research approaches, the goal is to present the required stages for the application in scientific studies and research. To do so, it was opted to consider the social constructivist perspective as an alternative to be explored by political science, especially because it represents the most dissonant modality in relation to the positivist model.

A common characteristic of this method is the intense interaction between the researcher and the realities experienced by subjects, as well as the simultaneous involvement of data collection and analysis. Another point to highlight is the importance of the researcher in this approach who, by adopting the social constructivist tradition, needs to stay vigilant regarding the conditions under which differences and distinctions between people occur, to identify how they come about and are sustained, from an analysis that is "contextually situated in time, place, culture, and situation" (CHARMAZ, 2006, p.131). The concern with the context is clear, especially regarding the comprehension of contributions to the construction of the observed reality.

Another characteristic of the social constructivist perspective is the adoption of abductive logic in the development of ideas, where reasoning begins from data, and from their analysis, codes and categories are formulated in an iterative process, until a plausible interpretation about the results observed is reached (CHARMAZ, 2014). Figure 1, elaborated from the case developed by Ward, Gott, and Hoare (2017), represents this iterative relationship in the researchers reasoning process between the inductive, deductive, and abductive stages in the theoretical construction according to grounded theory.

Figure 1 – reasoning in grounded theory research



Source: Ward, Gott, and Hoare (2017)

Considering what is presented in the figure above, the importance of recognising the context as an essential element in theoretical construction is highlighted. It becomes essential to understand actors in the performance of their activities, the interest lying on how they act, awakening the peculiar curiosity to the origins of scientific finds.

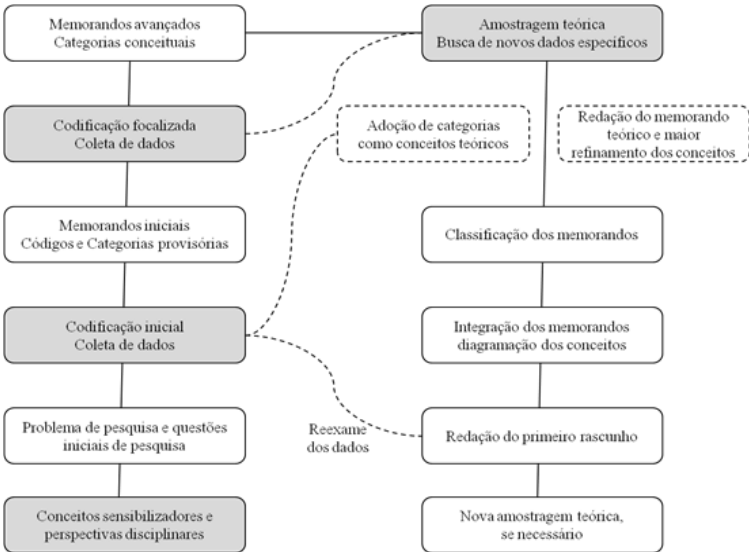
According to Charmaz (2014, p.239), the constructivist approach "places priority on the studied phenomenon and sees both data and analysis as created from shared experiences and relationships with participants and other sources of data". From this idea follows the proposition of applying the constant comparison method with results obtained in the field. The

interpretations of observable realities emerge from a process based on successive interactions between data collection and analysis, starting from comparisons, codings, and categorizations of the studied phenomena, enabled by behaviors, actions, and words extracted from actors researched (STRAUSS; CORBIN, 1994; GOULDING, 2002; STRAUSS; CORBIN, 2008; CHARMAZ, 2014). The next section is dedicated to explaining the necessary stages to the grounded theory process according to the social constructivist tradition.

3. The grounded theory process

This section of the paper will present the main stages to elaborate research with the grounded theory approach. From the start, it is important to highlight that, given the perspectives presented previously, for the construction of this paper the social constructivists perspective was considered, developed by Kathy Charmaz (2006, 2014). Figure 2 explains the grounded theory process according to the perspective adopted in this paper, indicating the sequence of steps to be developed during research.

Figure 2. The grounded theory process



Source: adapted from Charmaz (2009, p. 26)

The distinction between the perspectives is notable in the first stage in the work being developed, since in the chosen perspective It is necessary to choose literature to construct sensitizing concepts, something which is not present in the models proposed by Glaser and by Strauss.

3.1. Sensitizing concepts

The initiative to develop sensitizing concepts is originated in Blumer's (1969) symbolic interactionism by considering that researchers "begin their studies with certain research interests and a set of general concepts " (CHARMAZ, 2006, p. 16). Considering the characteristics interpretivist research, Blumer (1969) proposes that the logic of social research is altered and, instead of aiming for finding definitive concepts, there is a break with this dictum, establishing what he names "sensitizing concepts". In contrast with definitive concepts, a sensitizing concept is a broad term, without definitive characteristics, functioning as a spark so that aspects of a question are reflected, serving as a general frame of reference and guidance in the empirical approach. In this way they enable the researcher an initial attempt to pursue and question the appearance of codes and cat-

egories, seeking the development of ideas on the processes that the data will define (CHARMAZ, 2014).

Sensitizing concepts represent a preliminary direction of what is expected to be observed in the field, thus becoming ephemeral and vulnerable to confrontation with the peculiarities of the context studied and revision with data gathered in the field (BLUMER, 1969). Supported by this conception, sensitizing concepts are characterized as provisional tools that will drive the researcher to confront these concepts with the data gathered during collection, to begin to construct the study's initial codes. They are constructed from disciplinary perspectives adopted by the researchers and, therefore, work as triggers, as starting points for the study's development, originating from prior literature review related to the main terms that guide the investigator's research interests.

Currently, with available technological resources, it is possible to conduct this review at national and international research databases, as a way to broaden the initial scope of the study being developed. However, sensitizing concepts do not necessarily need to come from formal references. Although they do need to represent the researcher's view on the aspect being analyzed, since they will contribute to the construction of the research problem and initial research questions, fundamental elements for the conception of the research project being elaborated.

Nevertheless, it is important to highlight that an important contribution from grounded theory resides precisely in the support coming from data collected for the formulation of new hypotheses that will confront or reinforce existing knowledge. In this way, sensitizing concepts need to be confronted with field data from the researcher's initial contact, given the intense interaction aspect between researcher and lived realities of subjects, in addition to the simultaneous involvement in data collection and analysis.

3.2. Initial sample definition

The circularity between the data collection and analysis stages indicates that these processes occur in a simultaneous and iterative way, until theoretical saturation is reached, that is, when no new properties or theoretical insights can be drawn from the collection of more data (CHARMAZ, 2006; BANDEIRA-DE-MELO; CUNHA, 2010). However, to enable data collection to begin, the selection of the initial sample is necessary, and it should be "aimed toward theory construction, not for population representativeness" (CHARMAZ, 2014, p.8). In this way, theoretical sampling is characterized as non-probabilistic, in distinction to the traditional perspective of statistical or probabilistic sampling.

For the sample selection to occur, a preliminary stage is required, which is defining the object of research, which necessarily needs to be a "social phenomenon that one seeks to deepen theoretical analysis on" (LAPERRIÈRE, 2010, p. 357). The chosen social phenomenon needs to reflect a process and one should explore its evolution in a context or bounded situation. When the research involves a case study, sample selection requires two stages, one being the selection of the case itself and the second, the definition of the individuals to research (MERRIAM; TISDELL, 2016).

The choice in case to be studied depends on multiple factors, with a case study "be selected because it is an instance of some process, issue, or concern" (MERRIAM, 2009, p. 41). Studies under the grounded theory mantle involve intentionality, established by criteria that enable the identification of participants that have experiences with the studied phenomenon and may contribute to the theoretical elaboration; this can be started with a homogeneous sample of individuals and later on encompass heterogeneous groups (CRESWELL, 2007; STRAUSS; CORBIN, 2008).

It is crucial to make an important distinction between initial sample and theoretical sampling, a component of the grounded theory process. The initial sample, comprised by the non-probabilistic definition of case(s) and subject(s) to participate in the research, represents only the starting point from which data will be obtained. Theoretical sampling represents refining the data into codes, developing the categories' properties that enable explaining the phenomenon studied. Thus, "[w]hen engaging in theoretical sampling, the researcher seeks people, events or information to illuminate and define the boundaries and relevance of the categories" gathering samples for the construction of theoretical categories (CHARMAZ, 2006, 189).

3.3. Data collection and analysis

With the initial sample defined, the investigator can begin the data collection stage, establishing the most appropriate techniques for the defined sample. Although Charmaz (2006) emphasizes the possibility of collecting data through ethnographic observation processes, with in-depth interviews and text analysis of extracted or existent texts, it is evident that other possibilities of data collection can be adopted in the grounded theory process. The possibility of conducting semi-structured interviews for research based on grounded theory within political science has indeed been pointed out by Rocha (2020). Adopting different data sources enriches the study, enabling comparisons, confrontations, and questioning that leads to greater refinement of the categories formulated. Regardless of the data collection technique adopted, it is important that "how participants invoke ideas, practices, and accounts from both the larger and local cultures of which they are a part [of]" is considered, because this will contribute to understanding the social context, making the coding and categorization process easier (CHARMAZ, 2006, p. 40).

Along with the data collection process begins the data analysis process, which are, in the context of grounded theory, indissociable stages (BANDEIRA-DE-MELO; CUNHA, 2010). According to Charmaz (2006, p. 43), data analysis begins with the coding process, which "means categorizing segments of data with a short name that simultaneously summarizes and accounts for each piece of data". Thus, the central procedure of the analytical process of research conducted according to grounded theory is systematic comparative thinking across data obtained throughout each step of collection and the codes and categories constructed in this process, aiming for the identification of convergences and divergences that contribute to theoretical saturation.

Coding, described by Charmaz (2014) as the process of selection, separation, classification, and denomination that summarizes and represents the data, aims for the construction of the study's analytical structure, working as an intermediary phase between the data and the theory. Coding is conducted in stages, with the first one named initial coding, corresponding to the phase when naming the data obtained happens, in a way that an initial definition and classification is developed in the shape of a synthesis. The initial coding needs to be centered around the data, allowing preliminary codes generated to be temporary, but enabling comparisons from the available data. This stage is understood as the one that "generates the bones" of the analysis, that is, that allow the occurrence of the theoretical integration process, whose role is to "assemble these bones into a working skeleton" (CHARMAZ, 2006, p. 45).

The initial coding can be developed in different ways, such as word-by-word coding, which allows the observation of the structure and flow of words; line-by-line, which enables the identification of codes emerging from words expressed in that process; and also, incident-to-incident coding, in which incidents observed in the interviewee's speech are compared (CHAR-

MAZ, 2006). The possibility exists of adopting computational programmes to support the coding process, given that they have "great capacity for organizing massive amounts of data, facilitating analysis, and assisting communication among the members of a research team" (MERRIAM; TISDELL, 2016, p. 221). During this stage, the emergence of *in vivo* codes can also be observed, which are widely used terms spoken by interviewees and that can represent the meaning given by them to their opinions and attitudes. The flexibility of this coding type enables capturing these language aspects, without these *in vivo* codes getting lost in the researcher's interpretation.

The conduction of the initial coding enables adjusting the study to the empirical world in the codes' construction process and, later, in the development of categories. It also enables the perception of the relevance of the study being developed, given that it serves as a "skeleton" for the researcher's interpretation regarding the phenomenon studied, allowing others to understand the way that it "makes relationships between implicit processes and structures visible" (CHARMAZ, 2006, p. 54). The way through which this becomes possible is the elaboration of initial memos, which are records of the researcher's thoughts on the data, as a way to generate new ideas, crystallize questions and directions assumed throughout the study, through comparison between the data, codes, categories, and sensitizing concepts (CHARMAZ, 2014).

A second stage in the analysis is named focused coding, which "means using the most significant and/or frequent earlier codes to sift through large amounts of data" (CHARMAZ, 2006, p. 56), with the aim of generating and validating provisional properties and categories by constant comparisons with new empirical data. The codes originally found are grouped, as a way of enabling the discovery of possible associations, disconnected or not perceived in the prior stage, so that these can allow understanding of the studied phenomenon from the emergence of new ideas.

The refinement of provisional data, codes, and categories is done through the elaboration of advanced memos, where relationships between categories and subcategories are described, in a regrouping process of the codes generated previously. This process is named by Strauss and Corbin (2008) as axial coding and seeks to explore questions regarding when, where, why, who, and how, as a way of generating more precise explanations on the phenomenon and providing coherence to the analysis (CEPELLOS; TONELLI, 2020). The systematic process of constant comparison of the data generates confrontation between the propositions created, verifying the consistency of the analysis conducted, giving greater robustness to the categories established and their properties (CHARMAZ, 2014).

As mentioned previously, this category refinement process with data saturation is named theoretical sampling. The goal of theoretical sampling is to elaborate and refine categories that make up the theoretical proposition, using new data, which can be obtained through follow-up interviews to fill possible gaps in the formulation of categories, as well as enable its analytical development, reinforcing its properties until these are consolidated with available data (CHARMAZ, 2006).

Theoretical sampling follows an iterative process of collection and analysis that ends with the classification and integration of memos, the formulation of definitive categories, and the proposition of a diagram elaborated using a process of abductive inference, for which thinking about the experience enables the elaboration of a substantive theory, developed for the immediate field of study (GOULDING, 2002; CHARMAZ, 2006). This process is developed until no new significant contributions from the data obtained that aid in the explanatory capacity of the formulation elaborated are presented.

3.4. Validity and reliability of the research

In studies of a qualitative nature, an issue that demands additional effort by the researcher is the definition of the study's validity and reliability criteria. This concern transcends the structure of stages representative of the grounded theory process constructed under the social constructive view that was previously presented.

Before anything, it is important to highlight the inherent ethical aspect to this kind of research. The adoption of an ethical conduct is essential, given that it is what allows the construction of the researcher's credibility to be built. Merriam and Tisdell (2016) highlight that ethical behavior needs to be present in the main forms of data collection in qualitative research. In interviews, individuals may feel that their privacy is being invaded, maybe embarrassed by certain questions posed, or yet may say things that they are not allowed to reveal. Likewise, observation procedures may pose traps such as the researcher's excessive involvement on issues related to the research field, such as having their judgement influenced by the situation observed. Similar risk is also present in data made available through documents, given that these can introduce information that is confidential to the organization involved.

This stage passed, it must be understood that qualitative studies are different from quantitative studies regarding validity and reliability criteria. This logic is clear when is considered the way reality is interpreted. Thus, studies that take on a constructive view assume that reality is multiple, originated from social constructions based on the experiences lived by individuals. It is sought to interpret the meanings constructed by subjects so that a formulation of the researcher's own conception on the object of investigation is possible. Given this peculiarity, criteria that ensure quality in qualitative research are presented, especially highlighting those that enable the recognition of validity and reliability.

Specifically considering the grounded theory approach, Strauss and Corbin (2008) recommend some criteria for verifying the quality of the research, presented as questionings. for these authors, it is necessary to have rigour in the coding and research process, quality in the elaboration of concepts, systematic relationships between the concepts formulated, conceptual density, variation and specificity, significance of the finds, and theoretical sensitivity.

More broadly, Creswell (2010) points out that questions on the validity of research permeate all phases of research work, given that they represent a researcher who adopts procedures that verify the precision of results. Merriam and Tisdell (2016) defend that the validity of qualitative research takes on internal and external characteristics. Internal validity (or credibility) is related to the congruence of data obtained with the reality researched. External validity (or transferability) corresponds to the condition where results reached in the study can be reproduced in other contexts or situations, delimiting the scope of application of the findings obtained in the study.

Merriam and Tisdell (2016) highlight as main possibilities of internal validity: triangulation, respondent validation, adequate engagement by the researcher in data collection, researcher reflexivity, peer review, audit trail, rich, thick descriptions of the phenomenon, and maximum variation in the application of the findings. Cho and Lee (2014) point out that, in studies on grounded theory approach, in addition to the possibilities presented, there is also the presentation of dissonant or negative information identify throughout the study.

When it comes to reliability (or consistency), Merriam and Tisdell (2016) state that this refers to the possibil-

ity of getting the same (or similar) findings by replicating the research, under similar circumstances. However, the authors argue that reliability does not apply to qualitative research, given that the conception of multiple socially constructed realities it is not conducive with the possibility of their replication. "[Qualitative] researchers seek to describe and explain the world as those in the world experience it" (MERRIAM; TISDELL, 2016, p. 250). Consequently, the consistency and security of the study is evaluated, associating these with issues of internal validity, hoping that other researchers can, from the same collected data, come to similar understandings.

4. Possibilities for the application of grounded theory in Political Science

After presenting the grounded theory approach trajectory as an alternative for qualitative research and the steps in the methodological process linked to the social constructivists tradition, the challenge is verifying the possibilities for its application in the political science field. So that it was possible to identify, albeit in a preliminary perspective, the scientific production using grounded theory in the field of political science, a review of published papers recorded in CAPES Journal Repository[4] was conducted. To do so, search terms considered were "grounded theory" and "political science". The result of this search indicated a total of 62 articles, with only 23 in Portuguese.

However, when reading the selected articles in more detail, it was verified that in fact, only three articles had the desired profile. The remaining papers despite having the chosen search terms, used some reference that had the terms, were predominantly linked to other areas, or were originally written in another language. Thus, the following papers should be registered: "Electoral defeats and political careers: the case of the Rio Grande do Sul Workers' Party (PT)", by Frédéric Louault (2011), published in the Revista de Sociologia e Política; "Casting a light on bureaucratic determinants and management of social welfare policy in Brazilian municipalities", by Natália Sátyro and Eleonora Cunha (2019), published in Opinião Pública; and "Demand as a determinant factor in the execution of social assistance services in Brazilian municipalities", by Raquel D'Albuquerque and Natália Sátyro (2020), also published in Opinião Pública.

The reality explored by the data in this preliminary search Campos at least one question: are field researchers in political science really not interested in developing studies based on grounded theory or do they not recognise it as a valuable alternative to conduct qualitative or mixed research? So that it is possible to answer this question it is necessary to better understand what the data gathered in the search want to represent.

To do so, the scope of the previous search was altered, replacing only the term "grounded theory" for the term "qualitative". The number of articles selected rises to 1,201, with 625 written in Portuguese, leading to the understanding that, given the volume of qualitative works identified, the academic investment in grounded theory in political science in Brazil is still small, possibly due to an absence of recognising it as a viable qualitative alternative. Would this be another piece of evidence of deficiencies in the methodological formation of political scientists, pointed out by Soares (2005)? Indeed, the work developed by Barberia, Godoy, and Barboza (2014) regarding the evolution of research methods and techniques teaching in Brazilian political science ratifies the need for formation in formal model analysis and qualitative research techniques.

4 Coordination for the Improvement of Higher Education Personnel.

Additionally, it is also noted that critiques were formulated on the application of grounded theory in political science. When synthesizing them, Becker (2012) questions above all the classic tradition for not making possible the the use of previous references, which is overcome by remaining traditions. It is also questioned the limitation of its application on studies in international relations, whose social unit are a nation-state or region of the world, which is confronted by Glaser and Strauss when defending this possibility with the grounded theory approach. McNabb (2015), when reporting the study developed by Cook and Barry (1995), indicates as a disadvantage of grounded theory an overabundance of raw data generated by research, something which currently can be overcome with the adoption of computational programmes to support the coding process, as described in the previous section.

Despite these critiques being founded, these same researchers identify potential in the application of grounded theory in political science. In addition to the research by Cook and Barry (1995), on the interactions of public policies between the public and private sectors, McNabb (2015) dedicates a chapter of his book "Research methods for political science: quantitative and qualitative approaches" to discuss grounded theory's potential, using as reference the example of the work done by King, Feltey, and Susel (1998) on the underlying causes of public antipathy in the political process and ways of improving participation in the policy-making decisions of public administration.

More broadly, Becker (2012) sees a larger potential in the application of grounded theory in political sociology and public policy studies, as well as studying political culture, such as electoral behavior or responses to political campaigns. Furthermore, she recognizes that grounded theory could fulfil some needs in the field of political science, especially in theory generation processes, as a way of coming closer to the phenomena that it seeks to explain. Besana (2018) also indicates the possibility of applying grounded theory in studies on political participation at a micropolitical level, which could be difficult in reaching by adopting other analytical perspectives.

Another field where an elevated potential for exploring qualitative methods is feasible is the one of public policy evaluation. Regarding this, Rossi, Lipsey, and Henry (2019, p. 304) understand that social researchers that choosing this perspective are guided "formative evaluation, that is, making a program work better by feeding information to its managers and sponsors. In addition, they tend to rely on information about the lived experiences of those being served by the program". Under this perspective, studies conducted by adopting grounded theory could contribute to the evaluation process of public policies, both in an *ex ante* analysis, aiming to identify problems and needs experienced by individuals and thus contribute to policy-making, as well as in *ex post* analysis, to verify if policy effects effectively contributed in solving problems identified, based on the lived experience of policy beneficiaries.

Therefore, it is observed that despite there being critiques to the use of qualitative methods in political science research and, specifically, directed at grounded theory, these have been continually combatted by the increase in their application in the field, as well as the incorporation of new computational tools to support research that incorporate techniques associated with the approach. Beyond the tradition in the development of qualitative research, it is also necessary that there is increased investment in the formation of researchers on qualitative and mixed methods. It is only with this investment a stimulus to scientific production from works developed with qualitative (including grounded theory) or mixed methods that we can expect overcoming gaps that still remain in the political science field, due to not taking advantage of the application possibilities that were presented in this section.

5. Final remarks

The main goal in constructing this paper was to encourage qualitative research in political science, especially through the use of the social constructivist approach of grounded theory. The concern with this gap was present in works by Giovanni Sartori and other political scientists, who emphasize the value of quantitative research in this research field.

For this debate to be possible, the origin and evolution of grounded theory as a methodological approach was presented. It was demonstrated that its roots are planted in the social sciences, through the work of sociologists Barney Glaser and Anselm Strauss, under the influence of Herbert Blumer's symbolic interactionism. The break between its creators resulted in the creation of three main traditions: the classic stream, followed by Glaser, the systematic procedures stream, developed by Strauss and Corbin (1994), and the social constructivist chain, proposed by Kathy Charmaz. Taking into account that the latter is the one that most coalesces with political science research, in order to better answer criticism from researchers in this field, it was chosen as reference for this paper.

Thus, the main stages of the grounded theory process according to Charmaz (2006) were presented. It was highlighted the importance of constructing sensitising as a starting point for the development of research, as well as criteria of creating the initial sample. The indissociable nature between the data collection and analysis process was also discussed, emphasizing the iterative nature between these stages until theoretical saturation is reached and the substantive theory is elaborated. In addition to these steps, the validity and reliability of research conducted using the grounded theory approach was debated, underscoring the importance of application of techniques to reinforce these quality criteria of scientific production.

Lastly, a review was presented that indicated the limited production of knowledge using the grounded theory approach in Brazilian political science. Although there are many criticisms of this methodological approach, it is observed that they are beginning to be overcome, whether by the evolution of techniques and tools, or by the adoption itself in research in this field. Therefore, some possibilities for the application of grounded theory were shown, especially in the areas of political culture, political institutions, and public policies.

Consequently, a contribution for researchers linked to political science is hoped for, so that they can develop qualitative and/or mixed studies based on the grounded theory approach. Likewise, a significant effort is needed to incorporate this perspective in the formation of new researchers in the field. This is an important challenge for academia, so that potential new knowledge that can be generated through initiatives that converge with this possibility is used.

BIBLIOGRAPHICAL REFERENCES

Bandeira-de-Mello, R.; Cunha, C. (2010), "Grounded theory", in C. Godoi; R. Bandeira-de-Mello; A. Silva, A. (org.). *Pesquisa qualitativa em estudos organizacionais: paradigmas, estratégias e métodos*. São Paulo, Saraiva.

Barberia, L.; Godoy, S.; Barboza, D. (2014), "Novas Perspectivas sobre o 'Calcanhar Metodológico': O Ensino de Métodos de Pesquisa em ciência política no Brasil", *Teoria & Sociedade*, vol. 22, no. 2: 156-186.

Becker, Brittney. (2012), "The Grounded Theory Method and its uses for Political Science", in B. Becker; G. Kaufmann (org.), Series of Paper: Methods of Field Research.

Besana, P. (2018), "Notas sobre el uso de la etnografía y la teoría fundamentada en ciencia política. Un análisis amplio de la participación política y el Estado en asentamientos informales de la periferia de Buenos Aires, Argentina", *Universitas humanística*, vol. 86, nº. 2:107-137.

Blumer, H. (1969), *Symbolic interactionism: perspective and method*. Englewood Cliffs, Prentice-Hall.

Burrell, G.; Morgan, G. (1979), *Sociological paradigms and organisational analysis.* Aldershot, Ashgate.

Cansino, C. (2007), "Adiós a la ciencia política: crónica de uma morte anunciada", *Temas y debates*, vol. 14, nº 2:13-29.

Cepellos, V.; Tonelli, M. (2020), "Grounded theory: passo a passo e questões metodológicas na prática", *Revista de Administração Mackenzie*, vol. 21, nº 5:1–29.

Charmaz, K. (2006), *Constructing grounded theory: A practical guide through qualitative analysis.* London, Sage.

Charmaz, K. (2014), *Constructing grounded theory*. London, Sage.

Cho, J.; Lee, E. (2014), "Reducing confusion about grounded theory and qualitative content analysis: similarities and differences", *The Qualitative Report*, v. 19, nº 32:1-21.

Clarke, A. (1991), "Social words/arenas theory as organizational theory", in D. Maines (org.) *Social organization and social process: essays in honor of Anselm Strauss*. New York, Aldine de Gruyter.

Creswell, J. W. (2007), *Qualitative inquiry and research design*. Thousand Oaks, Sage.

Creswell, J. W. (2010), *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. Porto Alegre, Artmed.

D'albuquerque, R.; Sátyro, N. (2020), "A demanda como determinante na execução dos serviços socioassistenciais nos municípios brasileiros", *Opinião Pública*, vol. 26, nº 2:246-282.

Eslava, A. (2014), "Análisis cualitativo y cuantitativo para los estudios políticos: trilateralismo, metodología experimental y grounded theory", *Cinta de moebio*, vol. 3, nº. 51:111-126.

Feldman-Bianco, B. (1987), *A Antropologia das sociedades contemporâneas*, Global, São Paulo.

Geertz, C. (2008), *A interpretação das culturas*. Rio de Janeiro, LTC.

Gephart, R. (2004), "From the editors: qualitative research and the academy of management journal". *Academy of Management Journal*, v. 47, nº. 4:454-462.

Glaser, B. (2002), "Constructivist grounded theory?", *Forum: qualitative social research*, vol. 3, nº. 3:1-14.

Glaser, B.; Strauss, A. (1967), *The discovery of grounded theory: strategies for qualitative research*. New Brunswick, Aldine.

Goertz, G.; Mahoney, J. (2012), *A tale of two cultures: qualitative and quantitative research in the social sciences*. Oxfordshire, Princeton University Press.

Goulding, C. (2002), *Grounded theory: a practical guide for management, business and market researchers*. Thousand Oaks, Sage.

Guba, E.; Lincoln, Y. (1994), "Competing paradigms in qualitative research", in N. Denzin; Y. Lincoln, (ed.). *Handbook of qualitative research*. Thousand Oaks, Sage.

Kenny, M.; Fourie, R. (2015), "Contrasting classic, straussian, and constructivist grounded theory: methodological and philosophical conflicts", *The qualitative report*, vol. 20, nº 8:1270-1289.

Laperrière, A. (2010), "A teorização enraizada (grounded theory): procedimento analítico e comparação com outras abordagens similares", in J. Poupart; J. Deslauriers; L. Groulx; A. Laperrière; R. Mayer; A. Pires. *A pesquisa qualitativa: enfoques epistemológicos e metodológicos*. Petrópolis, Vozes.

Louault, F. (2011), "Derrotas eleitorais e carreiras políticas: o caso do PT no Rio Grande do Sul", *Revista de Sociologia e Política*, vol. 19, nº 40:233-253.

Malinowski, B. (1992), *Los argonautas del Pacífico Occidental*, Península, Barcelona.

Marcus, G. (2001), Etnografía en/del sistema mundo: el surgimiento de la etnografía multilocal, *Alteridades*, vol. 11, nº 22: 111-127.

McNabb, D. (2015), *Research methods for political science: quantitative and qualitative approaches*. New York, Routledge.

Merriam, S. (2009), *Qualitative research: a guide to design and implementation*. San Francisco, Jossey-Bass.

Merriam, S.; Tisdell, E. (2016), *Qualitative research: a guide to design and implementation.* San Francisco, Jossey-Bass.

Peci, A. (2004), "Além da dicotomia objetividade-subjetividade", in M. Vieira; D. Zouain (org.), *Pesquisa qualitativa em Administração*. Rio de Janeiro, FGV.

Patton, M. (2002), *Qualitative research and evaluation methods*. Thousand Oaks, Sage.

Radcliff-Brown, A. (1958), *El método de la Antropología Social*, Editorial Anagrama, Barcelona.

Rezende, F. (2015), "Transformações metodológicas na ciência política contemporânea", *Política Hoje*, vol. 1, nº. 24:13-45.

Rocha, V. (2020), "Da teoria à análise: uma introdução ao uso de entrevistas individuais semiestruturadas na ciência política", *Política Hoje*, vol. 1, nº. 29:1-26.

Rossi, P.; Lipsey, M.; Henry, G. (2019), *Evaluation: a systematic approach*. Thousand Oaks, Sage.

Sartori, G. (2004), "Where is political science going?", *Political science and politics*, vol. 37, nº. 4: 785-789.

Sátyro, N.; Cunha, E. (2019), "Jogando luzes sobre os determinantes burocráticos e de gestão da política de assistência social nos municípios brasileiros", *Opinião Pública*, vol. 25, nº 2:401-432.

Soares, G. (2005), "O calcanhar metodológico da ciência política no Brasil", *Sociologia, problemas e práticas*, vol. 1, nº. 48:27-52.

Strauss, A. (1993), *Continuous permutations of action*. New York, Aldine de Gruyter.

Strauss, A.; Corbin, J. (1994), "Grounded theory methodology: an overview", in N. Denzin; Y. Lincoln (ed.). *Handbook of qualitative research*. Thousand Oaks, Sage.

Strauss, A.; Corbin, J. (2008), *Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento de teoria fundamentada*. Porto Alegre, Artmed.

Vergara, S.; Caldas, M. (2005), "Paradigma interpretativista: a busca da superação do objetivismo funcionalista nos anos 1980 e 1990", *Revista de Administração de Empresas*, vol. 45, nº. 4:66-72.

Ward, K.; Gott, M.; Hoare, K. (2017), "Analysis in grounded theory: how is it done?", *SAGE Research Methods Cases*. Thousand Oaks, Sage.

DA TEORIA À ANÁLISE: Uma introdução ao uso de entrevistas individuais semiestruturadas na ciência política.

Virgínia Rocha - Universidade Federal de Pernambuco

Resumo

O que são entrevistas semiestruturadas? Como elas podem ser usadas na ciência política? Neste artigo abordo os principais aspectos definidores das entrevistas em profundidade como técnicas qualitativas de coleta de dados focadas na compreensão de emoções, crenças, percepções, motivações e atitudes dos indivíduos. Especificamente, trato da combinação entre estrutura e flexibilidade presente nas entrevistas semiestruturadas e discuto as consequências disso para a validade e a confiabilidade dos dados coletados. Em seguida, apresento um guia prático para a aplicação desse método na ciência política. Detalho as principais etapas do uso de entrevistas semiestruturadas, da teoria à análise dos dados, trazendo exemplos práticos e permeando discussões relevantes na área, como o debate entre as tradições quantitativa e qualitativa e a necessidade de transparência na aplicação de entrevistas e na pesquisa qualitativa de forma geral. A expectativa é oferecer um conteúdo introdutório para estudantes e pesquisadores(as) da área e, assim, contribuir para avançar no treinamento formal e no uso rigoroso de técnicas qualitativas, ambos ainda incipientes no Brasil.

Palavras-chave: entrevistas semiestruturadas; entrevistas individuais; entrevistas em profundidade; métodos qualitativos; coleta de dados.

Abstract

What are semi-structured interviews? How can they be used in political science? In this article, I discuss the main defining aspects of in-depth interviews as qualitative data collection techniques focused on understanding the emotions, beliefs, perceptions, motivations, and attitudes of individuals. Specifically, I deal with the combination of structure and flexibility present in semi-structured interviews and discuss the consequences of this for the validity and reliability of the collected data. Next, I present a practical guide to the application of this method in political science. I discuss the main stages of the use of semi-structured interviews, from theory to data analysis, bringing practical examples and permeating relevant discussions in the area, such as the debate between quantitative and qualitative traditions and the need for transparency in the application of interviews and qualitative research in general. The expectation is to offer introductory content to students and researchers in the area and, thus, contribute to advance in formal training and rigorous use of qualitative techniques, both still incipient in Brazil.

Keywords: semi-structured interviews; individual interviews; in-depth interviews; qualitative methods; data collection.

1. Introdução[1]

"No mínimo, os dados qualitativos são apresentados a título de ilustração; no máximo, dados qualitativos são essenciais para a inferência causal" (GERRING, 2017, p. 15). A importância e os benefícios do uso de dados qualitativos para inferências na ciência política são crescentes (GOERTZ; MAHONEY, 2012; GERRING, 2017; MOSLEY, 2013). Nesse âmbito, as entrevistas em profundidade podem ser aplicadas em diversas áreas da disciplina. Para citar apenas alguns exemplos, Batista (2017) combina análise estatística com entrevistas semiestruturadas com gestores federais e municipais, em uma abordagem multimétodo, para compreender quais são os maiores incentivos e desafios à difusão da Lei de Acesso à Informação nos municípios brasileiros. Lotta et al. (2018), ao investigar a implementação das Supervisões de Assistência Social (SAS) em São Paulo, executam entrevistas semiestruturadas com Supervisoras de três regiões diferentes para melhor compreender o funcionamento das SAS na cidade, obtendo assim maior entendimento sobre esse meio social. Maglia (2020), por sua vez, entrevista as lideranças dos 25 partidos políticos criados no Brasil entre 1990 e 2018, com o objetivo de identificar, na percepção dos respondentes, quais as motivações para a criação de novas legendas partidárias.

Apesar da sua contribuição, o ensino de métodos qualitativos ainda está longe do desejável no Brasil. A situação é análoga para a rigorosidade na qual técnicas qualitativas de coleta e análise de dados têm sido utilizadas (SOARES, 2005; BARBERIA et al., 2014). Essa conjuntura não é específica do Brasil. Mosley (2013) argumenta que a noção de que técnicas de entrevista são métodos menos úteis ou relevantes pode ser uma consequência da ausência de treinamentos mais formais nestes métodos. Apesar de ser amplamente usada em áreas como sociologia, história, antropologia, psicologia, entre outras, cientistas políticos(as) possuem poucas referências metodológicas sobre o uso de entrevistas na disciplina.

Além disso, existem questões epistemológicas associadas ao uso dos métodos qualitativos na ciência política. Especialistas destacam o continuum "positivistas - interpretativistas". Os primeiros, embora reconheçam a complexidade da realidade social, defendem a identificação de relações de causa - efeito, hipóteses falseáveis e empiricamente testáveis. Consideram ainda que os métodos quantitativos e qualitativos são ambos úteis nessa empreitada. Para os interpretativistas, dada a realidade social, seria impossível gerar conhecimento sem considerar o contexto histórico e as relações de poder (MOSLEY, 2013).

As implicações disso para o uso de entrevistas é que interpretativistas trazem grande atenção aos desafios de distanciar o conhecimento e o contexto do(a) entrevistador(a) do processo de coletar e sistematizar os dados oriundos de entrevistas. Já para aqueles(as) em uma posição mais puramente positivista, entrevistas são meios para, objetivamente, testar ou gerar hipóteses falseáveis. Isso não significa que eles(as) desconsideram questões como os efeitos do(a) entrevistador(a) (*interviewer effects*), isto é, como o comportamento ou as características do(a) entrevistador(a) podem afetar as respostas obtidas, mas sim que interpretativistas são mais sensíveis às nuances desse processo do que positivistas. Atualmente, a ciência política é predominantemente positivista, o que afeta os principais desafios enfrentados no uso de entrevistas na área (LYNCH, 2013; MOSLEY, 2013).

Neste artigo, parto de uma perspectiva mais positivista do uso das entrevistas semiestruturadas, mas sem deixar de reconhecer preocupações importantes sobre

1 Agradeço aos(às) pareceristas que avaliaram este trabalho pelos excelentes comentários, aos(às) colegas que sugeriram melhorias em versões anteriores desse estudo, em especial Palloma Marciano e Rebeca Rocha, e à CAPES pelo apoio financeiro via Bolsa de Doutorado. Evidentemente, qualquer equívoco neste trabalho é de minha inteira responsabilidade.

como a relação entre entrevistador(a) e informante pode afetar os dados coletados. Portanto, mobilizo aqui referências interdisciplinares que ajudam a esclarecer a complexidade dessa técnica. Ainda nesta perspectiva positivista, entendo as análises qualitativas e quantitativas como complementares. Nesse sentido, a partir dos argumentos trazidos por King et al. (1994), diversos estudos demonstraram a necessidade de compreender as abordagens qualitativa e quantitativa não a partir do paradigma quantitativo, como fazem os citados autores, mas sim de acordo com as especificidades e objetivos distintos de cada método (BRADY; COLLIER, 2010; GERRING, 2017; GOERTZ; MAHONEY, 2012; MAHONEY, 2010; RAGIN, 2008). Essa perspectiva não implica na ausência de padrões de cientificidade na análise qualitativa, mas sim que eles diferem daqueles estabelecidos na abordagem quantitativa (MINAYO, 2017).

Assim, apresento discussões gerais sobre entrevistas em profundidade, mas com foco na abordagem semiestruturada, que é a técnica de entrevista mais frequente nas análises qualitativas (KALLIO et al., 2016). Procuro responder as seguintes perguntas: *o que são entrevistas semiestruturadas? Como elas podem ser usadas na ciência política?* O objetivo é trazer os principais aspectos de entrevistas individuais semiestruturadas como técnica qualitativa de coleta de dados (BERRY, 1999; DUARTE, 2005) e apresentar uma introdução ao seu uso. Embora existam diversos trabalhos, inclusive no formato de manuais ou guias, sobre como fazer entrevistas, eles são, em geral, voltados para as ciências sociais, com maior ênfase na sociologia ou antropologia (BONI; QUARESMA, 2005; DUARTE, 2005; GASKELL, 2003; LEGARD et al., 2003; MINAYO, 2012), focados em etapas específicas da execução de entrevistas (BERRY, 1999; CROUCH; MCKENZIE, 2006; DWORKIN, 2012; KALLIO et al., 2016; MARSHALL, 1996) ou, quando compreendem a técnica como um todo, são livros que aprofundam cada aspecto do método em diferentes capítulos (MOSLEY, 2013; WENGRAF, 2001).

Evidentemente, esses estudos são importantes e trazem recomendações valorosas para o tema em questão, mas considerado necessário oferecer um material que sintetize os principais aspectos gerais da técnica de entrevistas qualitativas com foco na sua aplicação na ciência política. Conteúdos assim são relevantes tanto para tentar preencher a aparente lacuna de materiais didáticos sobre o tema na referida área, quanto para oferecer um suporte técnico para a decisão de pesquisadores(as) sobre qual método devem utilizar. Se no Brasil a escolha do método a ser aplicado em uma pesquisa tem sido, muitas vezes, pautada na afinidade do(a) acadêmico(a) com a metodologia ou em suas noções pré-definidas, a ideia é que ela passe a ser tecnicamente embasada (MARSHALL, 1996; REZENDE, 2011). Assim, a expectativa é que este artigo possa contribuir para que estudantes e pesquisadores(as) da área pouco familiarizados(as) com essa abordagem, possam encontrar aqui uma visão geral da sua aplicação e obter conhecimento suficiente para i) avaliar a adequação dessa técnica ao seu problema de pesquisa e, caso seja adequado, ii) ter referências iniciais para se aprofundar na técnica e usá-la com o rigor necessário. Assim, assumo o *trade-off* de tangenciar algumas questões de maneira mais superficial para que seja possível cobrir todas as etapas essenciais do uso de entrevistas - da teoria à análise.

Dito isso, este artigo está organizado da seguinte forma. Além desta introdução, na segunda seção, apresento a abordagem semiestruturada vis-à-vis outras técnicas de entrevista, destacando os principais aspectos que a diferenciam das demais e quais são os desafios centrais para a sua execução. Na terceira seção, exponho as principais etapas da realização de entrevistas - teoria, roteiro/tópicos guia, seleção de entrevistados(as), aplicação das entrevistas e consolidação e análise dos dados. Foco nas discussões centrais para cada passo do processo e expo-

nho exemplos práticos de estudos da área que utilizam a referida técnica. Na quarta e última seção, trago reflexões finais e concluo o trabalho.

Entrevistas:ão?

Por meio das entrevistas é possível perguntar aquilo que não conseguimos observar: sentimentos, intenções e pensamentos, por exemplo. Assim, entrevistas são feitas para compreender a perspectiva do outro, partindo do pressuposto de que tal perspectiva é significativa e pode ser conhecida e explicitada (PATTON, 2002). Elas permitem a melhor compreensão de mecanismos causais, podem servir como principal fonte de dados e, ainda, podem gerar dados que serão analisados de forma quantitativa posteriormente (LYNCH, 2013; MOSLEY, 2013). Além disso, entrevistas em profundidade são apropriadas quando se tem como objetivo descrever de forma detalhada um dado meio social a fim de compreendê-lo, quando se deseja usar seus resultados para desenvolver um referencial conceitual para pesquisas subsequentes ou quando se deseja testar conceitos. Isto é, utilizar os dados obtidos para testar expectativas que foram geradas fora de uma teoria já estabelecida (GASKELL, 2003).

Existem variadas classificações de entrevistas. Uma primeira distinção relevante é entre as entrevistas individuais e as grupais. Embora compartilhem pontos norteadores gerais, pelo fato de ambas usarem abordagem semiestruturada (GASKELL, 2003; HERTEL, 2009; MOSLEY, 2013), os dois tipos possuem algumas especificidades. Entrevistas grupais ou grupos focais também possuem foco no comportamento, atitudes e opiniões dos indivíduos, mas sob a perspectiva da construção de consensos e reação a pontos de divergência. Alguns estudiosos defendem que entrevistas grupais podem ser preferíveis às individuais, pois permitem observar como os indivíduos interagem e reagem a situações de conflito, além da possibilidade de usar recursos extras (projeções, metáforas etc.) (GASKELL, 2003). No entanto, é importante ponderar os custos logísticos que elas exigem e as suas limitações metodológicas (HERTEL, 2009).

Já entre as entrevistas individuais, diversas são as possibilidades: entrevista de *survey*, entrevista de aconselhamento, entrevista "diário", entrevista de história de vida, entrevista etnográfica, entrevista não estruturada/informal e conversações. Também é possível definir entrevistas a partir de três abordagens principais: entrevista informal de conversação, abordagem geral de entrevista guiada e entrevista aberta padronizada (BERRY, 1999; PATTON, 2002). O primeiro tipo se refere a uma entrevista em estilo de conversa continuada, também chamada de entrevista não estruturada ou etnográfica, apropriada para explorar um dado tema de interesse de maneira mais fluida, sendo tipicamente usada emção (DUARTE, 2005; PATTON, 2002). Seus resultados podem, inclusive, ajudar a construir roteiros semiestruturados ou questionários para entrevistas fechadas posteriores (DUARTE, 2005). No segundo tipo, a entrevista é guiada por uma espécie de lista de tópicos, cujo objetivo é garantir que todos os pontos relevantes foram cobertos. Desta forma, é uma abordagem adequada para adquirir informações específicas sobre o objeto de pesquisa. Por fim, na última abordagem as perguntas são estruturadas de forma mais sistematizada e o objetivo é reduzir o grau de variação nas respostas dadas pelos(as) entrevistados(as) em cada pergunta (BERRY, 1999; GASKELL, ; WOOFFITT; WIDDICOMBE, 2006).

De maneira geral, então, temos entrevistas abertas, semiabertas ou fechadas, sendo as duas primeiras, respectivamente, entrevistas guiadas por um tema central e por um roteiro-base. Ambas apresentam uma abordagem em profundidade, sendo qualitativas. Já as entrevistas fechadas são quantitativas em sua maioria e possuem abordagem linear (DUARTE, 2005). Na técnica

qualitativa, a ordem das perguntas e as palavras usadas podem variar, ao passo que na quantitativa, a ordem e as palavras utilizadas importam e devem ser padronizadas (WOOFFITT; WIDDICOMBE, 2006). Neste artigo, tratarei das entrevistas semiestruturadas, isto é, entrevistas qualitativas em profundidade e semiabertas (DUARTE, 2005).

ão é, , semiestruturada? É"conversaósito"(LEGARD et al., 2003), sendo .,ãodos(as) entrevistados(as) (GASKELL, 2003). Ao contrário da pesquisa quantitativa que foca na generalização e no efeito médio (GERRING, 2017; GOERTZ; MAHONEY, 2012), esse método foca em símbolos, significados, crenças, atitudes, valores e motivações (GASKELL, 2003; LEGARD et al., 2003; DUARTE, 2005).

Além desta definição geral, tais entrevistas podem ser delimitadas de acordo com algumas características principais (LEGARD et al., 2003). Um primeiro ponto é que elas apresentam uma combinação entre estrutura e flexibilidade: na medida em que seguem um tópico guia ou roteiro com os principais temas a serem abordados (estrutura), o(a) pesquisador(a) tem liberdade para voltar numa pergunta anterior, aprofundar um determinado tópico etc. (flexibilidade). É diferente de um questionário estruturado de *survey*, mas possui estrutura por meio dos tópicos guia ou do seu roteiro (DUARTE, 2005; LEGARD et al., 2003). Neste sentido, duas outras características se destacam: a natureza interativa e a possibilidade de aprofundar questões. No primeiro caso, a entrevista semiestruturada é fruto da interação entre o(a) pesquisador(a) e o(a) entrevistado(a). No segundo, o(a) entrevistador(a) procura aprofundar as respostas obtidas a partir das chamadas *follow up questions*, isto é, perguntas que pedem para que o(a) entrevistado(a) desenvolva uma resposta dada em um determinado tópico (LEGARD et al., 2003) (ex.: "poderia me dizer algo mais sobre isso?" (GASKELL, 2003)).

Algumas questões decorrem dessa natureza interativa da técnica, como a maneira pela qual o formato da entrevista afeta os dados coletados e os efeitos causados pelo(a) próprio(a) entrevistador(a) nas respostas obtidas. No primeiro caso, discute-se que a entrevista pode não só mapear opiniões-pré existentes, como também participar da formação dessa percepção. Ou seja, é um método que além de compreender percepções, também gera conhecimento e informação. Isso ocorre porque o indivíduo pode ser questionado sobre temas sobre os quais nunca havia pensado antes (CROUCH; MCKENZIE, 2006; LEGARD et al., 2003). Assim, as entrevistas podem ser consideradas um processo de co-construção dos dados, na medida em que há construção de realidade e sentimentos (BERRY, 1999; GASKELL, 2003; ROULSTON et al., 2003; SIONEK et al., 2020). No segundo ponto, relacionado aos chamados efeitos do(a) entrevistador(a), o comportamento de quem entrevista, sua linguagem e nuances, pode afetar as respostas recebidas e, portanto, os dados coletados. Por isso, a variação no *rapport* é um dos problemas associados à natureza interativa das entrevistas (MOSLEY, 2013). Desta forma, gravar as entrevistas é essencial, uma vez que tomar notas ao longo do processo poderia afetar a forma dos dados (LEGARD et al., 2003).

A questão dos efeitos do(a) entrevistador(a) é considerada positivista e se manifesta, então, de duas maneiras. A primeira é relacionada à forma como os(as) entrevistados(as) enxergam quem entrevista e, potencialmente, mudam suas respostas a partir disso. A segunda, considera os vieses do(a) entrevistador(a) ao interpretar as respostas obtidas, o que será influenciado não só pela teoria, mas pelo seu próprio contexto e percepções. Para os(as) interpretativistas, isso está ligado à maneira como o(a) entrevistador(a) se posiciona em relação ao(à) entrevistado(a), tendo consciência de todas essas variações (MOSLEY, 2013). Alguns especialistas ressaltam as questões de intersubjetividade e reflexividade. Neste sentido, há o

entendimento de que as entrevistas propiciam a geração de um material, e não apenas a coleta de dados (MINAYO; GUERRIERO, 2014).

É pertinente destacar, também, que as entrevistas semiestruturadas podem ser usadas em uma abordagem multimétodo (DUARTE; 2005; GASKELL; 2003; LYNCH, 2013). Numa mesma entrevista é possível utilizar perguntas qualitativas e quantitativas (DUARTE, 2005). Ainda, existe a alternativa de usar entrevistas grupais para observar opiniões e processos de consenso e divergência combinadas às entrevistas individuais para compreender melhor as motivações e o contexto particular do indivíduo (GASKELL, 2003). Por conseguinte, as entrevistas em profundidade podem ser aplicadas de forma combinada a outros métodos, seja para apoiar um desenho de *survey*, por exemplo, e dar suporte para sua interpretação, seja para levantar informação contextual a fim de entender os resultados de uma pesquisa (*ibidem*).

Na ciência política, entrevistas em profundidade também são usualmente utilizadas para a construção de experimentos de campo (GERRING, 2017). Informações qualitativas também podem ser usadas para melhorar a inferência causal em desenhos que usam técnicas quase-experimentais, como o pareamento (*matching*), por exemplo. Elas auxiliam na redução de erros de mensuração, permitem a construção de um intervalo de confiança qualitativo sobre a magnitude do efeito e ajudam a lidar com *confounders* (GLYNN; ICHINO, 2015). Dentro dos métodos qualitativos, há também o uso de entrevistas de elites em desenhos de *process-tracing* (TANSEY, 2007).

Além dessas áreas, as entrevistas em profundidade podem ser utilizadas como técnica qualitativa de avaliação de impacto de políticas públicas (BATISTA; DOMINGOS, 2017), no campo do marketing político (VEIGA; GONDIM, 2001) e em pesquisas sobre diversos outros temas como *lobbying*, violência vivida durante a guerra, preferências de empresários sobre políticas de bem-estar social (MOSLEY, 2013), estudos sobre a implementação de políticas públicas (GUIMARÃES, 2018, LOTTA et al., 2018) ou ainda, o estudo de partidos políticos no Brasil (MAGLIA, 2020; ROCHA, 2019), para citar apenas alguns exemplos.

Nas avaliações, entrevistas deste tipo são especialmente relevantes para entender como o impacto de um programa afeta cada pessoa de forma distinta (BATISTA; DOMINGOS, 2017). Isto é, para compreender as perspectivas de participantes de um determinado programa, da equipe que implementa o programa e de outros grupos relevantes à execução de uma política pública ou programa social. Para tanto, o uso desta técnica lança luz sobre as experiências das pessoas com o programa, como elas percebem e sentem as ações desenvolvidas e assim por diante (PATTON, 2002). Nessa mesma linha, um estudo sobre a implementação de uma política específica com foco nos(as) burocratas de nível de rua, como o de Guimarães (2018), ajuda a observar diferentes aspectos relevantes das perspectivas de pessoas diretamente responsáveis por entregar aquele serviço às beneficiárias. No marketing político, a exploração das experiências de vida e cotidiano de uma pessoa podem ajudar a compreender melhor as suas visões sobre a política. Em comum, está o foco nas crenças, decisões, valores, expectativas, desejos etc. É um "diálogo assimétrico" no qual o(a) entrevistador(a) busca coletar dados e o(a) entrevistado(a) é a fonte das informações. O foco, consiste, então, na fala do(a) entrevistado(a) (BATISTA; DOMINGOS, 2017; VEIGA; GONDIM, 2001).

Assim, no lugar de pesquisas mais imersivas, como entrevistas participantes, as aplicações de entrevistas na ciência política tendem a ser mais concisas, mais pontuais. Na medida em que isso afeta o nível de profundidade atingido, influencia também a validade interna

dos dados coletados. Por outro lado, pode haver ganho de validade externa, pela possibilidade de contemplar mais grupos sociais ou mais locais. Neste sentido, existem quatro desafios principais para a aplicação de entrevistas na ciência política: ética, amostragem (seleção de entrevistados(as)), validade e confiabilidade.

A questão ética está associada ao risco que uma dada pesquisa pode gerar aos(às) seus(suas) participantes. Esse risco é considerado pequeno em comparação a pesquisas médicas, por exemplo. Todavia, eles existem e devem ser considerados nas ciências sociais. No âmbito de entrevistas em profundidade, alguns exemplos desse risco são tocar em temas sensíveis que podem relembrar experiências traumáticas dos(as) respondentes ou a falta de anonimato ao reportar os dados, colocando em risco a integridade do(a) entrevistado(a). Nesse caso, o consentimento informado é central para garantir a ética em pesquisas que envolvem interações com seres humanos. Ele deve comunicar ao(à) informante sobre os objetivos da pesquisa, esclarecer que sua participação é voluntária e que ele(a) é livre para deixar de participar a qualquer momento e não sofrerá nenhuma punição por isso, e apresentar os riscos de participação e o que será feito para mitigá-los. Também é preciso garantir a confidencialidade dos(as) respondentes (MOSLEY, 2013).

Muitas vezes, esse consentimento é oficializado por um Termo de Consentimento, podendo ocorrer, em alguns casos, de forma oral - quando o(a) entrevistado(a) não é alfabetizado(a), por exemplo. A construção e a assinatura deste termo trazem diversos desafios para os(as) pesquisadores(as). O primeiro deles é que as instituições responsáveis pela aprovação de uma pesquisa, que no caso das Universidades brasileiras seriam os Comitês de Ética em Pesquisa (CEPs) e, a nível internacional, os *Institutional Review Boards* (IRBs), tendem a preferir uma linguagem mais formal e técnica, que pode ser menos acessível aos(às) entrevistados(as). Existe uma crítica de que, em alguns casos, os documentos são escritos de tal forma que protegem mais a instituição que financia ou está vinculada à pesquisa do que a pessoa que participa do estudo em si. Com isso, é papel do(a) pesquisador(a) garantir que o Termo seja interpretável, isto é, explicar ao(à) informante, em linguagem acessível e clara, do que se trata aquele documento (BROOKS, 2013; MOSLEY, 2013).

Um outro ponto, é que a própria assinatura do Termo pode afetar a motivação do(a) entrevistado(a) de participar da pesquisa. Há indicação de que o consentimento oral está associado a maior motivação a participar do estudo quando comparado ao consentimento escrito. Dado que o consentimento informado é irrevogável, por ser condição básica de respeito ao ser humano que integra um estudo, o(a) pesquisador(a) deve estar atento(a) à forma como apresenta essa etapa ao entrevistado(a), com honestidade e tranquilidade (BROOKS, 2013; MOSLEY, 2013).

Minayo e Guerriero (2014) alertam que a pesquisa qualitativa envolve desafios éticos que vão além dos instrumentos e procedimentos definidos pelos Comitês de Ética - eles envolvem o processo da pesquisa como um todo. As autoras destacam três fatores relacionados às questões de ética em pesquisa: o comportamento do(a) pesquisador(a) em campo; a forma como o material é analisado; e o relato feito a partir deste material. Nesse sentido, o(a) pesquisador(a) deve sempre manter o respeito em relação aos indivíduos com quem interage. Isso se estende às demais etapas da pesquisa. Com relação à publicação dos dados ou do relatório de pesquisa, a decisão do que deve ser incluído e exposto na análise é, também, ética. É crucial que o(a) pesquisador(a) seja transparente e descreva em detalhes o seu processo de pesquisa, como foram estabelecidas as relações em campo, a fim de que suas conclusões possam ser avaliadas nesse contexto. Um exemplo prático de medidas éticas que podem ser tomadas na apresentação

dos resultados é suprimir ou mudar os nomes das pessoas e/ou instituições para nomes fictícios (MINAYO; GUERRIERO, 2014).

Os próximos três desafios - seleção dos(as) entrevistados(as), validade e confiabilidade - estão diretamente ligados a algumas das etapas de planejamento e execução das entrevistas e, portanto, serão pormenorizados ao longo da próxima seção. Numa visão inicial, o desafio da seleção dos(as) entrevistados(as) se refere a como definir a população das pessoas que poderão ser entrevistadas; e como escolher os indivíduos a partir dessa população, algo que depende da construção da teoria e de como as hipóteses serão testadas. Já o desafio da validade interna está associado a fatores como a amostra/seleção de entrevistados(as) e a forma como a entrevista foi conduzida (se as perguntas conseguiram capturar a informação relevante para a pesquisa, se as respostas dos(as) respondentes foi verdadeira e, havendo viés nas respostas, se o(a) entrevistador(a) foi capaz de detectá-lo). Em suma, refere-se a todos os aspectos que compõem o processo de planejamento e aplicação das entrevistas e que podem afetar os dados coletados quanto a sua capacidade de capturar o objeto de estudo da pesquisa (MINAYO, 2012; MOSLEY, 2013). Finalmente, o último desafio ao realizar entrevistas é a confiabilidade das respostas recebidas. Isso está relacionado ao quão consistente é o instrumento de mensuração. No caso das entrevistas, se podemos confiar que ao repetir uma mesma entrevista com um mesmo indivíduo nós receberíamos respostas similares, consistentes (MOSLEY, 2013).

2. Como Usar Entrevistas? Da teoria à análise dos dados

A figura 1, a seguir, apresenta um esquema visual que resume as principais etapas aqui abordadas. Cada passo será discutido em um subtópico específico, exceto pela etapa de submissão da pesquisa a um comitê de ética, a qual já foi discutida na seção anterior. O consentimento informado é crucial para realizar entrevistas, mas a necessidade de submeter o estudo para aprovação depende da instituição à qual o(a) pesquisador(a) está vinculado(a)[2]. Tal etapa seria, naturalmente, anterior à aplicação das entrevistas e está caracterizada em cinza na figura 1 por ser parte importante do planejamento de entrevistas.

Figura 1: Etapas para a realização de Entrevistas Semiestruturadas.



Fonte: Elaboração própria com base em Gaskell (2003) e Mosley (2013).

2 Os documentos exigidos pelos Comitês de Ética para a realização de pesquisa com seres humanos são vários, em um processo bastante exigente, por ser pautado predominantemente nas pesquisas biomédicas (MINAYO; GUERRIERO, 2014). Embora algumas resoluções mais recentes flexibilizem um pouco esse processo para pesquisas nas ciências sociais, o(a) pesquisador(a) deve se planejar para coletar e entregar os documentos exigidos pelo Comitê da sua instituição e considerar o tempo de avaliação da sua solicitação para realizar a pesquisa. Como isso afeta o cronograma da análise em si, é recomendável checar logo no início da pesquisa se seria necessário submetê-la a um CEP, quais documentos são exigidos e quanto tempo pode levar até que a pesquisa seja aprovada, antes de partir para a aplicação das entrevistas em si.

Teoria

O primeiro passo para realizar uma entrevista semiestruturada deve ser o desenvolvimento de um referencial teórico ou conceitual cujo conteúdo vai nortear a análise. Desta forma, é importante que conceitos e temas centrais sejam previamente identificados (GASKELL, 2003; CROUCH; MCKENZIE, 2006). Assim, o objeto de estudo deve ser definido a partir de uma pergunta de pesquisa ou problematização e situado em um debate teórico (MINAYO, 2012). Tal ponto pode parecer evidente, considerando que um problema de pesquisa deve contrapor de maneira clara as contribuições da análise proposta em relação ao que já foi discutido na literatura especializada sobre o tema (GUSTAFSSON; HAGSTROM, 2018). Mesmo assim, por ser crucial, esse ponto deve ser lembrado e destacado.

Além disso, outra possibilidade é a utilização da teoria fundamentada (*grounded theory*) (GASKELL, 2003). A teoria fundamentada pode ser entendida como a descoberta da teoria a partir dos dados (GLASER; STRAUSS, 2006 [1976]). Ela é uma metodologia cujos direcionamentos para coletar e analisar os dados permitem gerar teorias de médio-alcance (CHARMAZ; BELGRAVE, 2015). A partir de uma abordagem indutiva, a ideia é gerar teorias que sejam ancoradas nos dados e capazes de explicar o fenômeno analisado. Para além da sociologia, a teoria fundamentada é bastante usada na antropologia e em estudos organizacionais, por exemplo. Na ciência política, dois pontos são contenciosos para o uso desse método: o debate sobre as metodologias qualitativa e quantitativa e a ênfase na abordagem dedutiva e na verificação de teorias, ao invés da geração destas. De maneira geral, o uso desse método na ciência política seria limitado, sendo possivelmente mais adequado para os estudos de sociologia política e políticas públicas (BECKER, 2012).

Uma vez definido o referencial teórico, é possível partir para duas questões essenciais para a aplicação do método: o que perguntar e a quem perguntar. A primeira questão está ligada à definição do tópico guia ou roteiro, ao passo que a segunda concerne a seleção de entrevistados(as) (GASKELL, 2003). Ambas são características definidoras das entrevistas semiestruturadas e não devem ser confundidas com técnicas quantitativas de construção de *survey*, tampouco com as questões de aleatorização e representatividade que permeiam a definição de amostras (DUARTE, 2005; GASKELL, 2003).

Tópicos Guia

A elaboração apropriada do tópico guia é crucial para a execução de uma boa entrevista. Como diz o próprio nome, os tópicos guiam a pesquisa e permitem que seja possível alcançar os objetivos da investigação. Neste sentido, devem ser construídos para permitir que a entrevista ocorra de forma lógica e plausível (GASKELL, 2003). O roteiro deve ser ancorado nas questões centrais da pesquisa e nas suas teorias e hipóteses. Essas hipóteses podem surgir à medida que as respostas são obtidas, o que pode provocar mudanças nos tópicos guia (DUARTE, 2005; MINAYO, 2012). Assim, o(a) pesquisador(a) deve pautar suas ações nas teorias e hipóteses, mas manter a mente aberta para questioná-las (MINAYO, 2012).

Além de se basear no referencial teórico apropriado, o(a) pesquisador(a) deve reconhecer o campo. Isto é, observar e conversar antecipadamente com pessoas que podem ser relevantes para a análise, discutir o tema com colegas e usar a criatividade (GASKELL, 2003; MINAYO, 2012). Quando estiver em campo, manter um diário com anotações das falas, depoimentos e visões apresentadas pelos(as) interlocutores(as) pode ajudar o(a) pesquisador(a) a tecer uma narrativa coletiva sobre as vivências e experiências dos indivíduos cuja realidade é analisada (MINAYO, 2012).

Os tópicos facilitam a estruturação da entrevista de uma forma que seja possível garantir que os mesmos pontos foram tratados com todos os(as) entrevistados(as).

Todavia, o(a) entrevistador(a) tem liberdade para explorar esses tópicos da forma que achar mais apropriada (PATTON, 2002). É a combinação entre estrutura e flexibilidade tratada por Legard et al. (2003).

No que se refere à formulação do roteiro da entrevista em si, é ideal que este seja construído como uma sequência de títulos de parágrafos. Não devem ser elaboradas perguntas fechadas a serem simplesmente repetidas para os(as) entrevistados(as), mas sim temas que se deseja conversar com eles. É necessário sempre relembrar o objetivo de compreender as percepções, visões de mundo, crenças e atitudes dos(as) entrevistados(as) sobre determinados temas (GASKELL, 2003). O roteiro desse tipo de entrevista não precisa ser extenso, tendo em geral entre 4 e 7 perguntas, sendo ideal que não exceda uma página. Tais questões precisam ser amplas o suficiente para permitir o aprofundamento do tópico e não devem ser redundantes. Quem entrevista deve explorar as respostas apresentadas (com perguntas de *follow up*, que visam aprofundar as respostas) e somente partir para a próxima pergunta quando o tópico tiver sido esgotado (DUARTE, 2005; GASKELL, 2003).

Os roteiros, então, tendem a ser compostos por dois níveis - um que aborda os temas centrais da pesquisa, com perguntas que permitem que os(as) informantes falem livremente sobre os tópicos, e outro que abrange as perguntas *de follow up* (KALLIO et al., 2016). Vale salientar, no entanto, que muitos roteiros possuem bem mais que sete perguntas, às vezes em torno de 15 delas (HORTON et al., 2004; ROCHA, 2019). Mais importante que seguir uma regra de quantidade de perguntas é garantir que todos os aspectos relevantes para compreensão do objeto de pesquisa foram inseridos.

De maneira prática, é ideal que o tópico guia possa ser usado em certa medida para monitorar a entrevista, caso haja alguma limitação ou estimativa de tempo e se saiba quantos tópicos ainda precisam ser abordados (GASKELL, 2003; HORTON et al., 2004). Além disso, os tópicos guia deverão ajudar não só a guiar a entrevista, mas também a orientar a sua transcrição posteriormente, servindo como um esquema preliminar. Todavia, os subtópicos não são inflexíveis, podendo ser modificados para entrevistas subsequentes. É possível que o(a) entrevistador(a) perceba que um dado tópico não contribuiu muito para o desenvolvimento da discussão ou até que havia um tópico importante que não constava na estrutura anterior. Consequentemente, é comum que o(a) pesquisador(a) inicie as entrevistas com um roteiro e as finalize com outro. Entretanto, qualquer mudança deve ser sempre registrada, junto com a justificativa que levou a essa alteração (DUARTE, 2005; GASKELL, 2003, MINAYO, 2012).

Com base em uma revisão sistemática da literatura, Kallio et al. (2016) identificaram cinco etapas no processo de elaboração dos tópicos guia que, em certa medida, sumarizam o que já foi discutido. Os passos são: i. identificar os pré-requisitos para aplicar uma entrevista semiestruturada; ii. considerar o conhecimento prévio já construído sobre aquele tema (teoria e material empírico); iii. elaborar um guia preliminar; iv. testar essa versão em campo, havendo três possibilidades para isso: testagem interna (com membros da equipe pesquisadora), avaliação de especialistas (pesquisadores(as) que não integram a equipe) e testagem em campo (com indivíduos que poderiam ser entrevistados(as)). E, então, com base nas informações obtidas sobre a abrangência e importância das perguntas feitas, aprimorar o roteiro; e v. por fim, apresentar a versão final do guia.

Seleção de Entrevistados

A partir do referencial teórico e da construção do tópico guia, o próximo passo é selecionar os(as) entrevista-

dos(as). Como já dito, este é um dos desafios enfrentados na utilização da técnica de entrevistas (MARSHALL, 1996; MOSLEY, 2013). Na aplicação de técnicas que envolvem interação entre pesquisador(a) e indivíduos analisados, os conceitos de amostra e saturação são centrais. Eles estão ligados a uma pergunta mais profunda, isto é, em que medida tais técnicas podem ser consideradas científicas? Até que ponto as falas coletadas em determinado número de entrevistas refletem a perspectiva de um grupo estudado? (MINAYO, 2017).

De maneira geral, dois caminhos principais podem ser tomados para selecionar os(as) entrevistados(as): a amostragem probabilística ou não probabilística (intencional) (LYNCH, 2013). Qual deles é o melhor? Alguns argumentos defendem que o uso prevalente de abordagem não probabilística está baseado numa condição ontológica impossível de homogeneidade do mundo social (LUCAS, 2012). Muitos(as) qualitativistas reconhecem as limitações das conclusões alcançadas por meio de amostras não aleatórias e que, portanto, estudos qualitativos não possuem intenção de generalização. Todavia, Lucas (2012) discute que muitas investigações dentro dessa abordagem acabam por trazer conclusões generalizáveis, mesmo com as limitações de amostras intencionais.

O argumento é de que entrevistas em profundidade perdem boa parte da sua capacidade de gerar conhecimento quando uma seleção teoricamente mais rigorosa dos(as) entrevistados(as) não é posta em vigor (LYNCH, 2013; LUCAS, 2012). Embora essa discussão seja muito importante, é interessante observá-la com base no debate metodológico da ciência política entre as tradições quantitativas e qualitativas. A análise qualitativa foca nas singularidades e na intensidade de um fenômeno, e não na identificação de um padrão, como na análise quantitativa.

Numa distinção rápida, as análises qualitativas são expressas em linguagem natural; com abordagem *Small-N*, casos selecionados de forma não aleatória, mas sim proposital, e foco em eventos específicos. Nesse caso, é especialmente importante que os(as) entrevistados(os) selecionados(as) representem todos os pontos de vista relevantes para a pesquisa, ou seja, a diversidade nas fontes é essencial. Assim, os(as) participantes seriam selecionados(as) não de forma probabilística, mas sim intencional. A seleção seria, então, baseada na teoria e no potencial de informação que tais indivíduos podem agregar a análise (MACDOUGALL; FUDGE, 2001; GASKELL, 2003). Por outro lado, análises quantitativas utilizam abordagem probabilística e estatística, *Large-N*, amostras aleatórias e focam na generalização (GERRING, 2017). Enquanto análises quantitativas buscam a identificação do efeito causal médio (*effects-of-causes*), os estudos qualitativos objetivam compreender o que causou determinado efeito (*causes-of-effects*) (GOERTZ; MAHONEY, 2012). Se para a análise quantitativa selecionar os casos pela variável dependente não é indicado, para a análise qualitativa tal seleção pode ser a mais apropriada para entender os mecanismos causais que levaram a um dado evento *(ibidem)*.

Assim, reconhecer os limites do uso de técnicas qualitativas como as entrevistas em profundidade é essencial, mas precisamos ter cuidado para não sobrepor uma lógica quantitativa sobre um desenho de pesquisa qualitativo. Posto isso, a resposta para pergunta sobre qual o melhor caminho a seguir, se probabilístico ou intencional, é: depende. Essa decisão deve ser guiada pela pergunta de pesquisa e pelo objetivo do estudo (MARSHALL, 1996; MINAYO, 2017; MOSLEY, 2013; LYNCH, 2013).

Como explica Mosley (2013), se a intenção é generalizar a partir dos dados coletados, transformando-os posteriormente em *data set observations* (DSO), isto é, observações quantitativas, uma amostragem probabilística pode ser mais adequada. Todavia, se o desenho

de pesquisa é qualitativo, isto é, a pergunta de pesquisa procura entender por que um determinado evento ocorreu ou como ele ocorreu, os dados serão analisados como *causal process observations* (CPO)[3]. Neste caso, não falamos em *amostragem*, mas sim na *seleção de entrevistados(as)* (BRADY, 2010; GASKELL, 2003; MOSLEY, 2013).

Pesquisas qualitativas por vezes selecionam aleatoriamente o objeto no qual se está interessado(a) em estudar, por exemplo, firmas, ou escolas, e dentro dessas organizações selecionam intencionalmente as pessoas que poderiam trazer maior riqueza de informações para a análise. Assim, não é uma completa aleatorização. A amostra aleatória pode ser desenhada de variadas formas para as entrevistas em profundidade. Uma das possibilidades é desenhar amostras aleatórias estratificadas que podem ser usadas para identificar os chamados casos raros. No entanto, amostras aleatórias nem sempre são factíveis por diversos motivos (falta de recurso, entrevistados(as) que desistem de participar e assim por diante). Nesses casos, uma opção é combinar aleatorização com a técnica de "bola de neve", que consiste em selecionar alguns indivíduos aleatoriamente e usar sua rede pessoal para tentar recrutar mais pessoas (LYNCH, 2013).

Já na abordagem não aleatória, para a qual o termo correto seria "seleção" e não "amostragem" (GASKELL, 2003), o objetivo não é entrevistar indivíduos de um mesmo perfil, mas sim "variantes de um conjunto social particular" (CROUCH; MCKENZIE, 2006, p. 493). Esse é um ponto crucial para a validade e confiabilidade dos dados levantados via entrevista (DUARTE, 2005). É também uma questão fundamental na análise qualitativa de maneira geral, na qual a inferência decorre de poucos casos intencionalmente selecionados (GERRING, 2017). Nesse caso, como esclarece Minayo (2017), a fidedignidade da análise qualitativa está mais relacionada com a diversidade do objeto e com o desenho de pesquisa em si, do que com o tamanho da amostra. Em suas palavras, "(...) uma amostra qualitativa ideal é a que reflete, em quantidade e intensidade, as múltiplas dimensões de determinado fenômeno e busca a qualidade das ações e das interações em todo o decorrer do processo." (MINAYO, 2017, p. 10).

Sendo não probabilístico, quatro tipos de seleção de entrevistados(as) podem ser destacados: por conveniência (ou acidental), intencional (ou de julgamento), bola de neve e a de contatos intersticiais. No primeiro, a seleção depende da viabilidade, quando as fontes são selecionadas por proximidade ou disponibilidade. Podem ser muito úteis quando o acesso aos(às) potenciais entrevistados(as) é difícil ou na fase preliminar da pesquisa. No segundo, com base em seu julgamento, o(a) pesquisador(a) define os(as) entrevistados(as) a partir do conhecimento que eles(elas) possuem sobre o tema ou almejando alcançar representação subjetiva. Pode ser usada para garantir que casos raros ou negativos façam parte da pesquisa. No terceiro, chamado seleção bola de neve, ou guiada pelo(a) respondente ou de referência em cadeia, consiste em incluir entrevistados(as) à medida que as entrevistas são realizadas, com base na indicação de outros(as) respondentes. Assim, ao longo da pesquisa novos(as) entrevistados(as) podem ser selecionados(as), caso seja identificada a necessidade da sua inclusão para a pesquisa (DUARTE, 2005; LYNCH, 2013). Rocha (2019), por exemplo, ao investigar a formação de uma nova Direita no Brasil, utiliza a técnica de bola de neve ao questionar em diferentes grupos sociais quais seriam as pessoas consideradas por cada um deles como mais importantes para compreender o fenômeno estudado. No quarto e último tipo, os(as) entrevistados(as) podem ser pessoas numa fila, num

3 Brady (2010) discute a complementaridade das abordagens qualitativa e quantitativa com base em duas definições de observação, respectivamente quantitativas e qualitativas: *data set observations (DSO)* e *causal process observation (CPO)*. Não obstante, o termo *data set observations* seria abrangente, englobando observações qualitativas e quantitativas, visto que uma observação isolada de processo causal também pode ser inserida em um *data set* retangular (ou seja, todos os valores de uma linha em um formato retangular). Já para Gerring (2017), observações quantitativas são comparáveis, enquanto as qualitativas, não.

táxi, pessoas que estão em situações corriqueiras que podem espontaneamente compartilhar informações. É útil tomar notar e, se possível, gravar (com consentimento) estas falas, as quais podem trazer informações que seriam possivelmente inacessíveis por outros meios (LYNCH, 2013).

MacDougall e Fudge (2001) expõem três etapas para selecionar informantes: preparar, contatar e acompanhar. A primeira etapa, *preparar*, se inicia com a descrição da amostra: quais são os locais, características e experiências que esses(as) entrevistados(as) deveriam ter, com base nos objetivos da pesquisa? A ideia é segmentar o meio com relação ao tema abordado. Por exemplo, se uma entrevista em profundidade está sendo aplicada para entender a percepção dos indivíduos sobre uma determinada política pública de educação, todos os grupos que podem ser afetados pela intervenção no meio escolar devem ser ouvidos. Esses grupos podem ser os(as) estudantes beneficiados(as), os(as) professores(as), a equipe gestora e técnica da escola e os pais dos(as) estudantes participantes, por exemplo (GASKELL, 2003). Maglia (2020), em sua pesquisa sobre as razões que levaram à criação de mais partidos de direita do que de esquerda no Brasil (de 1990 a 2018), entrevista lideranças de todos os 25 partidos políticos criados. Guimarães (2018), em sua análise sobre o papel das burocratas de nível de rua em uma política pública de abrigamento para mulheres vítimas de violência em Pernambuco, entrevista profissionais em diferentes níveis (educadoras, coordenadoras e diretoras). O objetivo é obter visões diversificadas sobre o tema (DUARTE, 2005).

Tomando como referência o exemplo da política educacional, para que esses grupos possam ser identificados, é necessário que se conheça o meio escolar. Nesse sentido, um primeiro passo do(a) pesquisador(a) seria conversar com alguém que conhece o meio, por exemplo uma pessoa que trabalha em outra escola da rede de ensino, e conversar com essa pessoa sobre como é a dinâmica escolar, quem seriam os grupos afetados pela política e assim por diante. Quando há o envolvimento de mais de um meio social, a complexidade da seleção aumenta (GASKELL, 2003). Alguns critérios podem ser utilizados para definir os grupos que devem ser inseridos na seleção de entrevistados(as). Um primeiro critério é o sociodemográfico, como idade ou faixa de renda. Outras possibilidades são os grupos estatísticos ou taxonômicos. E há ainda a possibilidade de grupos naturais, que são pessoas que compartilham um passado ou projeto futuro em comum, ou compartilham interesses e valores similares etc. A premissa central é que a fonte possa agregar informação relevante, confiável e diversificada sobre a realidade estudada (GASKELL, 2003; DUARTE, 2005; DWORKIN, 2012).

Depois disso, é necessário encontrar informações sobre a fonte (listar os principais grupos e fontes de informações sobre eles), quais seriam os contatos-chave para alcançar essas pessoas (pessoas que conhecem as comunidades) ou até *champions*, ou seja, pessoas que poderiam recrutar ou ajudar a recrutar entrevistados(as). O passo seguinte dentro dessa etapa é procurar por projetos correlatos ao seu, entrando em contato com colegas ou membros da comunidade e, se existirem, aprender com esses estudos anteriores sobre como eles entrevistaram, quais vínculos foram travados etc. Por fim, é importante fazer seleções alternativas: fazer esse processo para criar cerca de 3 a 4 listas de contatos, caso haja problemas no contato com a seleção inicial (MACDOUGALL; FUDGE, 2001).

Guimarães (2018) e Maglia (2020) trazem exemplos práticos da necessidade de listar contatos alternativos. No primeiro caso, a pesquisadora havia acordado, inclusive em termo assinado, que a Secretaria da Mulher de Pernambuco iria compartilhar o contato de profissionais de diferentes casas abrigos do estado. Como esse acordo não foi cumprido, a autora buscou outros meios de contatar

as profissionais, se valendo de contatos que ela mesma possuía por já conhecer o meio social das casas abrigo. Já Maglia (2020), planejava inicialmente entrevistar os presidentes fundadores de todos os partidos políticos investigados. Contudo, alguns deles haviam falecido e outros não retornaram a tentativa de contato. A autora decidiu, então, entrevistar lideranças que participaram do processo de criação do partido ou que, atualmente, exercem a função de presidentes da legenda.

Na etapa seguinte, *contatar*, é essencial travar uma boa relação com contatos chave, mas sem que haja grande dependência desse contato para realizar as entrevistas. Nesta etapa o passo-a-passo seria o seguinte: 1. abordagem inicial com o contato chave ou *champion* e informá-los de forma breve sobre o objetivo da pesquisa; 2. negociar com esses contatos, procurando o apoio deles; 3. negociação direta, isto é, contato direto com aqueles(as) que serão entrevistados(as); 4. confirmação: definir como será abordagem (por e-mail, por ligação etc.); e 5. envolvimento: se houve acordo para manter contato com essas pessoas. Se sim, há planejamento e recursos para tal? Entrevistados(as) podem ter interesse em conhecer o resultado da pesquisa (MACDOUGALL; FUDGE, 2001). Assim, é essencial considerar o tema estudado e a disponibilidade de tempo e a predisposição das fontes a falar (DUARTE, 2005). A última etapa, *acompanhar*, consiste, de forma geral, em informar participantes sobre resultados da pesquisa; definir meios de comunicação para atualizar sobre o andamento do estudo, se o projeto for contínuo; ou, ainda, convidar participantes para eventos públicos sobre a pesquisa e, por fim; definir algum(a) responsável por possível *advocacy* que seja resultante do projeto (MACDOUGALL; FUDGE, 2001).

Finalmente, quantas pessoas devem ser entrevistadas? Ao contrário da abordagem quantitativa, na qual a representatividade da amostra em relação à população é crucial, aqui poucas entrevistas podem ser suficientes para obter uma descrição aprofundada do fenômeno analisado (DUARTE, 2005). Existem algumas estimativas de quantidade que variam entre 5 e 50 entrevistas (GASKELL, 2003). Dworkin (2012) sugere um número entre 25 e 30 participantes, o qual permitiria i) examinar certas características ligadas às questões de pesquisa e conceitos; ii) maximizar os dados a serem coletados; e iii) aumentar a probabilidade de que casos e hipóteses negativos sejam explorados (p. 1320). Já Baker e Edwards (2012) sugerem uma variação mais ampla que vai de 12 a 60 entrevistados(as).

Assim, a quantidade de entrevistas ideal depende do escopo da pesquisa, de questões práticas como quantas entrevistas é possível transcrever em um determinado período e quantas entrevistas são necessárias para obter informações úteis para a investigação (BAKER; EDWARDS, 2012; GASKELL, 2003). Uma alternativa é realizar entrevistas até o ponto de saturação, no qual mais entrevistas não adicionam informação útil à pesquisa (GASKELL, 2003; MACDOUGALL; FUDGE, 2001). Esse nível saturação deve ser reportado, assim como todas as outras etapas da seleção de entrevistados(as) (BLEICH; PEKKANEN, 2013).

O conceito de saturação tem sido bastante debatido no campo da análise qualitativa. Um dos pontos discutidos é de que não há simplesmente um ponto de corte para uma amostra, visto que quanto mais informações são obtidas sobre um objeto, é possível que mais perguntas surjam, dando continuidade à necessidade de mais entrevistas, por exemplo (MINAYO, 2017). O fator essencial é que este número seja justificado com base no seu objeto de pesquisa - em sua complexidade e extensão, mas também apoiado em referências teóricas caras à pesquisa. Mas é essencial lembrar que a definição prévia dos instrumentos, como os tópicos guia ou o número de entrevistas, é importante, mas

não imutável (GASKELL, 2003; MINAYO, 2017). Sendo um processo dinâmico, o(a) pesquisador(a) pode identificar a necessidade de entrevistar mais pessoas a partir das informações obtidas em campo (MINAYO, 2017). Sugiro que, caso o(a) pesquisador(a) defina o número de entrevistas previamente, justifique em seu projeto que esse número pode passar por modificações - tanto por questões práticas, como a impossibilidade de acessar algumas pessoas, quanto pelas razões teóricas já citadas. Se essa mudança ocorrer, ela deve ser justificada e apresentada de forma transparente a(o) leitor(a).

Aplicação de Entrevistas

Uma vez que os instrumentos da pesquisa foram construídos - definição do quê e a quem perguntar - e a proposta de pesquisa já foi devidamente aprovada de acordo com os critérios éticos necessários, é hora de partir para a aplicação prática das entrevistas. A qualidade das informações coletadas em uma entrevista depende sobremaneira de quem entrevista (PATTON, 2002). Assim, o(a) entrevistador(a) precisa ser bem treinado(a) para aplicar esta técnica: habilidade e técnica são fatores essenciais para que uma entrevista seja executada com rigor (BATISTA; DOMINGOS, 2017; PATTON, 2002). Nesta subseção, vou explorar os principais aspectos para a boa execução de uma entrevista em profundidade.

De acordo com Gaskell (2003), entrevistas em profundidade podem durar entre 1 hora e 1 hora e meia. No caso de entrevistas por telefone, embora haja algumas evidências de que seu tempo e profundidade não variam significativamente em relação às entrevistas presenciais (CACHIA; MILLWARD, 2011; IRVINE, 2011; VOGL, 2013), uma duração que ultrapassa 1 hora pode se tornar longa demais nesse formato. Seria recomendável, então, planejar a entrevista para um período menor.

O início da entrevista deve abordar temas leves, nunca se deve iniciar pelo tópico de interesse. É importante que o(a) entrevistado(a) se sinta à vontade para falar. Deve-se criar um ambiente com clima amistoso e que permita o estabelecimento de uma relação de confiança com o(a) pesquisador(a) (BONI; QUARESMA, 2005). Neste ponto, a postura do(a) entrevistador(a) (*rapport*) demonstrando atenção e interesse à fala do(a) entrevistado(a) é crucial (DUARTE, 2005; GASKELL, 2003). Um ponto que merece destaque, e vai além das perguntas, é evitar considerar apenas respostas que confirmam seu argumento (DUARTE, 2005). Na lógica qualitativa, o objetivo é sempre buscar por informações que *desconfirmem ou desafiem* a teoria analisada (GERRING, 2017). Esta postura deve ocorrer não só ao analisar e reportar resultados, mas também na própria execução das entrevistas. Ademais, é importante salientar que essa técnica pode ser realizada não só pessoalmente, mas também por outros meios, como o telefone e a internet (DUARTE, 2005; LEGARD et al., 2003).

Segundo Patton (2002), há seis tipos de perguntas que podem ser feitas em uma entrevista. A primeira delas são perguntas sobre comportamento e experiência do(a) entrevistado(a) para entender as ações e as atividades empreendidas e as experiências vividas pela pessoa em um determinado contexto ("Se eu te seguisse em um dia típico, o que eu veria?", p. 350). O segundo tipo são perguntas sobre valores e opiniões, no qual o foco sai da ação e vai para o âmbito cognitivo, interpretativo do indivíduo ("No que você acredita?"; "Qual a sua opinião sobre _____?", p. 350). O terceiro tipo são as perguntas sobre sentimentos. Aqui se lança luz sobre as emoções e não sobre os processos cognitivos da pessoa ("Como você se sente com relação a isso?", p. 350). O quarto tipo são perguntas de conhecimento, cujo objetivo é identificar o conhecimento factual do(a) entrevistado(a). O quinto tipo

consiste nas perguntas sensoriais que procuram entender os sentidos do(a) informante, ou seja, aquilo que ele(a) viu, tocou, cheirou ou ouviu ("Quando você passa pelas portas do Programa, o que você vê?", p. 350). Por fim, o sexto tipo são as perguntas que visam entender o contexto e os aspectos demográficos dos indivíduos. Tais perguntas focam em idade, escolaridade e ocupação, por exemplo.

Posto isso, uma etapa central na aplicação de entrevistas é saber como perguntar. Isto é, entender melhor as diferentes técnicas de questionamento. Berry (1991), Boni e Quaresma (2005), Duarte (2005), Gaskell (2003) e Patton (2002) destacam alguns dos principais aspectos com relação às técnicas de questionamento e boas práticas na aplicação de entrevistas, os quais foram sumarizados no quadro 1, a seguir[4].

Quadro 1: Técnicas para Entrevistas Qualitativas.

Técnica	**Descrição**
Antes da entrevista	Alguns passos importantes: Reler o roteiro da entrevista e memorizá-lo Considerar a agenda do(a) entrevistado(a), permitir que escolha local/hora mais conveniente Quando sua pesquisa depender de poucas fontes, checar previamente se elas têm interesse e disponibilidade Caso tenham disponibilidade, deixar as fontes mais relevantes para serem entrevistadas por último, quando mais informações já foram levantadas sobre o tema Identificar e buscar reduzir fatores relacionados ao(à) próprio(a) entrevistador(a) (a sua instituição ou o vocabulário usado, por exemplo) que possam influenciar as respostas obtidas
Início da entrevista	Começar com o agradecimento à disponibilidade e participação do(a) entrevistado(a)
Explicar objetivo da entrevista	Explicar para o(a) entrevistado(a) a natureza ou o objetivo da entrevista, mas de uma forma que não gere viés nas respostas obtidas
Autorização para gravar	Se houver gravação, é preciso avisar à(ao) respondente antes e pedir a sua autorização, abordando a gravação de forma tranquila. Sempre checar se o gravador está funcionando
Utilizar uma sequência de perguntas	Chamada "técnica de afunilamento", na qual se fazem perguntas que vão do geral para o específico
Iniciar com perguntas simples	Entrevistas não devem ser iniciadas pelo tópico de interesse, mas sim com perguntas simples e interessantes que deixem o(a) entrevistado(a) à vontade
Fazer perguntas sobre experiência e/ou comportamento antes das perguntas sobre sentimento e/ou opiniões	Ajuda a dar contexto antes de perguntas mais sensíveis
Fazer perguntas claras	Usar um vocabulário familiar ao contexto do(a) respondente, fazer perguntas curtas e evitar jargões
Perguntar uma coisa de cada vez	Fazer várias perguntas de uma vez só pode gerar confusão no(a) entrevistado(a)
Fazer perguntas realmente abertas	As perguntas devem ser feitas de tal forma que o(a) entrevistado(a) tenha espaço para se expressar como quiser. O(a) entrevistador(a) deve dar pausas para que o(a) entrevistado(a) possa pensar em sua resposta. O silêncio também é importante
Fazer sondagem e perguntas de follow-up	O objetivo é aprofundar as respostas dadas e aumentar a quantidade de informações coletadas

4 Ver Gaskell (2003), p. 83-84, para alguns exemplos práticos de perguntas, e ver Martin (2013), p. 249 - 253, para exemplo de questionário de entrevista semiestruturada.

Interpretar as perguntas	Fazer perguntas que ajudem a esclarecer ou confirmar declarações dadas pelo(a) entrevistado(a). Por exemplo: "Seria correto afirmar que a sua visão sobre este assunto é ..."
*Evitar perguntas sensíveis*5	Tais perguntas podem deixar o(a) respondente desconfortável e afetar as respostas dadas ou até levar ao encerramento abrupto da entrevista
Estimular respostas livres, mas manter o foco no interesse da pesquisa	Permitir que entrevistados(as) tenham a liberdade para sair um pouco do foco da pergunta, mas direcionar a conversa para os temas que são relevantes para a pesquisa
Rapport	Respeitar a opinião do(a) entrevistado(a), demonstrar interesse na sua resposta etc. Isso pode ser demonstrado por meio das expressões, da postura, dos gestos e do tom de voz adotados por quem entrevista. É recomendado que o(a) pesquisador(a) estude estas técnicas antes de aplicar suas entrevistas. Um conselho geral é demonstrar interesse e atenção no que seu(sua) entrevistado(a) diz, ser educado(a), atencioso(a) e utilizar uma linguagem adaptada ao(à) interlocutor(a)
Atenção à forma como entrevistado(a) responde	Considere adaptar suas perguntas a cada entrevistado(a). Esteja atento(a) e disposto(a) à necessidade de esclarecer algumas palavras que podem não ser conhecidas ou ter um significado distinto para eles(elas). Observe também a linguagem não verbal da fonte (expressões, postura etc.)
Uso do tópico guia	Deve apenas guiar e servir como base para monitorar o andamento da entrevista (tempo, temas abordados etc.). É importante memorizar as perguntas e utilizar os tópicos apenas quando necessário. Foco deve ser na escuta e entendimento do que o(a) entrevistado(a) diz
Encerramento da entrevista	Dar indícios de que vai encerrar a entrevista para que esta não seja finalizada abruptamente Finalizar a entrevista com uma nota positiva de agradecimento e garantia de confidencialidade e dar tempo para que entrevistado (a) "saia" da entrevista Perguntar se o(a) respondente gostaria de fazer algum comentário adicional após o gravador ser desligado Explicar como a informação será usada e o andamento da pesquisa, caso necessário

Fonte: Elaboração própria a partir de Berry (1991, p. 3 - 4), Boni e Quaresma (2005), Duarte (2005), Gaskell (2003) e Patton (2002).

Além desses pontos, Berry (1991) também apresenta nove técnicas adicionais identificadas durante a aplicação de entrevistas com um grupo de estudantes. A primeira delas é *contradição*, isto é, apresentar uma opinião contrária àquela expressa pelo(a) entrevistado(a) para estimular mais comentários sobre o tópico. Uma segunda técnica é *conexão*: conectar o comentário do(a) respondente com o tópico que o(a) pesquisador(a) tem interesse. A terceira é o *falso puzzle*: demonstrar não entendimento da resposta apresentada para que a pessoa elabore melhor seu pensamento. Uma quarta técnica é *desafiar*: pedir mais informações como forma de validar o ponto apresentado pelo(a) respondente. A quinta técnica é o *encorajar* por meio de elogios para que a pessoa continue a falar. A sexta técnica consiste em *demonstrar compreensão e dar tempo para elaboração*, isto é, reagir às respostas dadas com compreensão e de forma que a pessoa se sinta livre para tecer comentários adicionais sobre o tópico. Uma sétima técnica é o *reconhecimento* que seria repetir a resposta dada pelo(a) entrevistado(a) para demonstrar atenção no que está sendo dito. A oitava técnica é fazer *perguntas diretas* para obter mais informações. Por fim, a nona técnica observada é *buscar por mais detalhes* a partir de perguntas adicionais.

Ademais, vale lembrar que a forma como as entrevistas são aplicadas pode afetar a sua validade e a sua confiabilidade. Algumas estratégias para aumentar a confiabilidade dos dados coletados consiste na qualidade da gra-

5 Algumas entrevistas abordam um tema que é sensível de maneira geral, como a ansiedade gerada por vivências com crime/situações de violência, por exemplo. Nesse caso, há a sugestão de usar um tipo específico de entrevista, chamada entrevista de narrativa, na qual o(a) respondente narra um determinado evento da sua vida. Ver Hollway e Jefferson (1997) e Crouch e McKenzie (2006) para uma discussão mais aprofundada.

vação da entrevista (GASKELL, 2003; LEGARD et al., 2003; MOSLEY, 2013). Além da boa prática de garantir meios efetivos para realizar a gravação, pedir autorização do entrevistado para gravar e garantir a confidencialidade das suas respostas, é interessante transcrever a entrevista logo após a sua execução. Isso pode ajudar a esclarecer eventuais dúvidas na transcrição, visto que as informações estarão recentes na memória. Além disso, a gravação poderá ser consultada para verificar a confiabilidade e até a validade das respostas. Dado que uma gravação nem sempre é viável, seja por questões de confidencialidade, por potencial risco ao(à) entrevistado(a) ou a pedido da própria fonte, é recomendável realizar a transcrição logo após o encerramento da entrevista. Portanto, é interessante garantir um espaço de tempo razoável entre as entrevistas para que seja possível complementar as notas feitas durante a entrevista (MOSLEY, 2013). Ademais, a transcrição vai muito além do processo de escrever as falas dos(as) informantes. Os momentos de silêncio, os gestos feitos ao longo da entrevista, a postura, o tom de voz, risos e outros aspectos não verbais do(a) respondente também devem compor a transcrição. Esse conjunto de informações ajuda a compreender os sentimentos e as emoções dos(as) entrevistados(as) (BONI; QUARESMA, 2005).

Para além da gravação, ter o domínio do tema tratado é crucial para identificar possíveis inconsistências nas respostas apresentadas. Também é indicado repetir a entrevista com o mesmo indivíduo e o mesmo roteiro, variando quem entrevista. Ainda, é possível fazer entrevistas de acompanhamento, ou seja, repetir a entrevista com o mesmo indivíduo e verificar como sua resposta varia ao longo do tempo. Essa variação pode decorrer de efeitos do(a) entrevistador(a) ou de mudanças que aconteceram no período entre as entrevistas. É salutar destacar que há um *trade-off* nesse acompanhamento: é preciso escolher entre entrevistar poucas pessoas de forma mais intensa ou mais pessoas em menos vezes. A primeira fortalece validade interna, enquanto a segunda é útil para a validade externa. Como sempre, a decisão depende do objetivo da sua pesquisa (MOSLEY, 2013).

Por fim, um bom exercício para o(a) leitor(a) é ler transcrições de entrevistas que estejam disponíveis em repositórios acadêmicos e científicos para compreender melhor como essas técnicas são aplicadas na prática. Um exemplo é o projeto "Mulheres na ciência política", da Associação Brasileira de Ciência Política, no qual foram realizadas 30 entrevistas com pesquisadoras de diferentes áreas de estudo, regiões do país e tempo de carreira, como parte do objetivo da associação de recuperar sua memória institucional e ressaltar as contribuições das cientistas políticas para a área. Todas as transcrições das entrevistas estão publicamente disponíveis[6]. As estudiosas foram entrevistadas por outras 23 colegas da área, as quais se basearam em um mesmo roteiro semiestruturado (BIROLI et al., 2020). Assim, é possível observar como a condução da entrevista varia a depender do(a) entrevistador(a) e de quem está sendo entrevistado(a), algo intrínseco a natureza das entrevistas semiestruturadas.

Compilação e Análise dos Dados

Como já discutido, o roteiro que guia a entrevista permite que seja criada uma estrutura para comparação das respostas apresentadas em cada tópico, o que ajuda a sistematizar as informações extraídas, descrever os resultados obtidos e analisá-los em categorias (DUARTE, 2005). Assim como a construção dos tópicos guia deve ser pautada no problema de pesquisa e no objetivo da investigação, a análise dos dados coletados também deve ser teoricamente orientada (DUARTE, 2005; GASKELL, 2003).

O primeiro passo para essa análise é fazer uma transcrição de qualidade e bem detalhada. A transcrição pode

6 Os *links* para as transcrições podem ser acessados na página a seguir: <https://cienciapolitica.org.br/projetos/projeto-mulheres-ciencia-politica> (BIROLI et al., 2020).

ser transferida para uma tabela na qual as linhas representam cada entrevistado(a) e as colunas são preenchidas com as respostas dadas a cada pergunta/tópico abordado. Esta estrutura permite identificar com maior facilidade padrões de regularidades entre as respostas. Uma vez que seus dados foram coletados, ou seja, gravados, transcritos e sistematizados, com a sua pergunta de pesquisa em mente você pode avaliar se os dados permitem uma avaliação quantitativa ou qualitativa (GASKELL, 2003).

A abordagem utilizada depende do desenho de pesquisa adotado. As tradições quantitativa e qualitativa têm lógicas diferentes no que tange ao peso que dão às variáveis, divergindo quanto ao seu processo causal e as suas observações (*data set observations x causal process observations*). Na análise quantitativa, as observações têm o mesmo peso e o objetivo é identificar padrões e testá-los em relação à hipótese nula. Na análise qualitativa, os(as) pesquisadores(as) são como detetives que buscam por evidências que possam ajudá-los(as) a resolver um determinado quebra-cabeça. Cada pedaço de informação importa e um único dado pode servir para rejeitar uma explicação (*smoking gun evidence*) (GERRING, 2017).

Desta forma, embora as entrevistas semiestruturadas sejam consideradas uma técnica *qualitativa* de *coleta* de dados, dados qualitativos também podem ser quantificados. Porém, uma vez que dados qualitativos são transformados em observações quantitativas, não é possível fazer o processo inverso[7]. Ocorre que tal transformação de quali para quanti significa reduzir informações detalhadas e complexas em dimensões menores que são observadas de forma consistente entre as unidades de análise. Se várias observações do mesmo tipo são geradas, então temos observações quantitativas (GERRING, 2017).

Maglia (2020), por exemplo, a partir das entrevistas com dirigentes das 25 novas legendas partidárias criadas entre 1990 e 2018, identificou três lógicas de criação de novos partidos, segundo a perspectiva das elites: lógica de herança partidária, lógica pragmática e lógica ideológica. Numa ilustração hipotética, a partir dessa classificação e a depender da pergunta de pesquisa, seria possível realizar análises qualitativas com foco nos mecanismos de cada lógica ou transformar essa classificação numa variável categórica, inseri-la em um banco dados e realizar análises quantitativas sobre se partidos com lógicas diferentes de criação estão associados a maior particularismo ou a variações na performance do governo, por exemplo.

Sob uma perspectiva qualitativa, a análise dos dados - isto é, das falas dos(as) entrevistados(as) - se inicia a partir da busca por temas que possuem conteúdo comum e as respectivas funções desses temas (DUARTE, 2005; GASKELL, 2003). Neste processo, o(a) pesquisador(a) precisará dedicar bastante tempo à análise. Talvez seja necessário revisitar as transcrições e gravações e "reviver a entrevista", como explica Gaskell (2003). Durante esta etapa, anotar as ideias que se têm ao longo da análise é um ponto importante. Essas notas, assim como a análise de maneira geral, devem sempre estar orientadas pelos objetivos da pesquisa. Por fim, o(a) pesquisador(a) deve procurar não só por padrões, mas também por contradições dentro de um mesmo tópico (GASKELL, 2003).

Esse é um aspecto crucial já que a forma como o(a) pesquisador(a) sistematiza e interpreta os dados pode afetar a sua validade. Isso pode ocorrer, por exemplo, se os dados forem avaliados a partir de um viés que busca confirmar a teoria analisada. Assim, ao reportar os resultados, caso sejam escolhidas respostas específicas, é preciso justificar porque elas foram usadas. Afinal, seria possível escolher *outliers* que tragam informações favoráveis à teoria a ser testada.

7 Embora seja possível retornar ao material original (gravações e transcrições) para análise qualitativa (GERRING, 2017).

Outro fator que influencia a validade é a precisão da informação obtida. Há a possibilidade do(a) informante responder de forma propositalmente evasiva ou errônea. Neste caso, é possível utilizar informações adicionais ou de entrevistas anteriores para validar as respostas obtidas. Usar triangulação, combinando entrevistas com outros materiais empíricos e comparar os achados com o que a literatura sobre o tema ou evidências anteriores indicam, são possíveis caminhos na busca pela precisão das informações coletadas (MINAYO, 2012; MOSLEY, 2013).

De acordo com Minayo (2012), considerar os seus termos estruturantes é aspecto central da análise qualitativa. Os primeiros quatro termos seriam a experiência, a vivência, o senso comum e a ação e se referem ao objeto de estudo da análise. A experiência é a compreensão do indivíduo sobre seu lugar no mundo. Tal compreensão é afetada pela cultura e se expressa na linguagem do indivíduo. Já a vivência seria resultado de uma reflexão, uma interpretação do indivíduo sobre uma experiência, sendo influenciada por traços pessoais. É possível compartilhar uma mesma experiência, mas com vivências distintas. Essa vivência, entretanto, será afetada pelo coletivo. E aí entra o senso comum, definido por Minayo (2012) como um conjunto de conhecimentos que o indivíduo detém a partir das próprias experiências e vivências e se traduz nas crenças, nas opiniões, na forma como os indivíduos pensam e se relacionam. Por conseguinte, o senso comum é a base do que a análise qualitativa investiga. É a expressão dessas vivências e experiências dos indivíduos que será manifestada nas suas atitudes e na sua linguagem. Por fim, a ação estaria ligada à prática de indivíduos, grupos e instituições para construir ou transformar o mundo em que vivem.

Esses quatro substantivos trazidos por Minayo (2012), ajudam a esclarecer para o(a) leitor(a) os desafios enfrentados na análise qualitativa, dada a complexidade do seu objeto de pesquisa. Os dados coletados em uma entrevista em profundidade são pautados nas percepções, crenças, opiniões e visões de mundo dos indivíduos. Ter isso em mente é crucial para compreender os limites e alcances da análise em si. No argumento de Minayo (2012), essa análise deve ser pautada em dois verbos: compreender e interpretar. Compreender significa se colocar no lugar do outro, considerar sua subjetividade. Além disso, ter em mente que a compreensão do indivíduo sobre sua vida, seu mundo é limitada. Analogamente, a compreensão do(a) pesquisador(a) também o é. As condições e as dificuldades nas quais os(as) pesquisadores(as) chegaram na sua interpretação devem ser apresentadas de forma clara. Toda interpretação possui suas limitações e elas devem ser expostas (MINAYO, 2012).

Numa abordagem quantitativa, a ciência política tem discutido de forma crescente nos últimos anos as possibilidades do texto como dado (*text as data*) (MONROE; SCHRODT, 2008; ROBERTS, 2016). A análise automatizada de texto permitiria estudar fenômenos políticos que se manifestam em formato verbal ou escrito (como discursos em campanhas eleitorais, documentos, debates e proposições legislativas etc.) e que eram antes praticamente impossíveis de investigar pela sua extensão (GRIMMER; STEWART, 2013). Outras vantagens do uso da análise automatizada de texto são: a possibilidade de aplicar diferentes técnicas a despeito do idioma do conteúdo estudado; calcular medidas de incerteza a fim de identificar se há, de fato, diferenças relevantes nos textos e não decorrentes de erro de mensuração ou variação amostral; e diminuir a necessidade de trabalho manual, assim facilitando a replicação dos resultados (IZUMI; MOREIRA, 2018).

Não entrarei em detalhes sobre os diferentes modelos automatizados de análise de texto, já muito bem discutidos em outros estudos[8]. Recomendo que o(a) leitor(a)

8 Ver Grimmer e Stewart (2013), Izumi e Moreira (2018) e as introduções de dois números da *Political Analysis* sobre o tema: Monroe e Schrodt (2008) e Roberts (2016). Para materiais práticos e didáticos, ver Moreira (2019)

interessado(a) no uso de tais técnicas se aprofunde nos pontos de atenção mais relevantes ao emprego deste tipo de análise. Dito isso, é importante ressaltar que a análise automatizada não substitui o exame minucioso que a leitura dos textos é capaz de trazer. Os chamados *text methods* têm, então, o papel de ampliar e potencializar essa leitura cuidadosa (GRIMMER; STEWART, 2013; IZUMI; MOREIRA, 2018). Ademais, a análise automatizada de conteúdo é um método incorreto de linguagem. Significa dizer que, dada a complexidade da linguagem, todos os modelos quantitativos que buscam analisá-la serão errados em alguma medida. Não é possível simplesmente assumir, assim como para qualquer outro método, que os resultados alcançados são válidos. Portanto, é essencial empregar boas práticas de validação (GRIMMER; STEWART, 2013; ROBERTS, 2016). Ainda, vale destacar que não há um método claramente superior para a análise quantitativa de texto. Como sempre, isso depende primordialmente da pergunta de pesquisa abordada e dos dados em si, para os quais certos métodos podem ser mais adequados que outros (GRIMMER; STEWART, 2013; IZUMI; MOREIRA, 2018).

Por fim, um ponto crucial na análise dos dados gerados pelas entrevistas está relacionado ao uso de métodos qualitativos de forma geral e a maneira como eles são reportados (BLEICH; PEKKANEN, 2013). Como já discutido, nas ciências sociais brasileiras houve um movimento de rejeição dos métodos quantitativos por um lado e de uso inadequado dos métodos qualitativos, por outro (SOARES, 2005). Uma aplicação pouco transparente de uma técnica qualitativa pode reforçar uma desconfiança com os seus resultados. Portanto, ao escrever um artigo científico qualitativo, o passo-a-passo do processo da pesquisa deve ser descrito de forma clara e transparente (CRESCENTINI; MAINARDI, 2009). Bleich e Pekkanen (2013) recomendam que as etapas do desenho de amostra das entrevistas sejam reportadas em detalhes (ex.: a definição da população de potenciais entrevistados(as); quantos(as) entrevistados(as) declararam participação etc.). Além disso, Mosley (2013) destaca que apesar dos desafios para transparência e replicabilidade no uso de entrevistas - como a questão da confidencialidade - demais informações técnicas podem ser providas, como o roteiro das entrevistas ou a descrição das técnicas usadas para conduzi-las.

A ausência de transparência nos estudos qualitativos contribui para a visão de que tais análises são pouco rigorosas e de que dados qualitativos não podem ser considerados evidências (CRESCENTINI; MAINARDI, 2009). Por uma questão de escopo e espaço, não aprofundarei a discussão sobre a necessidade e os limites da transparência e replicabilidade nos métodos qualitativos[9]. No entanto, recomendo que o(a) leitor(a) interessado(a) na aplicação de entrevistas ou qualquer outra técnica qualitativa se familiarize com essa discussão *antes* de realizar sua pesquisa, a fim de reproduzir boas práticas em sua análise.

3. Conclusão

O que são entrevistas semiestruturadas? Como elas podem ser usadas na ciência política? Ao longo deste artigo procurei responder a essas duas perguntas a partir de uma perspectiva prática e didática da utilização de entrevistas semiestruturadas. Alguns aspectos frisados no texto merecem ser revisitados. O primeiro deles é que as entrevistas semiestruturadas são uma técnica *qualitativa* utilizada para *coletar dados*. Sendo assim, seu foco

(em português) e Silge e Robinson (2020) (em inglês). O primeiro ainda está em construção, mas traz um material introdutório interessante. O segundo é atualizado constantemente.

9 Para uma visão mais aprofundada sobre o tema ver Elman et al. (2018) e os websites *Qualitative Data Repository* e *Qualitative Transparency Deliberations on behalf of the APSA Section for Qualitative and Multi-Method Research,* especialmente a seção "*Evidence from researcher interactions with human participants*".

recai nas crenças, percepções, atitudes, emoções e assim por diante. Uma vez coletados, esses dados podem ser analisados em uma abordagem qualitativa ou quantitativa. Um outro ponto importante é a interação entre flexibilidade e estrutura desse tipo de entrevista e as consequências que isso tem para a validade e confiabilidade dos dados coletados. É muito importante que o(a) leitor(a) tenha essas limitações em mente aos construir seus instrumentos de pesquisa e ao analisar os dados. Ainda, algo que permeia o uso das técnicas de entrevistas, e de qualquer outra técnica, é o lastro em teoria e o foco no problema de pesquisa. Desde a construção do roteiro à análise dos dados, o debate teórico e a pergunta a ser respondida guiam a pesquisa. Finalmente, por ser uma técnica de interação entre pessoas, há diversas questões éticas associadas ao uso de entrevistas. O(a) pesquisador(a) deve estar atento(a) a todas elas, não só no seu aspecto formal - no que diz respeito à documentação necessária e à eventual aprovação em comitês de pesquisa, mas também a todo o processo de planejamento e execução de entrevistas.

Retomando as discussões iniciadas por Soares (2005), espero ter trazido duas contribuições subjacentes a esse estudo. Em primeiro lugar, técnicas qualitativas são tão relevantes quanto técnicas quantitativas de coleta e análise dos dados. Existem diversas áreas de estudo da ciência política que podem se beneficiar do uso desta técnica. Especialmente quando abordamos objetos de pesquisa ainda poucos explorados na disciplina. Nesse sentido, estudos que utilizam entrevistas em profundidade podem contribuir sobremaneira para o avanço das agendas de pesquisa. Por outro lado, entrevistas semiestruturadas - ou qualquer outra técnica qualitativa - não devem ser utilizadas por simples afinidade ou por rejeição ao uso de técnicas quantitativas. A escolha do método a ser usado deve ser tecnicamente guiada. A ausência de abordagens quantitativas *não* configura o uso de métodos qualitativos. Analogamente, simplesmente perguntar algo a alguém também não significa usar entrevistas em profundidade. Para que resultados válidos e confiáveis sejam alcançados, a técnica em questão precisa ser aplicada de forma rigorosa e transparente.

Dito isso: a partir de agora, como seguir? Se você verificou que entrevistas semiestruturadas são adequadas para a sua pergunta de pesquisa e deseja incluí-la em seu desenho, minha sugestão é simples: procure aprender mais sobre a técnica. Neste artigo, eu trouxe aspectos gerais do uso de entrevistas, mas há diversas questões importantes que, por escopo e espaço, não foram devidamente aprofundadas. Felizmente, existe um bom número de estudos sobre cada uma das etapas de execução de uma entrevista (ex.: como definir a quantidade de entrevistas que devem ser feitas ou como e quais perguntas fazer) (BERRY, 1999; BAKER; EDWARDS, 2012; DWORKIN, 2012; LUCAS, 2012; PATTON, 2002). Como já citado, há também alguns trabalhos nos quais pesquisadores(as) compartilham experiências e aprendizados ao aplicar técnicas de entrevista (BERRY, 1999; MOSLEY, 2013; ROULSTON et al., 2003; SIONEK et al., 2020). Ler esses trabalhos pode ajudar a replicar boas práticas e evitar erros. Acho também interessante ler esse tipo de produção, bem como alguns roteiros de entrevistas, para observar exemplos práticos da aplicação da técnica. Lembre-se, no entanto, que uma das recomendações é testar os instrumentos em campo. Ou seja, você tem a oportunidade de testar um roteiro preliminar e aprender a partir dele.

Com isso, além de apresentar a técnica de entrevistas semiestruturadas e propor diretrizes práticas que facilitem o uso informado do método, a ideia é que

este artigo possa contribuir por meio de um material introdutório para a mudança do paradigma ainda incipiente de treinamento mais formal em métodos qualitativos no Brasil (SOARES, 2005; BARBERIA et al., 2014). Tal cenário também está presente em programas de pós-graduação internacionais (MOSLEY, 2013). Esses casos têm em comum uma lacuna em formação adequada em métodos qualitativos, gerando um ciclo vicioso no qual o uso inadequado das técnicas e a falta de treinamento para melhorar sua aplicação acabam por alimentar uma visão equivocada das técnicas qualitativas como fáceis e pouco rigorosas. A expectativa é que este trabalho possa ajudar a informar estudantes e pesquisadores(as) interessados(as) na aplicação da técnica de entrevistas semiestruturadas, bem como que ele possa, em alguma medida, inspirar outras pesquisas focadas em produzir material pedagógico e promover treinamento mais rigoroso na área.

REFERÊNCIAS BIBLIOGRÁFICAS

BAKER, Sarah Elsie.; EDWARDS, Rosalind. (2012), "How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research", *National Centre for Research Methods Review Paper*.

BARBERIA, Lorena; GODOY, Samuel Ralize de; BARBOZA, Danilo. (2014), "Novas Perspectivas sobre o 'Calcanhar Metodológico': O Ensino de Métodos de Pesquisa em ciência política no Brasil", *Teoria & Sociedade*, vol. 22, no. 2: 156-186.

BATISTA, Mariana. (2017), *A Difusão da Lei de Acesso à Informação nos municípios brasileiros: fatores internos e externos*. Brasília, Fundação Escola Nacional de Administração Pública (Enap).

BATISTA, Mariana.; DOMINGOS, Amanda. (2017), "Mais que boas intenções: Técnicas quantitativas e qualitativas na avaliação de impacto de políticas públicas", *Revista Brasileira de Ciências Sociais*, vol. 32, no. 94.

BECKER, Brittney. (2012), "The Grounded Theory Method and its uses for Political Science", *in* B. Becker; G. Kaufmann (org.), *Series of Paper: Methods of Field Research*.

BERRY, Rita S. Y. (1999), "Collecting data by in-depth interviewing". *British Educational Research Association Annual Conference*, University of Sussex, Brighton, UK. September 2 – 5.

BIROLI, Flávia; GUARNIERI, Fernando; TATAGIBA, Luciana. (2020), *Projeto Mulheres na Ciência Política*. Disponível em < https://cienciapolitica.org.br/projetos/projeto-mulheres-ciencia-politica >. Acessado em 14 out. 2020.

BLEICH; Erik; PEKKANEN, Robert. (2013), "How to report interview data", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

BONI, Valdete; QUARESMA; Sílvia Jurema. (2005), "Aprendendo a entrevistar: como fazer entrevistas em ciências sociais", *Em Tese*, vol. 2, no. 1.

BRADY, Henry E. (2010), "Data-set observations versus causal-process observations: the 2000 U.S. presidential election 2010", *in*: H. Brady; D. Collier. (org.). *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Rowman & Littlefield Publishers.

BROOKS, Sarah M. (2013), "The Ethical Treatment of Human Subjects and the Institutional Review Board Process", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

CACHIA, Moira; MILLWARD, Lynne. (2011), "The telephone medium and semi-structured interviews: a complementary fit", *Qualitative Research in Organizations and Management*, vol. 6, no. 3.

CHARMAZ, Kathy; BELGRAVE, Linda Liska. (2015), "Grounded Theory", *in* G. Ritzer. (org.), *The Blackwell Encyclopedia of Sociology*, John Wiley & Sons, Ltd.

CRESCENTINI, Alberto; MAINARDI, Giuditta C. A. (2009), "Qualitative research articles: guidelines, suggestions and needs", *Journal of Workplace Learning,* vol. 21, no. 5: 431-439.

CROUCH, Mira; MCKENZIE, Heather. (2006), "The logic of small samples in interview-based qualitative research", *Social Science Information*, vol. 45, no. 4: 483-499.

DUARTE, Jorge. (2005), "Entrevista em profundidade". *in* J. Duarte; A. Barros (org.), *Métodos e técnicas de pesquisa em comunicação*, São Paulo: Atlas.

DWORKIN, Shari L. (2012), "Sample Size Policy for Qualitative Studies Using In-Depth Interviews", *Archives of Sexual Behavior*, vol. 41, no. 6.

ELMAN, Colin; KAPISZEWSKI, Diana; LUPIA, Arthur. (2018), "Transparent Social Inquiry: Implications for Political Science", *Annual Review of Political Science*, vol. 21: 29-47.

GASKELL, George. (2003), "Entrevistas Individuais e Grupais", *in* M. W. Bauer; G. Gaskell (org.), *Pesquisa Qualitativa com Texto, Imagem e Som: um manual prático*. Petrópolis, Editora Vozes.

GERRING, John. (2017), "Qualitative Methods", *Annual Review of Political Science*, vol. 20, no. 2.

GLASER, Barney G.; STRAUSS, Anselmo L. (2006 [1976]), *The Discovery of Grounded Theory: Strategies for Qualitative Research*. USA, AldineTransaction.

GLYNN, Adam N.; ICHINO, Nahomi. (2015), "Using Qualitative Information to Improve Causal Inference". *American Journal of Political Science*, vol. 59, no. 4:1055–1071.

GOERTZ, Gary; MAHONEY, James. (2012), *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Oxfordshire, Princeton University Press.

GRIMMER, Justin; STEWART, Brandon. M. (2013), "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts", *Political Analysis*, vol. 21: 267–297.

GUIMARÃES, Natália Cordeiro. (2018), "Profissionais no olho do furacão: o papel das educadoras sociais na implementação da política de abrigamento para mulheres ameaçadas de morte em Pernambuco". *Dissertação de Mestrado. Universidade federal de Pernambuco.*

GUSTAFSSON, Karl; HAGSTROM, Linus. (2018), "What is the Point? Teaching graduate students how to construct political science research puzzles", *European Political Science*, vol. 17: 634-648.

HERTEL, Shareen; SINGER, Mathew M.; LEE VAN COTT, Donna. (2009), "Field Research in Developing Countries: Hitting the Road Running", *PS: Political Science and Politics*, vol. 42, no. 2: 305-309.

HOLLWAY, Wendy; JEFFERSON, Tony. (1997), "Eliciting Narrative Through In-depth Interview", *Qualitative Inquiry*, vol. 3, no. 53.

HORTON, Joanne; MACVE, Richard; STRUYVEN, Geert. (2004), "Qualitative Research: Experiences in Using Semi-Structured Interviews", *in* HUMPHREY, Christopher; LEE, Bill (org.), *The Real Guide to Accounting Research*, Elsevier Science.

IRVINE, Annie. (2011), "Duration, Dominance and Depth in Telephone and Face-to-Face Interviews: A Comparative Exploration", *International Journal of Qualitative Methods*.

IZUMI, Maurício. MOREIRA; Davi. (2018), "O texto como dado: desafios e oportunidades para as ciências sociais", *Revista Brasileira de Informações Bibliográficas em Ciências Sociais - BIB*.

KALLIO, Hanna, PIETILÄ, Anna-Maija; JOHNSON, Martin; KANGASNIEMI. (2016), "Systematic methodological review: developing a framework for a qualitative semi-structured interview guide", *Informing Practice and Policy Worldwide through Research and Scholarship.*

KING, Gary; KEOHANE, Robert O.; VERBA, Sidney. (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, New Jersey, Princeton University Press.

LEGARD, Robin; KEEGAN, Jill; WARD, Kit. (2003), "In-depth Interviews", *in* J. Ritchie; J. Lewis (org.), *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. SAGE Publications, London.

LOTTA, Gabriela; PEREIRA, Guilherme Nunes; BICHIR, Renata Mirandola. (2018), "Policy implementation in the intramunicipal level: the case of the Social Assistance Departments of the city of São Paulo", *Revista Brasileira de Políticas Públicas e Internacionais.*

LUCAS, Samuel R. (2012), "Beyond the existence proof: ontological conditions, epistemological implications, and in-depth interview research", *Qual Quanti*.

LYNCH, Julia. F. (2013), "Aligning Sampling Strategies with Analytic Goals", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

MACDOUGALL, Colin; FUDGE, Elizabeth. (2001), "Planning and Recruiting the Sample for Focus Groups and In-Depth Interviews", *Pearls, Pith and Provocation*.

MAGLIA, Cristiana. (2020), "Novos Partidos de Direita no Brasil (1990-2018): Ideologia, Estrutura Institucional e Mercado Eleitoral", *Tese de Doutorado*, Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação de Ciência Política.

MAHONEY, James. (2010), "After KKV: The New Methodology of Qualitative Research", *World Politics*, vol. 62, no. 1: 120-147.

MARTIN, Cathie Jo. (2013), "Crafting Interviews to Capture Cause and Effect", *in* L. Mosley (org.), *Interview Research in Political Science*, Cornell Univ. Press: Ithaca.

MARSHALL, Martin N. (1996), "Sampling for qualitative research", *Family Practice,* vol. 13, no. 6: 522-525.

MINAYO, Maria Cecília de Souza. (2012), "Análise qualitativa: teoria, passos e fidedignidade", *Ciência e Saúde Coletiva*, vol. 17, no. 3: 621-626.

MINAYO, Maria Cecília de Souza. (2017), "Amostragem e Saturação em Pesquisa Qualitativa: Consensos e Controvérsias", *Revista Pesquisa Qualitativa*, vol. 5, no. 7: 1-12.

MINAYO, Maria Cecília de Souza; GUERRIERO, Iara Coelho Zito. (2014), "Reflexividade como Éthos da Pesquisa Qualitativa", *Ciência e Saúde Coletiva*, vol. 19, no. 4: 1103-1112.

MONROE, Burt L.; SCHRODT, Philip A. (2008), "Introduction to the Special Issue: The Statistical Analysis of Political Text", *Political Analysis*, vol. 16: 351-355.

MOREIRA, Davi. (2019), *Text as Data para Ciências Sociais: guia prático com aplicações*. Disponível em: <https://bookdown.org/davi_moreira/txt4cs/>. Acessado em: 12 Jul. 2020.

MOSLEY, Layna. (2013), *Interview Research in Political Science*. Cornell Univ. Press: Ithaca.

PATTON, Michael Quinn. (2002), *Qualitative Research and Evaluation Methods*. Thousand Oaks, Sage Publications: USA.

QUALITATIVE DATA REPOSITORY. (2020), *The Qualitative Data* Repository. Disponível em <https://qdr.syr.edu/.>. Acessado em 14 Jul. 2020.

QUALITATIVE TRANSPARENCY DELIBERATIONS ON BEHALF OF THE APSA SECTION FOR QUALITATIVE AND MULTI-METHOD RESEARCH. *Home*. Disponível em <https://www.qualtd.net/>. Acessado em: 14 Jul. 2020.

RAGIN, Charles C. (2008), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago and London, University of Chicago Press.

REZENDE, Flavio da Cunha. (2011), "A "Nova Metodologia Qualitativa" e as Condições Essenciais de Demarcação entre Desenhos de Pesquisa na ciência política Comparada.", *Revista Política Hoje*. vol. 20, no. 1.

ROBERTS, Margaret E. (2016), "Introduction to the Virtual Issue: Recent Innovations in Text Analysis for Social Science", *Political Analysis*, vol. 24, no. v10: 1-5.

ROCHA, Camila. (2019), ""Imposto é Roubo!" A Formação de um Contrapúblico ultraliberal e os Protestos Pró-Impeachment de Dilma Rousseff", *Dados*, vol. 62, no. 3: 1-42.

ROULSTON, Kathryn; DEMARRAIS, Kathleen; LEWIS, Jamie B. (2003), "Learning to Interview in the Social Sciences", *Qualitative Inquiry*, vol. 9, no. 4: 643-668.

SILGE, Julia; ROBINSON, David. (2020), *Text Mining with R: A Tiddy Approach*. Disponível em: <https://www.tidytextmining.com/> acessado em 12 jul. 2020.

SIONEK, Luiza; ASSIS, Dafne Thaissa Mineguel; FREITAS, Joanneliese de Lucas. (2020), ""If I had known, I wouldn't have come": Implications and Challenges of Qualitative Interview", *Psicologia em Estudo*, vol. 25.

SOARES, Glaucio. (2005), "O Calcanhar Metodológico da ciência política no Brasil", *Sociologia, Problemas e Práticas*. no. 48: 27-52.

TANSEY, Oisín. (2007), "Process Tracing and Elite Interviewing: A Case for Non-probability Sampling", *PS: Political Science & Politics*, vol. 40, no. 4: 765-772.

VEIGA, Luciana; GONDIM, Sonia Maria Guedes. (2001), "A utilização de métodos qualitativos na ciência política e no Marketing Político", *Opinião Pública*, vol. 7, no. 1.

VOGL, Susanne. (2013), "Telephone Versus Face-to-Face Interviews: Mode Effect on Semistructured Interviews with Children", *Interviewing and Survey Analytic Methods*, vol. 43, no. 1: 133-177.

WENGRAF, Tom. (2001), *Qualitative Research Interviewing,* Thousand Oaks, Sage Publications: USA.

WOOFFITT, Robin; WIDDICOMBE, Sue. (2006), "Interactions in Interviews", *in* P. Drew; G. Reymond; D. Weinberg (org.), *Talk and Interaction in Social Research Methods*. SAGE Publications: London.

From theory to analysis: An introduction to using semi-structured individual interviews in political science

Virgínia Rocha - Universidade Federal de Pernambuco

Abstract

What are semi-structured interviews? How can they be used in political science? In this article, I discuss the main defining aspects of in-depth interviews as qualitative data collection techniques focused on understanding the emotions, beliefs, perceptions, motivations, and attitudes of individuals. Specifically, I deal with the combination of structure and flexibility present in semi-structured interviews and discuss the consequences of this for the validity and reliability of the collected data. Next, I present a practical guide to the application of this method in political science. I discuss the main stages of the use of semi-structured interviews, from theory to data analysis, bringing practical examples and permeating relevant discussions in the area, such as the debate between quantitative and qualitative traditions and the need for transparency in the application of interviews and qualitative research in general. The expectation is to offer introductory content to students and researchers in the area and, thus, contribute to advance in formal training and rigorous use of qualitative techniques, both still incipient in Brazil.

Keywords: semi-structured interviews; individual interviews; in-depth interviews; qualitative methods; data collection.

1. Introduction[1]

"At a minimum, qualitative data are trotted out by way of illustration; at a maximum, qualitative data are essential to causal inference." (GERRING, 2017, p. 29). The importance and benefits of using qualitative data for inferences in political science have been increasing (GOERTZ; MAHONEY, 2012; GERRING, 2017; MOSLEY, 2013). In this context, in-depth interviews may be used in several fields of the discipline. To cite only a few examples, Batista (2017) combines statistical analysis with semi-structured interviews with municipal and federal managers, in a multimethod approach, to understand the greatest incentives and challenges in diffusing the Access to Information Law in Brazilian municipalities. Lotta et al. (2018), when investigating the implementation of the Social Assistance Supervision (SAS) in São Paulo, conducted semi-structured interviews with Supervisors from three different regions to better understand the functioning of the SASs in the city, thus gaining greater knowledge of this social milieu. Maglia (2020) interviewed leaders of the 25 parties created in Brazil between 1990 and 2018 aiming to identify, in the perception of respondents, the motivations for creating these new parties.

Despite its contribution, qualitative methods teaching is still a long way from desirable in Brazil. The situation is analogous with the rigorous use of qualitative techniques of collection and data analysis (SOARES, 2005; BARBERIA et al., 2014). This situation is not specific to Brazil. Mosley (2013) argues that the notion that interview techniques are less useful or relevant may be a consequence of the absence of more formal training in these methods. Despite being broadly used in fields such as sociology, history, anthropology, psychology, among others, political scientists have few methodological references regarding the use of interviews in the discipline.

Moreover, there are epistemological issues associated with the use of qualitative methods in political science. Specialists highlight the "positivist – interpretivist" continuum. The former, although recognizing the complexity of social reality, defend the identification of cause-effect relationships and empirically testable, falsifiable hypotheses. They also consider that both quantitative and qualitative methods are useful in this endeavor. Given social reality, interpretivists argue it would be impossible to generate knowledge without considering the historical context and power relations (MOSLEY, 2013).

The implications of this in the use of interviews is that interpretivists highlight the challenges of creating distance between the knowledge and context of the interviewer from the collection and systematization of data created from interviews. For those in a more purely positivist position, interviews are ways to objectively test or generate falsifiable hypothesis. This does not mean that they disregard issues such as interviewer effects, that is, that the interviewer's behavior or characteristics may affect the answers received, but that interpretivists are more sensible to the nuances of this process than positivists are. Currently, political science is predominantly positivistic, which affects the main challenges faced in the use of interviews in the field (LYNCH, 2013; MOSLEY, 2013).

For this paper, I write from a more positivistic perspective regarding the use of semi-structured interviews, while recognizing the importance of concerns about how the relationship between interviewer and informant may affect the data collected. Therefore, I mobilize here interdisciplinary references that help

1 I would like to thank the peer-reviewers who evaluated this paper for their excellent comments, the colleagues who suggested improvements on prior versions of this study, Palloma Marciano and Rebeca Rocha in particular, and CAPES for the financial support via PhD grant. Of course, any error in this work is entirely my own responsibility.

clarify the complexity of this technique. Still from a positivistic perspective, I understand qualitative and quantitative analyses as complementary. In that sense, since the arguments made by King et al. (1994), several studies have demonstrated the need to understand qualitative and quantitative approaches not from a quantitative paradigm, as the cited authors do, but according to the specificities and distinct goals of each method (BRADY; COLLIER, 2010; GERRING, 2017; GOERTZ; MAHONEY, 2012; MAHONEY, 2010; RAGIN, 2008). This perspective does not imply the absence of scientificity standards in qualitative analysis, but that they differ from those established in the quantitative approach (MINAYO, 2017).

Thus, I present general discussions on in-depth interviews, but with a focus on the semi-structured approach, which is the most frequent interview technique in qualitative analyses (KALLIO et al., 2016). I seek to answer the following questions: *what are semi-structured interviews? How can they be used in political science?* The goal is to bring forth the main aspects of individual semi-structured interviews as qualitative technique of data collection (BERRY, 1999; DUARTE, 2005) and present an introduction to their use. Although there are several works, including in the form of manuals or guides, on how to conduct interviews, they are generally for the social sciences, with a greater emphasis on sociology or anthropology (BONI; QUARESMA, 2005; DUARTE, 2005; GASKELL, 2003; LEGARD et al., 2003; MINAYO, 2012), focused on specific steps of interview execution (BERRY, 1999; CROUCH; MCKENZIE, 2006; DWORKIN, 2012; KALLIO et al., 2016; MARSHALL, 1996) or, when they encompass the technique as a whole, are books that go into detail of each aspect of the method in different chapters (MOSLEY, 2013; WENGRAF, 2001).

Evidently, these studies are important and bring valuable recommendations to the matter at hand, but some material that synthesizes the main overall aspects of the qualitative interview technique with a focus on its application for political science is needed. Content such as this are relevant both to try and fill the clear gap in didactic materials on the topic in the aforementioned field, and to offer technical support for the decision researchers make regarding which method they must use. If, in Brazil, the choice in method to be applied in research has often been defined by an academic's affinity with the methodology or their pre-conceived notions, the idea is that it starts being technically founded (MARSHALL, 1996; REZENDE, 2011). Consequently, the expectation is that this article can contribute so that students and researchers in the field who are unfamiliar with this approach may find here an overall view on its application and gain enough knowledge to i) evaluate the appropriateness of this technique to their research problem and, if it is suitable, ii) have initial references to deepen their knowledge about the technique and use it with the necessary rigor. Thus, I accept the trade-off of dealing with some issues more superficially so that it is possible to cover all essential stages in using interviews – from theory to analysis.

That being said, the paper is organized thusly. Apart from this introduction, in the second section, I present the semi-structured approach *vis-à-vis* other interview techniques, highlighting the main aspects that differentiate it from the others and what are the key challenges for its execution. In the third section, I detail the main stages in using interviews – theory, script/interview guide, interviewee selection, conducting interviews, and consolidation and data analysis. I focus on central discussions for each step in the process and give practical examples from the field that use the technique. In the last and final section, I present final discussions and conclude the paper.

2. Semi-structured interviews: What are they and why do they matter?

Interviews make it possible to ask that which we cannot observe: feelings, intentions, and thoughts, for example. Thus, interviews are conducted to understand others' perspectives, from the assumption that such a perspective is "meaningful, knowable, and able to be known and made explicit" (PATTON, 2002, p. 341). They enable the better understanding of causal mechanisms, may serve as main data source, and may generate data that will be quantitatively analyzed at a later date (LYNCH, 2013; MOSLEY, 2013). In addition, in-depth interviews are appropriate when the goal is a detail description of social environment to understand it, when one wants to use the results to develop a conceptual framework for subsequent research, or to test concepts. That is, using the data collected to test expectations generated outside an established theory (GASKELL, 2003).

There are several interview classifications. A first relevant distinction is among individual and group interviews. Although they share general guidelines, given that both use the semi-structured approach (GASKELL, 2003; HERTEL, 2009; MOSLEY, 2013), they have some specificities. Group interviews or focal groups also have a focus on individuals' behavior, attitudes, and opinions, but under the perspective of consensus construction and reaction to points of divergence. Some scholars argue that group interviews are preferable to individual ones, because they enable the observation of how individuals interact and react to situations of conflict, as well as the possibility of using extra resources (projections, metaphors, etc.) (GASKELL, 2003). However, it is important to weigh the logistical costs that they required and their methodological limitations (HERTEL, 2009).

On the other hand, for individual interviews, the possibilities are diverse: survey, counseling, diary interview, life story, ethnographic, unstructured/informal, and conversation interviews. It is also possible to define interviews by three main approaches: informal conversation interview, general interview guide approach, and standardized open-ended interview (BERRY, 1999; PATTON, 2002). The first type refers to an interview style of continued conversation, also known as unstructured or ethnographic interview, appropriate for a more fluid exploration a topic of interest, being typically used in ethnography and participant observation (DUARTE, 2005; PATTON, 2002). Its results may even assist in constructing semi-structured scripts or questionnaires for later close-ended interviews (DUARTE, 2005). In the second type, the interview is guided by a kind of topic list, with the goal of ensuring that all relevant points are covered. In this way, it is an adequate approach to acquire specific information regarding the object of study. Lastly, in the final approach the questions are more systematically structured and the goal is to reduce the degree of variation in the responses given by the interviewers for each question (BERRY, 1999; GASKELL, ; WOOFFITT; WIDDICOMBE, 2006).

Overall, we then have open-ended, semi-open, or close-ended interviews, with the first two, respectively, being interviews guided by a central theme and a base-script. Both have an in-depth approach, being qualitative. Close-ended interviews are, in their majority, quantitative and have a linear approach (DUARTE, 2005). In the qualitative technique, the order of the questions and words used may vary, whereas in the quantitative, the order and words used matter and must be standardized (WOOFFITT; WIDDICOMBE, 2006). In this article, I will deal with semi-structured interviews, that is, qualitative, in-depth, semi open-ended interviews (DUARTE, 2005).

So, what is, after all, a semi-structured interview? It is a "conversation with purpose" (LEGARD et al., 2003), being a *qualitative methodology of data collection*. A semi-structured interview has the central goal of investigating different perspectives and points of view about a fact, through the interviewees' understanding of reality (GASKELL, 2003). Contrary to quantitative research, which focuses on generalization and average effects (GERRING, 2017; GOERTZ; MAHONEY, 2012), this method focuses on symbols, meanings, beliefs, attitudes, values, and motivations (GASKELL, 2003; LEGARD et al., 2003; DUARTE, 2005).

Beyond this general definition, these interviews may be defined according to some main characteristics (LEGARD et al., 2003). One first aspect is that they present a combination between structure and flexibility: as they follow an interview guide or a script with the main themes to be discussed (structure), the researcher has the freedom to go back to a previous question and go deeper in a certain topic, etc. (flexibility). It is different from a survey structured questionnaire, but it has structure via the interview guide or its script (DUARTE, 2005; LEGARD et al., 2003). Therefore, two other characteristics are highlighted: the interactive nature and the possibility to dive deeper into questions. In the former, the semi-structured interview is the product of the interaction between the researcher and the interviewee. In the latter, the interviewer seeks to deepen the answers received with follow-up questions, that is, questions that ask the interviewee to elaborate on an answer given to a certain topic (LEGARD et al., 2003) (e.g., "could you tell me more about that?" (GASKELL, 2003)).

Some questions are derived from the interactive nature of the technique, as the format of the interview affects the data collected and the effects caused by the researcher themselves in the answers obtained. In the former, it is argued that the interview may not only map pre-existing opinions, but also participate in forming that perception. In other words, it is a method that more than understanding perceptions, also generates knowledge and information. This happens because the individual may be questioned on issues they had never thought of before (CROUCH; MCKENZIE, 2006; LEGARD ET AL., 2003). Therefore, interviews may be considered a process of co-construction of data, insofar as there is a construction of reality and feelings (BERRY, 1999; GASKELL, 2003; ROULSTON ET AL., 2003; SIONEK ET AL., 2020). In the latter, related to the so-called interviewer effects, the behavior of who interviews, their language and nuances, may affect the answers received and, consequently, the data collected. That is why variation in rapport is one of the problems associated with the interactive nature of interviews (MOSLEY, 2013). Thus, recording interviews is essential, given that taking notes throughout the process may affect the data (LEGARD et al., 2003).

The issue of interviewer effects is considered positivistic and it is manifested in two ways. The first is related to how interviewees see who is interviewing and, potentially, change their answers because of it. The second considers the interviewer's biases when interpreting the answers received, which will be influenced not just by the theory, but also by their own context and perceptions. To interpretivists, this is connected to how the interviewer positions themselves in relation to the interviewee, being aware of all these variations (MOSLEY, 2013). Some specialists highlight issues of intersubjectivity and reflexivity. Thus, there is the comprehension that the interviews enable the creation of a material, not just data collection (MINAYO; GUERRIERO, 2014).

It is worth highlighting, also, that semi-structured interviews may be used in a mixed-method approach (DUARTE; 2005; GASKELL; 2003; LYNCH, 2013).

In the same interview, it is possible to use qualitative and quantitative questions (DUARTE, 2005). Moreover, the alternative of using group interviews exists, to observe opinions and processes of consensus and divergence combined with individual interviews to better understand an individual's motivations and particular context (GASKELL, 2003). Consequently, in-depth interviews may be applied in combination with other methods, whether to support a survey design, for example, and provide aid to its interpretation, or to gather contextual information in order to understand the results of a research (*ibidem*).

In political science, in-depth interviews are also commonly used to construct field experiments (GERRING, 2017). Qualitative information can also be used to improve causal inference in designs that use quasi-experimental techniques, such as matching, for instance. They assist in the reduction of measurement errors, enable the construction of a qualitative confidence interval on effect size, and help to handle confounders (GLYNN; ICHINO, 2015). Within qualitative methods, there is also the use of elite interviews in process-tracing designs (TANSEY, 2007).

In addition to these areas, in-depth interviews may be used as qualitative technique for evaluating the impact of public policies (BATISTA; DOMINGOS, 2017), in the field of political marketing (VEIGA; GONDIM, 2001), and in studies on several other topics such as lobbying, violence experienced during war, the preferences of businesspeople regarding welfare policies (MOSLEY, 2013), studies on the implementation of public policies (GUIMARÃES, 2018, LOTTA et al., 2018) and, in addition, the study of political parties in Brazil (MAGLIA, 2020; ROCHA, 2019), to name a few.

In evaluations, interviews of this kind are especially relevant to understand how the impact of a program affects each person differently (BATISTA; DOMINGOS, 2017). That is, understanding the perspectives of participants in a given program, the team that implements the program, and other groups relevant to the execution of a public policy or social program. To do so, the use of this technique sheds light on people's experiences with the program, how they perceive and feel about the actions developed and so on (PATTON, 2002). In that same vein, a study on the implementation of a specific policy focusing on street-level bureaucrats, such as the one by Guimarães (2018), helps with observing different relevant aspects of the perspectives of people directly responsible for delivering that service to beneficiaries. In political marketing, exploring a person's life experiences and routines may assist in better understanding their views on politics. In common is the focus on beliefs, decision, values, expectations, desires, etc. It is an "asymmetrical dialogue" for which the interviewer seeks to collect data and the interviewee is the information source. The focus, then, lies on the interviewee's speech (BATISTA; DOMINGOS, 2017; VEIGA; GONDIM, 2001).

Therefore, in lieu of more immersive studies, such as participant interviews, the application of interviews in political science tend to be more concise, more precise. Insofar that affects scope, it also influences the internal validity of the data collected. On the other hand, there may be external validity gains, through the possibility of contemplating more social groups or places. Thus, there are four main challenges for the application of interviews in political science: ethics, sampling (interviewee selection), validity, and reliability (MOSLEY, 2013).

The ethical issue is associated with the risk that a given study may create to its participants. That risk is considered small when compared to medical research, for example. Nonetheless, it exists and must

be considered in social sciences. Within in-depth interviews, some examples of that risk are sensitive topics that may cause respondents to remember traumatic experiences or the lack of anonymity when reporting the data, putting at risk the safety of the interviewee. In that case, informed consent is key to ensure ethics in studies that involve interactions with human beings. It must communicate to informants the goals of the study, clarify that their participation is voluntary and they are free to leave at any point and will not be punished for it, and present the risks of participating and will be done to mitigate them. Ensuring respondents' confidentiality is also necessary (MOSLEY, 2013).

Often, this consent is made official with a Consent Form, which can be, in some cases, verbal – when the interviewee is illiterate, for example. The construction and signing of this form bring about several challenges to researchers. The first of which is that institutions responsible for approving a research, which in the case of Brazilian universities would be Ethics in Research Committees (CEPs, in Portuguese) and, internationally, the Institutional Review Boards (IRBs), tend to prefer a language that is more formal and technical, which may be less accessible to interviewees. There is criticism that, in some cases, the documents are written in such a way to protect more the institution that funds or is connected with the study than the participant. Given that, it is the role of the researcher to ensure that the Form is interpretable, that is, to explain it to the informant, in an accessible and clear language, what that document is about (BROOKS, 2013; MOSLEY, 2013).

Another aspect is that signing the Form itself might affect an interviewee's motivation to participate in the research. There is indication that verbal consent is associated with greater motivation to participate in a study when compared with written consent. Given that written consent is irrevocable, being a basic condition of respect to the human being that is taking part in a study, the researcher must be careful as to how they present this step to the interviewee, honestly and calmly (BROOKS, 2013; MOSLEY, 2013).

Minayo and Guerriero (2014) warn that qualitative research involves ethical challenges that go beyond the instruments and procedures defined by Ethics Committees – they involve the process as whole. The authors highlight three factors related to ethical issues in research: the researcher's behavior in the field; the way the material is analyzed; and the report made from that material. Thus, a researcher must always be respectful of the individuals they interact with. This is extended to other research stages. Regarding the publication of the data or the research report, the decision of what must be included and exposed in the analysis is also ethical. It is crucial that the researcher is transparent and describes in detail their research process, how relationships were established in the field, so that their conclusions can be evaluated within that context. A practical example of ethical measures that may be taken in result presentation is removal or change of people's and/or institutions' names to fictional ones.

The next three challenges – interviewee selection, validity, and reliability – are directly linked to some of the stages of planning and conduction interviews and, therefore, will be detailed throughout the next section. Overall, the challenge of selecting interviewees is about how to define the population of people that can be interviewed; and how to select individuals from that population, something that depends on the construction of theory and how the hypotheses will be tested. Internal validity is associated with factors such as sample/interviewee selection and how the interview was conducted (if the questions managed to capture information relevant to the research,

if respondents' answers was truthful and, if there are biases present in the answers, if the interviewer was able to detect them). In short, it encompasses all the aspects of the process of planning and conducting interviews and that may affect the data collected as well as its capacity to capture the research's object of study (MINAYO, 2012; MOSLEY, 2013). Finally, the last challenge when conducting interviews is the reliability of the answers received. This is related to how consisted is the instrument of measurement. In the case of interviews, if we can trust that if we were to repeat the same interview with the same individual, we would get similar, consistent answers (MOSLEY, 2013).

3. How should interviews be used? From theory to data analysis

Figure 1 below presents a visual schematic that summarizes the main stages broached here. Each stage will be discussed in a specific subtopic, except for submission of the research to an ethics committee, which was discussed in the previous section. Informed consent is crucial to conduct interviews, but the need to submit the study for approval depends on the institution the researcher is connected to[2]. This stage would, naturally, come before conducting interviews and it appears in grey in figure 1 due to its importance in interview planning.

Figure 1: Stages to use Semi-structured Interviews.



Source: The author, based on Gaskell (2003) and Mosley (2013).

3.1 Theory

The first step in conducting semi-structured interviews must be the development of a theoretical or conceptual framework that will guide the analysis. Thus, it is important that concepts and central themes are previously identified (GASKELL, 2003; CROUCH; MCKENZIE, 2006). Therefore, the object of study must be defined from a research question or problem and situated within a theoretical debate (MINAYO, 2012). This may seem obvious, considering that a research problem must clearly counter the contributions of the proposed analysis in relation to what has already been discussed (GUSTAFSSON; HAGSTROM, 2018). Nevertheless, because it is crucial, this point must be remembered and highlighted.

In addition, another possibility is the use of grounded theory (GASKELL, 2003). Grounded theory may be understood as the discovery of theory from data (GLASER; STRAUSS, 2006 [1976]). This methodology's guidelines to collect and analyze data allows the generation of middle-range theories (CHARMAZ; BELGRAVE, 2015). With an inductive approach, the idea is generating theories that are anchored in data and capable of explaining the phenomenon analyzed. Beyond sociology, grounded theory is often used in anthropology and organizational studies, for example. In political science, there are two points of contention for the use of this method: the debate on qualitative and quantitative methodologies and the emphasis in the deductive approach and confirmation of theories, rather than their creation. Overall, the use of this

2 There are several documents required by Ethics Committees for research with human beings, in a demanding process, given that it is mostly guided by biomedical research (MINAYO; GUERRIERO, 2014). Although some recent resolutions ease this process for social science research, a researcher must plan to collect and deliver the documents required by the Committee of their institution and consider the time of evaluation of their request to conduct the study. Given that this affects the schedule of the analysis itself, it is recommended to check right at the start if it is necessary to submit the study to a CEP, which documents are required, and how long it may take until the research is approved, before setting out to conduct the interviews.

method in political science would be limited, being possibly more suited to the fields of political sociology and public policies (BECKER, 2012).

Once the theoretical framework is defined, it is possible to move on to two essential questions for the method's application: what to ask and who to ask it to. The first issue is connected to the definition of the interview guide, whereas the second is related to the selection of interviewees (GASKELL, 2003). Both are defining characteristics of semi-structured interviews and must not be confused with quantitative techniques for survey construction, much less issues of randomization and representativeness regarding sampling (DUARTE, 2005; GASKELL, 2003).

3.2 Interview guide

The adequate elaboration of the interview guide is crucial for the execution of a good interview. As the name itself says, the topics guide the research and make it possible to reach the goals of the investigation (GASKELL, 2003). The script must be anchored by the central research questions and their theories and hypotheses. These hypotheses may come up as the answers are obtained, which could provoke a change in the interview guides (DUARTE, 2005; MINAYO, 2012). Thus, the researcher must base their actions in theories and hypotheses but keep an open mind to question them (MINAYO, 2012).

In addition to basing themselves in the appropriate theoretical framework, a researcher must understand the field. That is, make and have prior observations and conversations with people that might be relevant to the analysis, discuss the topic with their colleagues and use their creativity (GASKELL, 2003; MINAYO, 2012). When out in the field, keeping a log of things said, accounts, and perspectives presented by people spoken with might aid the researcher in weaving a collective narrative on the lives and experiences of the individuals whose reality is being analyzed (MINAYO, 2012).

The guide facilitates structuring the interview in a way that is possible to ensure that all the same items are discussed with all interviewees. However, the interviewer is free to explore the topics in the way they deem more appropriate (PATTON, 2002). It is the combination between structure and flexibility that Legard et al. (2003) discuss.

When it comes to formulating the interview script itself, it is ideal that it is constructed as a sequence of paragraph titles. Close-ended questions that require mere repetition to interviewees should not be elaborated, replaced by topics for conversation. It is necessary to remember the goal of understanding interviewees' perceptions, world views, beliefs, and attitudes on certain themes (GASKELL, 2003). The script for this kind of interview does not need to be extensive, usually being four to seven questions, and ideally not exceeding one page. Such questions need to be broad enough to allow deep discussion of the topic and must not be redundant. The interviewer must explore the answers presented (with follow-up questions, that aim to go deeper with responses) and only move on to the next question when the topic is exhausted (DUARTE, 2005; GASKELL, 2003).

The scripts, then, tend to be comprised of two levels – one that broaches the central themes of the research, with questions that allow informants to speak freely on the topics, and another that encompasses follow-up questions (KALLIO et al., 2016). It is worth highlighting, however, that many scripts have much more than seven questions, sometimes approximately 15 of them (HORTON et al., 2004; ROCHA, 2019). More important than following a rule on the quantity of questions is ensuring that all relevant aspects to comprehend the research object have been added.

Practically, it is ideal that the interview guide may be used to a certain extent to monitor the interview, should there much a limitation or estimation of time and one is aware of how many topics are left to discuss (GASKELL, 2003; HORTON et al., 2004). In addition, scripts must help to not only guide the interview, but also their transcription later on, serving as a preliminary schematic. However, subtopics are not inflexible, and can be modified for subsequent interviews. It is possible that the interviewer realizes that a given topic did not contribute much to the development of the discussion or even that there was an important topic that was not present in the previous structure. Consequently, it is common that a researcher starts interviews with a script and ends them with another. Nevertheless, any change must always be registered, along with a justification as to why it happened (DUARTE, 2005; GASKELL, 2003, MINAYO, 2012).

Based on a systematic review of the literature, Kallio et al. (2016) identified five steps in the process of elaborating topic guides that, to an extent, summarize what has been discussed. The steps are: i. identify the pre-requisites for application of a semi-structured interview; ii. consider the previous knowledge already built on that topic (theoretical and empirical material); iii. elaborate a preliminary guide; iv. test that version in the field, with three possibilities: internal testing (with members of the research team), experts' evaluations (researchers not on the team), and field testing (with individuals that could be interviewed). And then, based on the information obtained in the breadth and importance of the questions asked, improve the script; and v. lastly, present the final version of the guide.

3.3 Interviewee Selection

With the theoretical framework and the construction of the interview guide, the next stage is selecting the interviewees. As said, this is one of the challenges faced in the use of the interview technique (MARSHALL, 1996; MOSLEY, 2013). In the application of techniques that involve interaction among researchers and individuals analyzed, the concepts of sampling and saturation are central. They are connected to a deeper question, that is, to what extent can these techniques be considered scientific? To what extent the collected speeches in a certain number of interviews reflect the perspective of a given group? (MINAYO, 2017).

Generally, two main paths can be taken to select the interviewees: probabilistic or non-probabilistic (purposive) sampling (LYNCH, 2013). Which one is better? Some arguments defend that the prevalent use of the non-probabilistic approach is based on an impossible ontological condition of homogeneity of the social world (LUCAS, 2012). Many qualitativists acknowledge the limitations of the conclusions reached by non-random samples and that, consequently, qualitative studies have no intention towards generalization. However, Lucas (2012) argues that many investigations within this approach end up with generalizable conclusions, even with the limitations of purposive samples.

The argument is that in-depth interviews lose much of their capacity to generate knowledge when a more theoretically rigorous selection of interviewees is not practiced (LYNCH, 2013; LUCAS, 2012). Although this discussion is very important, it is interesting to observe it framed by the political science debate between the quantitative and qualitative traditions. Qualitative analysis focuses on the singularities and the intensity of a phenomenon, and not in identifying a pattern, as with quantitative analysis.

In a fast distinction, qualitative analyses are expressed in natural language; with a Small-N approach, cases are selected in a non-random, purposive way, and the focus is on specific events. In this case, it is especially important that interviewees selected represent all the

relevant points of view for the research, that is, diversity of sources is essential. Thus, participants would be selected rather than in a probabilistic way, intentionally. Selection would, then, be based on theory and on the potential information these individuals could add to the analysis (MACDOUGALL; FUDGE, 2001; GASKELL, 2003). On the other hand, quantitative analyses use a probabilistic and statistical approach, Large-N, random samples, and focus on generalization (GERRING, 2017). Whereas quantitative analyses seek to identify average causal effect (effects-of-causes), qualitative studies aim to comprehend what caused a given effect (causes-of-effects) (GOERTZ; MAHONEY, 2012). If for quantitative analysis, selecting cases on the dependent variable is unsuitable, for quantitative analysis such a selection may be the most appropriate to understand the causal mechanisms that led to a given event (ibidem).

Thus, recognizing the limitations of using qualitative techniques such as in-depth interviews is essential, but we must be careful not to superimpose a quantitative logic on a qualitative research design. That said, the answer to the question of what the best path to take is, whether probabilistic or international is: it depends. This decision must be guided by the research question and the goal of the study (MARSHALL, 1996; MINAYO, 2017; MOSLEY, 2013; LYNCH, 2013).

As Mosley (2013) explains, if the intention is to generalize from the data collected, later transforming them into data set observations (DSO), that is, quantitative observations, a probabilistic sample might be more appropriate. However, if the research design is qualitative, that is, if the research question seeks to understand why a given event occurred or how it occurred, the data will be analyzed as causal process observations (CPO)[3]. In this case, we do not use sampling, but interviewee selection (BRADY, 2010; GASKELL, 2003; MOSLEY, 2013).

Occasionally, qualitative researchers randomly select the object they are interested in studying, such as businesses or schools, and within those organizations, intentionally select people who could bring the largest wealth of information to the analysis. Therefore, it is not complete randomization. A random sample can be designed in several ways for in-depth interviews. One of the possibilities is to design stratified random samples that can be used to identify the so-called rare cases. However, random samples are not always possible for several reasons (lack of resources, interviewees who decide against participating, so on). In these cases, one option is to combine randomization with the "snowball" technique, which consists of randomly selecting a few individuals and using their personal networks to try and recruit more people (LYNCH, 2013).

In non-random sampling, for which the correct term would be "selection" rather than "sampling" (GASKELL, 2003), the goal is not to interview individuals with the same profile but "variants of a particular social setting" (CROUCH; MCKENZIE, 2006, p. 493). This is a crucial point for validity and reliability of the data collected via interview (DUARTE, 2005). It is also a fundamental question in qualitative analysis overall, in which inference follows from a few intentionally selected cases (GERRING, 2017). In that case, as Minayo (2017) clarifies, the trustworthiness of qualitative analysis is related more with the object's diversity and with the research design itself than with the size of the sample. In her words, "(...) an ideal qualitative sample is one that reflects, in quantity and intensity, the multiple dimensions of a given phenomenon, and seeks the quality of the actions and interactions throughout the whole process". (MINAYO, 2017, p. 10, original in Portuguese).

Within non-probabilistic, four types of interviewee selection may be highlighted: convenience (or acciden-

3 Brady (2010) discusses the complementarity of the qualitative and quantitative approaches based on two definitions of observation, respectively quantitative and qualitative: data set observations (DSO) and causal process observation (CPO). Nonetheless, the expression *data set observations* is broad, encompassing qualitative and quantitative observations, given that an isolated observation of a causal process can also be inserted in a rectangular data set (that is, all values in a line in a rectangular shape). For Gerring (2017), quantitative observations are comparable, while qualitative ones are not.

tal), purposive (or by judgment), snowball, and interstitial contacts. In the first, selection depends on viability, when sources are selected by proximity or availability. They can be very useful when access to potential interviewees is difficult or in the preliminary stage of the research. In the second, based on their judgement, the researcher defines interviewees based on the knowledge they have on the theme or aiming to reach subjective representation. It may be used to ensure that rare or negative cases are a part of their research. The third, called snowball selection, respondent-driven or chain referral, consists of including interviewees as interviews are conducted, based on referrals by other respondents. Thus, throughout the research, new interviewees can be selected, if the need to include them in the study is identified (DUARTE, 2005; LYNCH, 2013). Rocha (2019), for example, when investigating the formation of a new right in Brazil, used the snowball technique when questioning different social groups which would be the people considered by them as the most important to understand the phenomenon studied. In the fourth and last type, interviewees may be people in a line, in a taxi, people who are in run-of-the-mill situations and can spontaneously share information. It is useful to take notes and, if possible, record (with consent) these speeches, which can bring information that would possibly be inaccessible by other means (LYNCH, 2013).

MacDougall and Fudge (2001) lay out three stages for informant selection: prepare, contact, and follow-up. The first stage, *prepare*, begins with describing the sample: what are the places, characteristics, and experiences these interviewees should have, based on the goals of the research? The aim is to segment the setting in relation to the topic being broached. For example, if an in-depth interview is being used to understand individuals' perception of a certain education policy, all groups that may be affected by an intervention in the school setting should be heard. These groups may of students benefitted, teachers, the school's managing and technical staff, and the parents of the students participating, for instance (GASKELL, 2003). Maglia (2020), in her research about the reasons that led to the creation of more right-wing than left-wing parties in Brazil (between 1990 and 2018), interviewed leaders from all 25 political parties created. Guimarães (2018), in her analysis of the role of street-level bureaucrats in a public policy to house women victims of domestic violence in Pernambuco, interviewed workers at different levels (educators, coordinators, directors). The goal is to get diverse views on the issue (DUARTE, 2005).

Using educational policy as reference, so that these groups are identified, one has to know the school environment. Thus, a researcher's first step would be to talk with someone who knows this environment, for instance, a person who works at another school in the education system, and discuss with them how the school dynamics is, who would be the groups affected by the policy and so on. When more than one social setting is involved, the complexity of selection increases (GASKELL, 2003). Some criteria can be used to define the groups that must be included in interviewee selection. A first criteria is sociodemographic, such as age or income bracket. Other possibilities are statistical or taxonomic groups. And there is still the possibility of natural groups, people who share a past or future project, similar interests and values, etc. The central premise is that the source can aggregate relevant, trustworthy, and diverse information on the reality studied (GASKELL, 2003; DUARTE, 2005; DWORKIN, 2012).

Afterwards, it is necessary to find information about the source (list the main groups and sources of information about them), who are the key contacts to reach these people (ones who know the communities) or even *champions*, that is, people who could recruit or help recruit interviewees. The next step in this stage is to look for similar projects to yours, by contacting colleagues or community members and, if there are, learn

from these previous studies on how they interviewed, which connections were made, etc. Lastly, it is important to make alternative selections: going through this process to develop roughly three to four contact lists, if there are problems in contacting the initial selection (MACDOUGALL; FUDGE, 2001).

Guimarães (2018) and Maglia (2020) bring up practical examples of the need to list alternative contacts. In the former, the researcher had an agreement, with a signed form, that the Women's Secretariat in Pernambuco would share the contacts of workers from different shelters in state. Since that agreement was not fulfilled, the author sought other means to contact the workers, using contacts she had herself due to being acquainted with the social setting of shelters. Maglia (2020) initially planned to interview founding presidents of all political parties researched. However, some had passed away and others did not return her attempts to contact them. The author then decided to interview leaders who had participated in the process of creating the party or who, currently, exercised the function of party president.

In the next stage, *contact*, it is essential to have a good relationship with the key contacts, but without depending heavily on those contacts to conduct interviews. During this stage, the step-by-step would be as follows: 1. initial approach with key contact or champion to briefly inform them of the research's goal; 2. negotiate with these contacts, seeking their support; 3. direct negotiation, that is, direct contact with those who will be interviewed; 4. confirmation: define how the approach will happen (e-mail, phone call, etc.); and 5. involvement: if there was an agreement to maintain contact with these people. If so, is there planning and resources for it? Interviewees may be interested in knowing the results of the research (MACDOUGALL; FUDGE, 2001). Thus, it is essential to consider the theme studied, time constraints, and the sources predisposition to speak (DUARTE, 2005). The last stage, *follow-up*, generally consists of informing participants of the research results; defining the means of communication to update on the study's progress, if it is a continuous project; or, still, invite participants to public events about the study and, lastly; define someone responsible for any possible advocacy that results from the project (MACDOUGALL; FUDGE, 2001).

Finally, how many people should be interviewed? Contrary to the quantitative approach, in which the representativeness of the sample in relation to the population is crucial, here a few interviews may be sufficient to obtain an in-depth description of the phenomenon analyzed (DUARTE, 2005). There are a few estimations of quantity that vary between five and fifty interviews (GASKELL, 2003). Dworkin (2012) suggests a number between 25 and 30 participants, which would enable to i) examine certain characteristics connected to research questions and concepts; ii) maximize the data being collected; and iii) increase the probability that negative cases and hypotheses are explored (p. 1320). Baker and Edwards (2012) suggest a broader variation that ranges between 12 to 60 interviewees.

Thus, the ideal number of interviews depends on the scope of the research, on practical issues such as how many interviews is it possible to transcribe in a given timeframe, and how many interviews are needed to obtain useful information for the investigation (BAKER; EDWARDS, 2012; GASKELL, 2003). One alternative is to conduct interviews to the point of saturation, in which more interviews do not add useful information to the research (GASKELL, 2003; MACDOUGALL; FUDGE, 2001). This saturation level must be reported, as well as all the other stages of interviewee selection (BLEICH; PEKKANEN, 2013).

The concept of saturation has been widely debated in the field of qualitative analysis. One of the points debated is that there simply is no cut-off point for a

sample, given that when more information is obtained on an object, it is possible that more questions arise, maintaining the need for more interviews (MINAYO, 2017). The key factor is that this number is justified based on the research object – on its complexity and extent, but also supported by theoretical framework important to the research. It is essential to remember that the previous definition of the instruments, such as interview guides or the number of interviews, is important but not immutable (GASKELL, 2003; MINAYO, 2017). Given this dynamic process, a researcher may identify the need to interview more people given the information obtained in the field (MINAYO, 2017). I suggest that, should a researcher previously define the number of interviews, they justify in their project why that number may go through modifications – both for practical reasons, such as the impossibility of reaching some people, and the theoretical reasons already mentioned. If that change does occur, it must be justified and presented in a transparent way to the reader.

3.4 Conducting Interviews

Once the research instruments have been built – the definition of what and who to ask – and the research proposal has been duly approved according to the necessary ethical criteria, it is time to move on to the practical application of the interviews. The quality of the information collected in an interview especially depends on who interviews (PATTON, 2002). Thus, the interviewer needs to be well-trained to use this technique: ability and technique are essential factors for an interview to be executed with rigor (BATISTA; DOMINGOS, 2017; PATTON, 2002). In this subsection, I will explore the main aspects for the good execution of an in-depth interview.

According to Gaskell (2003), in-depth interviews may last between one and one and a half hours. In the case of phone interviews, although there is some evidence that their time and depth do not vary significantly in relation to in-person interviews (CACHIA; MILLWARD, 2011; IRVINE, 2011; VOGL, 2013), a duration that goes over one hour may become too long in this format. It would be advisable to plan the interview for a shorter length of time.

The beginning of interview should be on lighter themes, it should never be the topic of interest. It is important that the interviewee feel at ease when speaking. A friendly environment should be created, that enables the establishment of a relationship of trust with the researcher (BONI; QUARESMA, 2005). During this, the interviewer's attitude (*rapport*) demonstrating attention and interest in what the interviewee is saying is crucial (DUARTE, 2005; GASKELL, 2003). An aspect that deserves highlighting, and goes beyond the questions, is avoiding acknowledging only answers that confirm your argument (DUARTE, 2005). In the qualitative logic, the goal is always to seek information that *disconfirm or challenge* the theory analyzed (GERRING, 2017). This attitude must be present not only when analyzing and reporting results, but also in the execution of the interviews themselves. Moreover, it is important to highlight that this technique may be conducted not just in person, but also by other means, such as phone or internet (DUARTE, 2005; LEGARD et al., 2003).

According to Patton (2002), there are six types of questions that can be asked in an interview. The first are questions about an interviewee's behavior and experience, to understand the actions and activities undertaken and their experiences in a given context ("If I followed you through a typical day, what would I see you doing?", p. 350). The second type are questions about values and opinions, where the focus moves from action to an individual's cognitive, interpretive area ("What do you believe?; "What

is your opinion of _____?", p. 350). The third type are questions about feelings. Here light is shed on emotions rather than a person's cognitive processes ("How do you feel about that?", p. 350). The fourth type are knowledge questions, with the aim of identifying an interviewee's factual knowledge. The fifth type consists of sensory questions that seek to understand an informant's senses, that is, what they saw, touched, smelled, or heard ("When you walk through the doors of the program, what do you see?", p. 350). Lastly, the sixth type are questions that seek to understand the context and demographic aspects of individuals. These focus on age, schooling, and occupation, for example.

That being said, a key stage in the conduction of interviews is knowing how to ask questions. That is, understanding well the different questioning techniques. Berry (1999), Boni and Quaresma (2005), Duarte (2005), Gaskell (2003), and Patton (2002) highlight some of the main aspects with relation to the techniques for asking questions and good practices in the conduction of interviews, which are summarized in table 1 below[4].

Table 1: Techniques for Qualitative Interviews

Technique	Description
Before the interview	Some important steps: Re-read the interview script and memorize it Consider the interviewee's schedule and allow them to pick the most convenient place/time When your research is dependent on a few sources, check previously if they are interested and available If they are available, leave the most relevant sources to be interviewed last, when more information on the theme has been gathered Identify and seek to reduce factors related to the interviewer (their institution or vocabulary, for instance), that might influence the answers obtained
Beginning of interview	Begin by thanking the availability and participation of the interviewee
Explain the aim of the interview	Explain to the interviewee the nature or objective of the research, in a way that will not generate bias in the answers obtained
Recording authorization	If there is a recording, it is necessary to calmly warn the respondent beforehand and ask for their authorization. Always check that the recorder is working
Use a sequence of questions	Called "funneling technique", in which questions that go from the broad to the specific are asked
Begin with simpler questions	Interviews must not be initiated with the topic of interest, rather with simpler and interesting questions that put the interviewee at ease
Ask questions about experiences and/or behavior before questions about feelings and/or opinions	Helps to give context before asking more sensitive questions
Ask clear-cut questions	Use a vocabular familiar to that of the respondent's, avoid long-winded questions and jargon
Ask one thing at a time	Asking too many questions at once may confuse the interviewee
Ask really open-ended questions	Questions must be asked in such a way that the interviewee is free to express themselves however they want. The interviewer should give breaks so that the interviewee can think about the answer. Silence is also important

4 See Gaskell (2003), p. 83-84, for a few practical examples of questions, and see Martin (2013), p. 249 - 253, for an example of a semi-structured interview questionnaire.

Use probing and follow-up questions	The goal is to deepen the answers given and increase the quantity of information collected
Interpret the questions	Ask questions that help clarify or confirm statements given by the interviewee. For instance: "Would it be correct to state that your view on the matter is..."
Avoid questions5 which are sensitive	These questions may make the respondent uncomfortable and affect the answers given, or even lead to the abrupt closing of the interview
Stimulate free responses, but maintain the focus on the research interest	Allow the interviewees to have the freedom to somewhat diverge from the question, but direct the conversation to the themes that are relevant to the research
Rapport	Respect the interviewee's opinion, demonstrate interest in their answer, etc. This may be demonstrated through expressions, posture, gestures, and tone adopted by the interviewer. It is recommended that the researcher study these techniques before conducting their interviews. A general advice is to demonstrate interest and attention in what their interviewee is saying, be polite, attentive, and use language that is adapted to them
Attention to how the interviewee answers	Consider adapting your questions to each interviewee. Be attentive and willing of the need to clarify a few words that might not be known or have a different meaning to them. Also observe the source's non-verbal language (expressions, posture, etc.)
Using the interview guide	It should only guide and serve as base to monitor the interview's progress (time, subjects discussed, etc.). It is important to memorize the questions and use the guide only when necessary. The focus should be on listening and understanding what the interviewee says
Closing the interview	Give indications that the interview is coming to close so that it is not finalized abruptly Finalize the interview with a positive note of gratitude and assurance of confidentiality, and give time for the interviewee to "step out" of the interview Ask if the respondent would like to make any additional comments after the recorder is turned off Explain how the information will be used and the progress of the research, if necessary

Source: the author, based on Berry (1991, p. 3 - 4), Boni and Quaresma (2005), Duarte (2005), Gaskell (2003), and Patton (2002).

In addition to these, Berry (1991, p. 6) also presents nine techniques identified during interview application with a group of students. The first is *contradiction*, that is, presenting an opinion contrary to that of the one expressed by the interviewee to stimulate more comments on the topic. A second technique is *linking*: connecting the respondent's comment to the topic the researcher is interested in. The third is the *false puzzle*: appearing to not understand the answer presented so that the person further elaborate their thinking. A fourth technique is *challenging*: ask for more information as a way to validate the point being made by the respondent. The fifth technique is *encouraging* through compliments, so that the person continues speaking. The sixth technique consists of *demonstrating comprehension and giving time for elaboration*, that is, reacting to the answers given with an understanding and in a way that the person feels free to make additional comments on the topic. A seventh technique is *recognition*, which would be repeating the answer given by the interviewee to demonstrate attention to what is being said. The eighth technique is asking *direct questions* to obtain more information. Lastly, the ninth technique observed is *seek more details* with additional questions.

Moreover, it is worth remembering that the way interviews are conducted can affect their validity and reliability. One strategy to increase the reliability of the collected data consists of the quality of the interview recording (GASKELL, 2003; LEGARD ET AL., 2003;

5 Some interviews discuss themes that are generally sensitive, such as the anxiety developed by experiences with crime/situations of violence. In that case, it is suggested to use a specific type of interview, called narrative interview, in which the respondent narrated a given event in their lives. See Hollway and Jefferson (1997) and Crouch and McKenzie (2006) for an in-depth discussion.

MOSLEY, 2013). In addition to the good practice of ensuring effective means of recording, asking authorization of the interviewee to record, and ensuring the confidentiality of their answers, it is sensible to transcribe the interview immediately after it is done. This can help clarify eventual doubts in the transcription, given that information will be recent in memory. Also, the recording can be consulted to check reliability and even validity of the answers. Given that a recording is not always viable, whether due to reasons of confidentiality, potential risk to the interviewee, or at the source's request, it is advisable to transcribe immediately after the interview ends. Therefore, it is a good idea to ensure a reasonable amount of time between interviews, so that it is possible to complement the notes taken during it (MOSLEY, 2013).

Beyond this recording, being knowledgeable on the topic discussed is crucial to identify possible inconsistencies in the answers given. Repeating the interview with the same individual and the same script, varying the interviewer, is also suggested. In addition, it is possible to conduct follow-up interviews, that is, repeat the interview with the same individual and check how their answers change over time. This variation may be due to interviewer effects or changes that happened in the meantime between interviews. It is important to highlight that there is a trade-off in this follow-up: it is necessary to choose to interview fewer people more intensely or more people less frequently. The former strengthens internal validity, while the latter is useful for external validity. As always, the decision depends on the goal of the research (MOSLEY, 2013).

Lastly, a good exercise for the reader is to read interview transcriptions that are available in academic and scientific repositories, to better understand how these techniques are applied in practice. One example is the project "Women in political science", by the Brazilian Association of Political Science, for which 30 interviews were conducted with researchers in different fields of study, regions in the country, and career length, as part of the association's goal to reclaim its institutional memory and highlight the contributions of women political scientists to the field. All interview transcriptions are publicly available[6]. The scholars were interviewed by 23 other colleagues in the field, based on the same semi-structured script (BIROLI et al., 2020). Thus, it is possible to observe how the conduction of the interview varies depending on the interviewer and on who is being interviewed, something intrinsic to the nature of semi-structured interviews.

3.5 Data Compilation and Analysis

As discussed, the script that guides the interviews enables the creation of a structure to compare answers given in each topic, which helps systematize the information extracted, describe the results obtained, and analyze them in categories (DUARTE, 2005). As the construction of interview guides must be based on the research problem and the goal of the investigation, the analysis of the data collected must also be theoretically guided (DUARTE, 2005; GASKELL, 2003).

The first step for this analysis is a well-done and detailed transcription. The transcription can be transferred to a table in which lines represent each interviewee and columns are filled in with the responses given to each question/topic broached. This structure enables to identify regularity patterns more easily among answers. Once your data is collected, that is, recorded, transcribed, and systematized, with your research question in mind, you can evaluate if the data allows for a quantitative or qualitative analysis (GASKELL, 2003).

6 The links for the transcriptions can be accessed in the following website: <https://cienciapolitica.org.br/projetos/projeto-mulheres-ciencia-politica> (BIROLI et al., 2020).

The approach used depends on the research design chosen. The quantitative and qualitative traditions have different logics when it comes to the weight they give to variables, diverging on their causal process and observations (data set observations x causal process observations). In quantitative analysis, observations carry the same weight, and the goal is to identify patterns and test them against the null hypothesis. In qualitative analysis, researchers are like detectives searching for evidence that may help them solve a puzzle. Each piece of information matters and a single data point can be enough to reject an explanation (smoking gun evidence) (GERRING, 2017).

Thus, although semi-structured interviews are considered a *qualitative* data *collection* technique, qualitative data can also be quantified. However, once qualitative data are transformed in quantitative observations, it is not possible to reverse the process[7]. A transformation from qualitative to quantitative means reducing detailed and complex information to smaller dimensions that will be observed consistently across the units of analysis. If several observations of the same kind are generated, then we have quantitative observations (GERRING, 2017).

Maglia (2020), for instance, using interviews with leaders of the 25 new political parties created between 1990 and 2018, identified three logics of creating new parties, according to the perspective of elites: party legacy, pragmatic, and ideological logics. In a hypothetical example, from this classification and depending on the research question, it would be possible to conduct qualitative analyses focusing on the mechanisms of each logic, or transform that classification into a categorical variable, add it to a database and conduct quantitative analyses on whether parties with different logics are associated with higher degrees of particularism or variations in government performance.

Under a qualitative perspective, data analysis – that is, of the speech of interviewees – begins at the search for themes that have common content and the respective function of these themes (DUARTE, 2005; GASKELL, 2003). In this process, the researcher will need to dedicate a lot of time to the analysis. It might be necessary to revisit the transcriptions and recordings and "relive the interview", as Gaskell (2003) explains. During this stage, writing down ideas had during the analysis is an important aspect. These notes, as well as the analysis overall, must always be guided by the research goals. Lastly, a researcher must always search not only for patterns, but also contradictions within the same topic (GASKELL, 2003).

This is a crucial aspect, since the way a researcher systematizes and interprets the data may affect their validity. This can occur, for example, if the data are evaluated from a bias seeking to confirm the theory analyzed. Therefore, when reporting results, if specific responses are chosen, it is necessary to justify why they were used. After all, it would be possible to select outliers that bring in information favorable to the theory being tested.

Another factor that influences validity is the precision of the information obtained. There is a possibility that the informant intentionally gives evasive or erroneous answers. In this case, it is possible to use additional information or from prior interviews to validate the answers obtained. Using triangulation, combining interviews with other empirical materials, and comparing the finds with what the literature on the theme or previous evidence indicate, is a possible path in the search for precision of the information collected (MINAYO, 2012; MOSLEY, 2013).

According to Minayo (2012), considering its structuring terms is a central aspect of qualitative research. The first four terms would be experience, understanding,

7 Although it is possible to go back to the original material (recordings and transcriptions) for qualitative analysis (GERRING, 2017).

common sense, and action and refer to the analysis's object of study. Experience is the individual's comprehension of their place in the world. This comprehension is affected by culture and is expressed in the individual's language. Understanding would be the result of a reflection, an interpretation of the individual about an experience, being influenced by personal traits. It is possible to share the same experience, but with different understandings. That understanding, however, will be affected by the collective. That is where common sense comes in, defined by Minayo (2012), as a set of knowledges that the individual possesses from their own experiences and understandings and it is translated in beliefs, opinions, in how they think and relate. Consequently, common sense is the basis of what qualitative analysis investigates. It is the expression of individuals' understandings and experiences which will be manifested in their attitudes and language. Lastly, action would be connected to the practice of individuals, groups, and institutions to build or transform the world they live in.

These four nouns discussed by Minayo (2012) help to clarify to the reader the challenges faced in qualitative analysis, given the complexity of their research object. The data collected in an in-depth interview are based on individuals' perceptions, beliefs, opinions, and worldviews. Having this in mind is crucial to understanding the limitations and possibilities of the analysis itself. In Minayo's argument, that analysis must be founded in two verbs: comprehend and interpret. Comprehend means placing oneself in the place of another, to consider their subjectivity. In addition, keeping in mind that an individual's comprehension about their life, their world, is limited. Similarly, so is the researcher's comprehension. The conditions and difficulties under which researchers reached their interpretation must be clearly presented. Every interpretation has their limitations, and they must be made explicit (MINAYO, 2012).

In a quantitative approach, political science has increased discussions in the last few years about the possibilities of text as data (MONROE; SCHRODT, 2008; ROBERTS, 2016). The automated analysis of text would enable the study of political phenomena that are manifested in verbal or written formats (such as electoral campaign speeches, documents, debates, legislative propositions, etc.) and that were practically impossible to investigate given their size (GRIMMER; STEWART, 2013). Other advantages of the use of automated text analysis are: the possibility of applying different techniques despite the language of the content being studied; calculating uncertainty measures with the aim of identifying if there are, indeed, relevant differences in the texts that are not due to measurement errors or sample variation; and diminishing the need for manual work, thus making the replication of the results easier (IZUMI; MOREIRA, 2018).

I will not go into details on the different automated models of text analysis, already well-discussed in other studies[8]. I recommend that the reader interested in using these techniques dive deeper into the most relevant aspects to employing this kind of analysis. That being said, it is important to highlight that automated analysis does not replace the careful examination of reading the texts. Text methods have, therefore, the role of broadening and potentializing this careful reading (GRIMMER; STEWART, 2013; IZUMI; MOREIRA, 2018). Moreover, the automated content analysis is an incorrect method of language. This means that, given the complexity of language, all the quantitative models that seek to analyze it will be wrong in some measure. It is not possible to simply assume, such as with any other method, that results reached are valid. Therefore, it is essential to employ good validation tech-

8 See Grimmer and Stewart (2013), Izumi and Moreira (2018) and the introductions to two issues of *Political Analysis* on the theme: Monroe and Schrodt (2008) and Roberts (2016). For practical and didactical materials, see Moreira (2019) (in Portuguese) and Silge and Robinson (2020) (in English). The former is still under construction but has interesting introductory material. The latter is constantly updated.

niques (GRIMMER; STEWART, 2013; ROBERTS, 2016). Also, it is worth highlighting that there is no method clearly superior for quantitative text analysis. As always, this primarily depends on the research question and the data themselves, for which certain methods may be more suited than others (GRIMMER; STEWART, 2013; IZUMI; MOREIRA, 2018).

Lastly, a crucial point in the analysis of data generated by interviews is related to the use of qualitative methods overall, and how they are reported (BLEICH; PEKKANEN, 2013). As discussed, in the Brazilian social sciences there was a rejection of quantitative methods on one hand and an inadequate use of qualitative methods on the other (SOARES, 2005). A non-transparent application of a qualitative technique may reinforce a distrust of their results. Therefore, when writing a qualitative scientific paper, the step-by-step process of the research must be described in a clear and transparent way (CRESCENTINI; MAINARDI, 2009). Bleich and Pekkanen (2013) recommend that the stages of the sample design for interviews are reported in detail (e.g., the definition of the population of potential interviewees; how many interviewees participated, etc.). In addition, Mosley (2013) highlights that despite transparency and replicability challenges in the use of interviews – such as the issue of confidentiality – other technical information can be provided, such as the interview script or a description of the techniques used during interviews.

The absence of transparency in qualitative studies contributes to a view that these analyses are not rigorous enough and that qualitative data cannot be considered evidence (CRESCENTINI; MAINARDI, 2009). Due to scope and space, I will not detail the discussion on the need and limits of transparency and replicability of qualitative methods[9]. However, I recommend that the reader interested in the application of interviews or any other qualitative technique familiarizes themselves with this discussion *before* conducting their research, in order to reproduce good practices in their analysis.

4. Conclusions

What are semi-structured interviews? How can they be used in political science? Throughout this paper I have sought to answer these two questions from a practical and didactical perspective on the use of semi-structured interviews. Some aspects emphasized in the text are worth revisiting. The first is that semi-structured interviews are a *qualitative* technique of *data collection*. Thus, they focus on beliefs, perceptions, attitudes, emotions, and so on. Once collected, these data can be analyzed with a qualitative or quantitative approach. Another important point is the interaction between flexibility and structure of this type of interview and the consequences this has for the validity and reliability of the data collected. It is very important that the reader have those limitations in mind when building their research instruments and analyzing the data. Also, something that permeates the use of interview techniques, and another other technique, is the support by theory and focus on the research problem. From script construction to data analysis, the research is guided by the theoretical debate and the question to be answered. Finally, because it is a technique of interaction with people, there are several ethical issues associated with the use of interviews. A researcher must pay attention to all of them, not just their formal aspects – regarding required documentation and the eventual approval in research committees, but also throughout the whole planning and execution process of interviews.

9 For a deeper view on the issue, see Elman et al. (2018) and the websites *Qualitative Data Repository* and *Qualitative Transparency Deliberations on behalf of the APSA Section for Qualitative and Multi-Method Research,* especially the section "Evidence from researcher interactions with human participants".

Going back to the discussions initiated by Soares (2005), I hope to have brought two underlying contributions to this study. First, qualitative techniques are just as relevant as quantitative techniques of collection and data analysis. There are several areas of study within political science that can benefit from using this technique. Especially when we broach research objects still underexplored in the discipline. Thus, studies that use in-depth interviews may especially contribute to advancing research agendas. On the other hand, semi-structured interviews – or any other qualitative technique – must not be used due simply to affinity or rejection of quantitative techniques. The choice in method must be guided technically. The absence of quantitative approaches does *not* mean the use of qualitative methods. Similarly, simply asking someone something also does not mean using in-depth interviews. So that valid and reliable results are achieved, the technique in question must be used rigorously and transparently.

That being said: now, how to proceed? If you have found that semi-structured interviews are adequate for your research question and want to include them in your design, my suggestion is simple: try to learn more about the technique. In this article, I brought general aspects of the use of interviews, but there are several important issues that, due to scope and space, were not duly detailed. Fortunately, there is a decent number of studies on each one the stages of executing an interview (e.g., how to define the number of interviews that must be done or how and which questions to ask) (BERRY, 1999; BAKER; EDWARDS, 2012; DWORKIN, 2012; LUCAS, 2012; PATTON, 2002). As mentioned, there are also some works in which researchers share experiences and what they learned when applying interview techniques (BERRY, 1999; MOSLEY, 2013; ROULSTON et. al., 2003; SIONEK et al., 2020). Reading these works may help to replicate good practices and avoid errors. I also find it interesting to read this kind of publication, as well as some interview scripts, to observe some practical examples of the technique's application. Remember, however, that one of the recommendations is to test the instruments in the field. That is, you have the opportunity to test a preliminary script and learn from it.

With that, in addition to presenting the technique of semi-structured interviews and proposing practical guidelines that facilitate the informed use of the method, the idea is that this paper can contribute as introductory material for a still incipient paradigm shift of more formal training in qualitative methods in Brazil (SOARES; BARBERIA ET AL., 2014). This context is also present in international graduate programs. These cases have in common a gap in the adequate formation on qualitative methods, creating a vicious cycle in which the inadequate use of the techniques and lack of training to improve their application end up feeding a mistaken view of qualitative techniques as easy and not rigorous enough. The expectation is that this paper can help inform students and researchers interested in the application of the semi-structured interview technique, as well as that it can, to some extent, inspire other research focused on producing pedagogical materials and promoting more rigorous training in the area.

BIBLIOGRAPHICAL REFERENCES

BAKER, Sarah Elsie.; EDWARDS, Rosalind. (2012), "How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research", *National Centre for Research Methods Review Paper*.

BARBERIA, Lorena; GODOY, Samuel Ralize de; BARBOZA, Danilo. (2014), "Novas Perspectivas sobre o 'Calcanhar Metodológico': O Ensino de Métodos de Pesquisa em ciência política no Brasil", *Teoria & Sociedade*, vol. 22, no. 2: 156-186.

BATISTA, Mariana. (2017), *A Difusão da Lei de Acesso à Informação nos municípios brasileiros: fatores internos e externos*. Brasília, Fundação Escola Nacional de Administração Pública (Enap).

BATISTA, Mariana.; DOMINGOS, Amanda. (2017), "Mais que boas intenções: Técnicas quantitativas e qualitativas na avaliação de impacto de políticas públicas", *Revista Brasileira de Ciências Sociais*, vol. 32, no. 94.

BECKER, Brittney. (2012), "The Grounded Theory Method and its uses for Political Science", *in* B. Becker; G. Kaufmann (org.), *Series of Paper: Methods of Field Research*.

BERRY, Rita S. Y. (1999), "Collecting data by in-depth interviewing". *British Educational Research Association Annual Conference*, University of Sussex, Brighton, UK. September 2 – 5.

BIROLI, Flávia; GUARNIERI, Fernando; TATAGIBA, Luciana. (2020), *Projeto Mulheres na Ciência Política*. Disponível em < https://cienciapolitica.org.br/projetos/projeto-mulheres-ciencia-politica >. Acessado em 14 out. 2020.

BLEICH; Erik; PEKKANEN, Robert. (2013), "How to report interview data", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

BONI, Valdete; QUARESMA; Sílvia Jurema. (2005), "Aprendendo a entrevistar: como fazer entrevistas em ciências sociais", *Em Tese*, vol. 2, no. 1.

BRADY, Henry E. (2010), "Data-set observations versus causal-process observations: the 2000 U.S. presidential election 2010", *in*: H. Brady; D. Collier. (org.). *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Rowman & Littlefield Publishers.

BROOKS, Sarah M. (2013), "The Ethical Treatment of Human Subjects and the Institutional Review Board Process", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

CACHIA, Moira; MILLWARD, Lynne. (2011), "The telephone medium and semi-structured interviews: a complementary fit", *Qualitative Research in Organizations and Management*, vol. 6, no. 3.

CHARMAZ, Kathy; BELGRAVE, Linda Liska. (2015), "Grounded Theory", *in* G. Ritzer. (org.), *The Blackwell Encyclopedia of Sociology*, John Wiley & Sons, Ltd.

CROUCH, Mira; MCKENZIE, Heather. (2006), "The logic of small samples in interview-based qualitative research", *Social Science Information*, vol. 45, no. 4: 483-499.

DUARTE, Jorge. (2005), "Entrevista em profundidade". *in* J. Duarte; A. Barros (org.), *Métodos e técnicas de pesquisa em comunicação*, São Paulo: Atlas.

DWORKIN, Shari L. (2012), "Sample Size Policy for Qualitative Studies Using In-Depth Interviews", *Archives of Sexual Behavior*, vol. 41, no. 6.

ELMAN, Colin; KAPISZEWSKI, Diana; LUPIA, Arthur. (2018), "Transparent Social Inquiry: Implications for Political Science", *Annual Review of Political Science*, vol. 21: 29-47.

GASKELL, George. (2003), "Entrevistas Individuais e Grupais", *in* M. W. Bauer; G. Gaskell (org.), *Pesquisa Qualitativa com Texto, Imagem e Som: um manual prático*. Petrópolis, Editora Vozes.

GERRING, John. (2017), "Qualitative Methods", *Annual Review of Political Science*, vol. 20, no. 2.

GLASER, Barney G.; STRAUSS, Anselmo L. (2006 [1976]), *The Discovery of Grounded Theory: Strategies for Qualitative Research*. USA, AldineTransaction.

GLYNN, Adam N.; ICHINO, Nahomi. (2015), "Using Qualitative Information to Improve Causal Inference". *American Journal of Political Science*, vol. 59, no. 4:1055–1071.

GOERTZ, Gary; MAHONEY, James. (2012), *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Oxfordshire, Princeton University Press.

GRIMMER, Justin; STEWART, Brandon. M. (2013), "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts", *Political Analysis*, vol. 21: 267–297.

GUIMARÃES, Natália Cordeiro. (2018), "Profissionais no olho do furacão: o papel das educadoras sociais na implementação da política de abrigamento para mulheres ameaçadas de morte em Pernambuco". *Dissertação de Mestrado. Universidade federal de Pernambuco.*

GUSTAFSSON, Karl; HAGSTROM, Linus. (2018), "What is the Point? Teaching graduate students how to construct political science research puzzles", *European Political Science*, vol. 17: 634-648.

HERTEL, Shareen; SINGER, Mathew M.; LEE VAN COTT, Donna. (2009), "Field Research in Developing Countries: Hitting the Road Running", *PS: Political Science and Politics*, vol. 42, no. 2: 305-309.

HOLLWAY, Wendy; JEFFERSON, Tony. (1997), "Eliciting Narrative Through In-depth Interview", *Qualitative Inquiry*, vol. 3, no. 53.

HORTON, Joanne; MACVE, Richard; STRUYVEN, Geert. (2004), "Qualitative Research: Experiences in Using Semi-Structured Interviews", *in* HUMPHREY, Christopher; LEE, Bill (org.), *The Real Guide to Accounting Research*, Elsevier Science.

IRVINE, Annie. (2011), "Duration, Dominance and Depth in Telephone and Face-to-Face Interviews: A Comparative Exploration", *International Journal of Qualitative Methods*.

IZUMI, Maurício. MOREIRA; Davi. (2018), "O texto como dado: desafios e oportunidades para as ciências sociais", *Revista Brasileira de Informações Bibliográficas em Ciências Sociais - BIB*.

KALLIO, Hanna, PIETILÄ, Anna-Maija; JOHNSON, Martin; KANGASNIEMI. (2016), "Systematic methodological review: developing a framework for a qualitative semi-structured interview guide", *Informing Practice and Policy Worldwide through Research and Scholarship.*

KING, Gary; KEOHANE, Robert O.; VERBA, Sidney. (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, New Jersey, Princeton University Press.

LEGARD, Robin; KEEGAN, Jill; WARD, Kit. (2003), "In-depth Interviews", *in* J. Ritchie; J. Lewis (org.), *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. SAGE Publications, London.

LOTTA, Gabriela; PEREIRA, Guilherme Nunes; BICHIR, Renata Mirandola. (2018), "Policy implementation in the intramunicipal level: the case of the Social Assistance Departments of the city of São Paulo", *Revista Brasileira de Políticas Públicas e Internacionais.*

LUCAS, Samuel R. (2012), "Beyond the existence proof: ontological conditions, epistemological implications, and in-depth interview research", *Qual Quanti*.

LYNCH, Julia. F. (2013), "Aligning Sampling Strategies with Analytic Goals", *in* L. Mosley (org.), *Interview Research in Political Science*, Ithaca, Cornell Univ. Press.

MACDOUGALL, Colin; FUDGE, Elizabeth. (2001), "Planning and Recruiting the Sample for Focus Groups and In-Depth Interviews", Pearls, Pith and Provocation.

MAGLIA, Cristiana. (2020), "Novos Partidos de Direita no Brasil (1990-2018): Ideologia, Estrutura Institucional e Mercado Eleitoral", *Tese de Doutorado*, Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação de Ciência Política.

MAHONEY, James. (2010), "After KKV: The New Methodology of Qualitative Research", *World Politics*, vol. 62, no. 1: 120-147.

CRESCENTINI, Alberto; MAINARDI, Giuditta C. A. (2009), "Qualitative research articles: guidelines, suggestions and needs", *Journal of Workplace Learning,* vol. 21, no. 5: 431-439.

MARTIN, Cathie Jo. (2013), "Crafting Interviews to Capture Cause and Effect", *in* L. Mosley (org.), *Interview Research in Political Science*, Cornell Univ. Press: Ithaca.

MARSHALL, Martin N. (1996), "Sampling for qualitative research", *Family Practice,* vol. 13, no. 6: 522-525.

MINAYO, Maria Cecília de Souza. (2012), "Análise qualitativa: teoria, passos e fidedignidade", *Ciência e Saúde Coletiva*, vol. 17, no. 3: 621-626.

MINAYO, Maria Cecília de Souza. (2017), "Amostragem e Saturação em Pesquisa Qualitativa: Consensos e Controvérsias", *Revista Pesquisa Qualitativa*, vol. 5, no. 7: 1-12.

MINAYO, Maria Cecília de Souza; GUERRIERO, Iara Coelho Zito. (2014), "Reflexividade como Éthos da Pesquisa Qualitativa", *Ciência e Saúde Coletiva*, vol. 19, no. 4: 1103-1112.

MONROE, Burt L.; SCHRODT, Philip A. (2008), "Introduction to the Special Issue: The Statistical Analysis of Political Text", *Political Analysis*, vol. 16: 351-355.

MOREIRA, Davi. (2019), *Text as Data para Ciências Sociais: guia prático com aplicações*. Disponível em: <https://bookdown.org/davi_moreira/txt4cs/>. Acessado em: 12 Jul. 2020.

MOSLEY, Layna. (2013), *Interview Research in Political Science*. Cornell Univ. Press: Ithaca.

PATTON, Michael Quinn. (2002), *Qualitative Research and Evaluation Methods*. Thousand Oaks, Sage Publications: USA.

QUALITATIVE DATA REPOSITORY. (2020), *The Qualitative Data* Repository. Disponível em <https://qdr.syr.edu/.>. Acessado em 14 Jul. 2020.

QUALITATIVE TRANSPARENCY DELIBERATIONS ON BEHALF OF THE APSA SECTION FOR QUALITATIVE AND MULTI-METHOD RESEARCH. *Home*. Disponível em <https://www.qualtd.net/>. Acessado em: 14 Jul. 2020.

RAGIN, Charles C. (2008), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago and London, University of Chicago Press.

REZENDE, Flavio da Cunha. (2011), "A "Nova Metodologia Qualitativa" e as Condições Essenciais de Demarcação entre Desenhos de Pesquisa na ciência política Comparada.", *Revista Política Hoje*. vol. 20, no. 1.

ROBERTS, Margaret E. (2016), "Introduction to the Virtual Issue: Recent Innovations in Text Analysis for Social Science", *Political Analysis*, vol. 24, no. v10: 1-5.

ROCHA, Camila. (2019), ""Imposto é Roubo!" A Formação de um Contrapúblico ultraliberal e os Protestos Pró-Impeachment de dilma Rousseff", *Dados*, vol. 62, no. 3: 1-42.

ROULSTON, Kathryn; DEMARRAIS, Kathleen; LEWIS, Jamie B. (2003), "Learning to Interview in the Social Sciences", *Qualitative Inquiry*, vol. 9, no. 4: 643-668.

SILGE, Julia; ROBINSON, David. (2020), *Text Mining with R: A Tidy Approach*. Disponível em: <https://www.tidytextmining.com/> acessado em 12 jul. 2020.

SIONEK, Luiza; ASSIS, Dafne Thaissa Mineguel; FREITAS, Joanneliese de Lucas. (2020), ""If I had known, I wouldn't have come": Implications and Challenges of Qualitative Interview", *Psicologia em Estudo*, vol. 25.

SOARES, Glaucio. (2005), "O Calcanhar Metodológico da ciência política no Brasil", *Sociologia, Problemas e Práticas*. no. 48: 27-52.

TANSEY, Oisín. (2007), "Process Tracing and Elite Interviewing: A Case for Non-probability Sampling", *PS: Political Science & Politics*, vol. 40, no. 4: 765-772.

VEIGA, Luciana; GONDIM, Sonia Maria Guedes. (2001), "A utilização de métodos qualitativos na ciência política e no Marketing Político", *Opinião Pública*, vol. 7, no. 1.

VOGL, Susanne. (2013), "Telephone Versus Face-to-Face Interviews: Mode Effect on Semistructured Interviews with Children", *Interviewing and Survey Analytic Methods*, vol. 43, no. 1: 133-177.

WENGRAF, Tom. (2001), *Qualitative Research Interviewing,* Thousand Oaks, Sage Publications: USA.

WOOFFITT, Robin; WIDDICOMBE, Sue. (2006), "Interactions in Interviews", *in* P. Drew; G. Reymond; D. Weinberg (org.), *Talk and Interaction in Social Research Methods*. SAGE Publications: London.

Ideologia e Partidos no Brasil: reflexão e prática a partir dos programas de governo

Nara Salles - Universidade Estadual do Rio de Janeiro

Resumo

O objetivo deste artigo é discutir a classificação ideológica dos partidos brasileiros a partir de seus programas de governo. Para isso, organizo a exposição a partir de dois eixos. O primeiro, de caráter teórico, debate se e de que maneira esses documentos podem ser utilizados como objetos de mensuração de posições ideológicas, sobretudo a partir das contribuições da teoria da saliência da competição partidária. Nesse âmbito, apresento uma síntese das classificações já elaboradas para o caso brasileiro por meio dos programas de governo. Os limites dessas categorizações dão origem ao segundo eixo da discussão proposta, de caráter metodológico, e se volta para a apresentação de técnicas de análise automatizada de texto, alternativas ao esquema de codificação manual elaborado pelo Manifesto Project/MARPOR. Dando corpo às questões levantadas ao longo do artigo, realizo um exercício de análise de 889 plataformas eleitorais registradas por candidatos a cargos executivos no Brasil nos três níveis de disputa desde 2010. Para isso, utilizo a técnica de escalonamento, wordfish, e concluo que, embora haja casos nas diversas arenas em que a competição programática é organizada ideologicamente, há outras dimensões que podem estruturar essas disputas.

Palavras-chave: Análise automatizada de texto; Ideologia; Partidos; Programas de governo; Wordfish.

Abstract

This article aims to discuss the ideological classification of Brazilian parties, based on their government programs. For this, I organize the exhibition based on two axes. The first, from a theoretical feature, debates whether and how these documents can be used as objects of measurement of ideological positions, especially from the contributions of the salience theory of party competition. In this context, I present a summary of the classifications already developed for the Brazilian case through government programs. The limits of these categorizations give rise to the second axis of discussion, from a methodological feature, and turns to the presentation of different automated text analysis techniques, alternatives to the manual coding scheme developed by the Manifesto Project/MARPOR. Giving substance to the questions raised throughout the article, I carry out an analysis exercise of 889 electoral platforms registered by candidates for executive positions in Brazil at the three levels of dispute since 2010. For this, I use the scaling technique, wordfish, concluding that, although there are cases in the various arenas in which the programmatic competition is organized ideologically, there are other dimensions that can structure these disputes.

Keywords: Automated text analysis; Government programs; Ideology; Parties; Wordfish.

1. Introdução

Esquerda e direita são importantes conceitos e instrumentos classificatórios do mundo político e social. Embora a ideologia tenha passado por uma reavaliação que apontava para o seu enfraquecimento após a queda do muro de Berlim e o surgimento das questões pós-materialistas, muitos países, como o Brasil, viram emergir nos últimos anos a força da (nova) direita, indicando que essas ainda são definições cruciais para a compreensão das agendas e posições políticas. Mas é possível classificar ideologicamente os partidos através dos programas de governo apresentados por seus candidatos?

Para a *teoria da saliência da competição partidária*, as agremiações disputam concedendo ênfases seletivas em diferentes propostas políticas compiladas em seus manifestos ou plataformas eleitorais (BUDGE; FARLIE, 1983; ROBERTSON, 1976). Se a análise de tais documentos permite conhecer essas ênfases, elas são suficientes para classificar os partidos em uma escala esquerda-direita? Quais categorizações foram propostas até agora para o caso brasileiro, utilizando os programas de governo enquanto objetos? O que dizem essas classificações?

O objetivo deste artigo é elaborar e discutir cada uma dessas questões. Neste esforço, colocam-se alguns desafios relacionados à classificação ideológica do ponto de vista da competição programática, isto é, da divergência entre as propostas políticas dos concorrentes. Essas questões não são novas, todavia persistem no debate travado entre disputas eleitorais e ideologia.

Em primeiro lugar, se tomamos esquerda e direita como categorias definidas a priori, sobretudo historicamente, corremos o risco de constituir uma perspectiva enrijecida, que dificulta identificar posições políticas multidimensionais e estratégias pontuais, emergentes de disputas e contextos específicos. Por outro lado, se abandonamos o espectro ideológico enquanto instrumento analítico, podemos incorrer no erro de desconsiderar preferências políticas mais gerais, dificultando a compreensão e previsão de novos debates e fenômenos, tal como a eleição de Bolsonaro, por exemplo.

Da perspectiva metodológica, a classificação ideológica dos partidos brasileiros a partir dos programas de governo tem levado a resultados curiosos, como a quase inexistência de partidos de direita. Uma das explicações para isso pode se relacionar com o método utilizado pelos estudos, a codificação do *Manifesto Project/MARPOR*. Entre uma série de limitações, essa técnica trabalha com categorias fixas de esquerda e direita, aplicáveis às mais de 45 democracias sobre as quais o projeto se volta desde 1979. Ou seja, negligencia as configurações relativas que esses conceitos podem assumir, não apenas no que se refere ao espaço, mas também ao tempo. Como seria possível avançar?

Depois de apresentar e discutir essas questões, proponho um exercício de analisar 889 programas de governo registrados por candidatos a cargos executivos no Brasil desde 2010, nas três diferentes arenas de disputa. Para isso, utilizo a técnica de análise automatizada de texto que atribui posições espaciais aos partidos através das palavras utilizadas em suas plataformas eleitorais, *wordfish* (SLAPIN; PROKSCH, 2008). Meu argumento principal é que nem sempre uma competição programática se organiza ideologicamente e, portanto, isso se converte em um problema empírico de pesquisa.

Neste artigo, investigo se essa dimensão é a que estrutura as disputas para a presidência, governos estaduais e prefeituras das capitais, comparando as posições atribuídas aos partidos com a classificação ideológica apontada pela literatura. Nesta análise, interessa-me não apenas a posição dos partidos em um espectro programático e, possivelmente, ideológico, mas também

a variação ou constância desse arranjo nos diferentes níveis de disputa.

As seções deste artigo estão organizadas da seguinte maneira: na primeira, discuto esquerda e direita sob a ótica da competição, refletindo se a classificação ideológica pode ser construída a partir dos programas de governo. Na segunda, sintetizo a quais resultados têm chegado os estudos que se propuseram a classificar os partidos brasileiros através desses documentos e suas implicações. Na terceira, apresento técnicas alternativas à codificação do *Manifesto Project/MARPOR* para a análise automatizada de texto e justifico o uso do *wordfish*. Na quarta, descrevo e analiso as posições espaciais atribuídas aos partidos em cada disputa no que se refere à dimensão ideológica. Na última, elaboro as considerações finais, destacando os principais pontos para uma agenda de pesquisa que articule ideologia e competição eleitoral no Brasil.

2. Ideologia, competição eleitoral e o lugar dos programas de governo

Antes de iniciar o debate específico sobre como os partidos brasileiros têm sido classificados ideologicamente a partir de seus programas de governo[1], há uma questão imprescindível: é possível mensurar posições ideológicas através desses documentos? Essa pergunta nos leva, diretamente, à natureza dessas plataformas e das disputas nas quais estão inseridas.

Não há democracia sem competição (SCMITTER; KARL, 1991), mas não há uma única maneira de entendê-la. O caminho mais explorado e longevo nos estudos eleitorais investiga as configurações dessas disputas a partir, sobretudo, do número de competidores e suas posições relativas, a fim de compreender seus condicionantes institucionais e sociais (DOWNS, 1957; RIKER, 1982; ROEMER, 2001). Outra abordagem, em parte derivada da anterior, procura identificar o conteúdo da competição, as questões – *issues* – em torno das quais ela se estrutura, quais as preferências políticas dos atores envolvidos e como elas são estrategicamente mobilizadas (BUDGE; FARLIE, 1983; ROBERTSON, 1976).

Não é que ideologia não teria impacto para determinar essas posições relativas nas quais os modelos espaciais estão interessados. O problema é que o clássico modelo downsiano de proximidade levaria ao que denomino de *paradoxo programático* (SALLES, 2020). A ideologia é, em Downs (1957), um atalho informacional para a identificação das preferências dos partidos com menor custo ao eleitorado. Entretanto, porque as legendas têm como objetivo exclusivo a maximização eleitoral, e sabendo que os eleitores tomam a sua decisão a partir da lógica de proximidade, o movimento centrípeto, visando o eleitor mediano, seria inevitável. Em outras palavras, porque a proximidade entre as posições políticas é o que define uma eleição e todos querem vencer, elas teriam incentivos para convergir[2].

Desse ponto de vista, faria pouco sentido investigar posições ideológicas a partir dos programas de governo. Afinal, o movimento centrípeto imposto pelo modelo de proximidade faria que eles convergissem mais do que se diferenciassem. No entanto, a literatura vem apontando que, no mundo real, o modelo downsiano não funciona tão perfeitamente – ao menos não sem a violação de pelo menos um de seus pressupostos. Grofman (2004) sintetiza cada um deles, apresentando aborda-

1 Opto, neste trabalho, pelo termo programas de governo ou plataformas eleitorais em vez de manifestos partidários, pois penso ser mais adequado para a referência a documentos registrados por eleição e não no momento de fundação dos partidos.

2 Não pretendo defender que não há diferença na proposta de Downs (1957) no que se refere a sistemas bipartidários ou multipartidários e a posse ou não de informações completas, pois há. Essas especificidades não invalidam a conclusão de que, grosso modo, ela aponta para a convergência entre os partidos.

gens alternativas. Para este artigo, basta adotar a ideia de que, embora os partidos disputem votos, isso não constitui objetivo exclusivo ou sequer um fim em si mesmo (ALESINA, 1988; CHAPPELL JR.; KEECH, 1986; SARTORI, 1976; WITTMAN, 1983).

Robertson (1976) foi pioneiro ao tratar dessas questões à luz dos manifestos partidários. Inspirado por análises de coberturas jornalísticas, ele classificou os partidos britânicos em termos das saliências concedidas em seus programas de governo entre 1922 e 1974, mensurando a quantidade de vezes que determinada área foi mencionada em relação ao número total de propostas de cada documento. Como resultado, ele apontou uma característica geral: o que orientava os partidos era a ênfase seletiva de diferentes políticas e não um embate direto. Isto é, Robertson (1976) identificou que, em vez de apresentar posições distintas sobre um mesmo tema, os competidores salientavam *issues* que lhes eram mais favoráveis e evitavam mencionar políticas salientes para seus oponentes.

Posteriormente, seus achados foram corroborados, estendidos e sistematizados por Budge e Farlie (1983), dando corpo à *teoria da saliência da competição partidária*, que, basicamente, tem a concepção fundamental de que partidos disputam através da ênfase seletiva de *issues* – o que denomino de competição programática. Além disso, para a *teoria da saliência*, as posições políticas que as agremiações assumem nessas disputas podem ser mensuradas nos programas de governo através das ênfases seletivas concedidas por cada competidor.

Nesse sentido, sim, é possível extrair posicionamentos políticos a partir dos programas de governo (BENOIT; LAVER, 2006; BUDGE; FARLIE, 1983; GRIMMER; STEWART, 2013; KLINGEMANN et al., 2006; LAVER, BENOIT; GARRY, 2003; MONROE; SCHRODT, 2008). No entanto, mensurar posições políticas equivale a mensurar posições ideológicas? Formulada de outra maneira: o que mede a análise de programas de governo? Trata-se, necessariamente, de ideologia?

Para a *teoria da saliência*, a ênfase seletiva de *issues* – e, então, o conteúdo desses documentos – seria determinada por três fatores: (1) o sucesso ou o consenso em torno das políticas; (2) a vinculação partidária das propostas (*issue ownership*); (3) a competitividade da eleição. O primeiro foi apresentado por Robertson (1976), ao argumentar que há *issues* que serão comuns a todos os partidos, pois são consensualmente tidos como relevantes e necessários pela população. Nesse sentido, ele afirma:

> O único momento em que partidos não só não precisam ser diferentes, mas não podem ser, é quando um conjunto de políticas parece estar funcionando [...]. Que é tolice advogar as mesmas políticas que o oponente quando elas fracassam e tolice diferir quando elas obtêm sucesso é uma regra [...] (ROBERTSON, 1976, p.130, *tradução nossa*)[3].

Se o bom desempenho e o apoio popular em torno das políticas atuariam como um incentivo centrípeto sobre as posições partidárias, a ideia de *issue ownership* mostra que os partidos têm agendas específicas (BUDGE; FARLIE, 1983; DE SIO; WEBER, 2014; PETROCIK, 1996; REPASS, 1971). Isso significa que as agremiações se associam prévia, historicamente e a longo prazo com determinados *issues* e que, embora tais posições possam ser moderadas, elas devem apresentar alguma consistência, atendendo de modo responsivo a expectativas de atividades e eleitores mais fiéis (ALDRICH, 1983). No contexto da discussão proposta neste trabalho, *issue ownership* estaria relacionado à ideologia?

Por fim, o último fator relacionado às ênfases concedidas pelos partidos em seus programas de governo:

3 No original: "The only time when parties not only need not to be different. But cannot afford to be, is where one set of policies appears to be working [...]. That it is foolish to advocate the same policies as one's opponent when they fail, and foolish to differ when they succeed is a rule [...]".

o grau de competitividade eleitoral. Tornar um *issue* saliente atrai uma entrada líquida de votos – para determinando competidor e para seu oponente. Logo, não é estratégico mencionar políticas enfatizadas por seus rivais, sob o risco de favorecê-los. *Issues* podem se tornar campo comum de ênfases, porém, quando as eleições são demasiadamente acirradas, em disputas cuja margem de vitória é pequena. Apesar de os partidos preferirem posicionamentos não competitivos,

mais próximos de suas preferências, a posição ocupada por eles dependerá de sua força eleitoral: onde há alta votação, há maior incentivo para posicionamento extremo (BUDGE; FARLIE, 1983; ROBERTSON, 1983).

É fundamental observar, por essa perspectiva, os programas de governo são, por natureza, instrumentos de posicionamentos políticos, mas não apenas isso. Eles são objetos *estratégicos* da competição eleitoral. Porque as preferências são dadas endogenamente – e não de modo determinista como postulava Downs (1957) –, a estrutura de saliências não é rígida, mas depende da avaliação dos líderes sobre quais dimensões são mais salientes para o eleitorado em cada pleito. Isso significa que os posicionamentos políticos expressos nas plataformas podem não ser sempre um retrato da posição ideológica dos partidos. Nesse sentido, os programas podem incluir fatores contextuais da realidade, elementos de natureza pessoal sobre os candidatos e seus adversários, entre outros. É especificamente por essa razão que eles são considerados espaços endógenos de disputa política.

A questão da estratégia traz um primeiro desafio para pensar a ideologia a partir dos programas de governo. Se os *issues* são mobilizados estrategicamente, em virtude do contexto e da agenda de cada disputa, todo e qualquer conteúdo proposto nas plataformas pode ser utilizado para a classificação ideológica dos partidos? Se sim, qual técnica garantiria que mudanças nos parâmetros das categorias esquerda, centro e direita e suas variantes sejam incorporadas na análise? Essas perguntas nos levam ao dilema mais clássico no que se refere à ideologia e textos políticos: posicionamentos podem ser distribuídos em uma única dimensão?

Esse debate não é nada novo e é uma das principais críticas dirigidas ao modelo clássico downsiano, por exemplo. Stokes (1963), em uma das primeiras e mais contundentes críticas ao trabalho do autor, argumentou que o eleitorado distribui suas preferências com relação não a uma, mas a várias questões e, mais do que isso, que elas variam de relevância ao longo do tempo e não podem ser subsumidas em um único eixo, direita-esquerda. Por isso, a já menciona importância de os partidos identificarem e mobilizarem estrategicamente dimensões latentes na disputa: "[...] as habilidades dos líderes políticos, que devem manobrar pelo apoio público em uma democracia, consistem, em parte, em saber quais as dimensões políticas são ou podem tornar-se salientes para o eleitorado através de propaganda adequada" (STOKES, 1963, p. 372, *tradução nossa*)[4].

O problema é que, se a competição não pode ser compreendida de maneira unidimensional, as categorias esquerda e direita assim o são. Se voltamos ao início do uso desses rótulos como sintetizadores de posicionamentos políticos, não é difícil perceber que se trata, por natureza, de categorias unidimensionais e espaciais (BOBBIO, 1995). Afinal, na Revolução Francesa, os que sentavam à esquerda na reunião dos Estados Gerais eram os que tinham afinidade com o igualitarismo e reformas sociais, enquanto os que se sentavam à direita eram mais próximos da aristocracia e do conservadorismo. Todavia, do final do século XVIII até os dias atuais, o que indicam as categorias esquerda e direita? Elas ainda são úteis para a classificação do mundo político e social?

4 No original: "[...] the skills of political leaders who must maneuver for public support in a democracy consist partly in knowing what issues dimensions are salient to the electorate or can be made salient by suitable propaganda".

Se temos como referência a clássica contribuição de Bobbio (1995), o conteúdo dessa categorização gira, principalmente, em torno da questão da igualdade. Enquanto a direita reconheceria a desigualdade como natural, limitando a ação política, a esquerda se orientaria pelo objetivo de superar tal ordem. Considerando os gastos sociais, por exemplo, governos de esquerda prefeririam medidas redistributivas, enquanto os de centro e direita teriam maior afinidade com políticas de ajuste (ALLAN; SCRUGGS, 2004). Esse binarismo, todavia, possui mediações, relacionadas a variáveis como liberdade, tradição, hierarquia, participação, ordem e autonomia (BOBBIO, 1995). Não seriam essas mediações novas dimensões do espaço político, relacionadas – porém não reduzidas – à ideologia?

Benoit e Laver (2012) parecem convencidos de que sim: "mesmo essa noção drasticamente simplificada de uma 'dimensão esquerda-direita' refere-se a duas questões potencialmente separáveis: uma diz respeito à política econômica e uma segunda se refere à regulação estatal do comportamento social" (BENOIT; LAVER, 2012, p. 195, *tradução nossa*)[5]. Contudo, eles também parecem acreditar que, se na teoria existem mais dimensões potenciais de distinção entre as preferências e posições políticas do que a ideológica, na prática elas podem estar correlacionadas. Isso significa que conhecer o posicionamento de alguém sobre determinado *issue* pode ser útil para prever sua preferência com relação a outros temas. Por isso, do ponto de vista analítico, surge a seguinte pergunta: como utilizar as categorias esquerda e direita sem reduzir significativamente o espaço político, mas, também, sem acrescentar tantas dimensões de modo que torne impossível apreendê-lo? Essa questão se soma a outras, como: o significado de esquerda e direta pode ser estabelecidos a priori? Essas definições variam de acordo com o contexto e a época de cada disputa?

Essa discussão é vital, do ponto de vista teórico e metodológico, e se relaciona diretamente à classificação ideológica atribuída aos partidos brasileiros por meio da análise dos programas de governo. Como demonstro a seguir, o uso, mesmo adaptado, de uma codificação preestabelecida há mais de 30 anos e fixa para a análise de qualquer democracia tem levado a resultados intrigantes. Se os programas de governo, então, podem ser utilizados para a mensuração de posições ideológicas, essa classificação não é garantida pelo simples uso dos objetos, mas exige, também, o emprego de técnicas adequadas.

3. Programas de governo e a classificação ideológica dos partidos no Brasil

Esses desafios que permeiam a obtenção de posições ideológicas através dos programas de governo se estendem aos poucos estudos que se debruçaram sobre isso até o momento no caso brasileiro. Entre eles, destacam-se, sobretudo, as contribuições de Tarouco (2011) e Tarouco e Madeira (2013) pelo maior volume, variedade e abrangência dos documentos analisados. Mais recentemente, somaram-se a eles, novos estudos que incorporaram outras eleições presidenciais (JORGE *et al.*, 2018) se propuseram a debater a questão da polarização (MADEIRA; VIEIRA; TAROUCO, 2017) ou aventaram abordagens comparativas entre programas de diferentes países na América Latina (BABIRESKI, 2014; OLIVEIRA, 2011). Entre eles, há algo em comum: a prevalência quase absoluta dos partidos brasileiros na esquerda do espectro ideológico (Quadro 1).

5 No original: "Even this drastically oversimplified notion of a 'left-right dimension' refers to two potentially separable issues: one concerns economic policy, a second concerns state regulation of social behavior".

Quadro 1 – Classificação ideológica dos partidos brasileiros pela codificação MARPOR

Referência	Direita	Esquerda
Lehmann *et al.* (2019b)	PSL 2018 (+19,8) PSDB 1989 (+19,8)	PT 1989 (-35,6) PSOL 2006 (-35,4) PT 1998 (-26,9) PV 2010 (-26,5) PT 2018 (-25,4) PRONA 1994 (-24,8) PT 2010 (-24,2) PPS 1998 (-24,1) PDT 2018 (-22,5) PT 2002 (-19,7) PT 1994 (-19,6) PT 2014 (-18,5) PT 2006 (-18,3) PSDB 1998 (-17,2) PSDB 2010 (-15,7) PPS 2002 (-15,7) PSDB 2002 (-15,4) PSDB 2006 (-13,3) PSB 2002 (-11,7) PSDB 2014 (-11,0) PSB 2014 (-11,0) PSDB 2018 (-2,7) PRN 1989 (-2,6) PSDB 1994 (-1,8)
Jorge *et al.* (2018)		PT 2014 (-0,13) PSDB 2014 (-0,06)
Babireski (2014)		PP (-15,2) DEM (-11,2)
Tarouco e Madeira (2013)	PFL 1995 (+6,5)	PDT 1979 (-37,3) PTB 1979 (-36,8) PT 1980 (-34,4) PTB 2001 (-27,2) PMDB 1981 (-25,0) PT 1990 (-22,0) PDS 1979 (-15,1) PP 2003 (-14,3) PPB 1995 (-14,1) PSDB 1988 (-13,6) PDT 1994 (-10,3) PMDB 1994 (-8,5) PFL fund. (-5,4) PFL 2005 (-3,6) PSDB 2001 (-0,4)

Fonte: elaboração própria.

As classificações dos partidos brasileiros podem causar certo desconforto. Por elas, praticamente não haveria legendas de direita no Brasil. Assim, mesmo agremiações consideradas consensualmente de direita, como PP, PFL/DEM e PRONA se encontram na esquerda do espectro ideológico. Por outro lado, partidos comumente localizados mais próximos ao centro, como PMDB e PSDB, possuem escores muito semelhantes ao do PT, por exemplo. Esses resultados podem indicar que o conteúdo dos documentos analisados versava sobre outros fatores que não ideológicos – ao menos, não de forma majoritária ou determinante – ou alguma insuficiência metodológica, já que o emprego da codificação da codificação do *Manifesto Project/MARPOR* como técnica[6], ainda que adaptada, é uma característica comum desses estudos.

Na prática, a mensuração é feita através da codificação manual de *quasi-sentences*, ou trechos que contenham apenas uma proposta política, em uma das 56 categorias, distribuídas em 7 domínios (Quadro 2). Algumas delas, como é possível observar, contém subcategorias e direções positiva ou negativa da proposta, porém não é uma regra.

Não há dúvidas de que o esforço do *MARPOR* é fundamental em muitos sentidos. Primeiramente, como já mencionado, pelo seu pressuposto fundador, que postula que competição partidária se estrutura a partir da ênfase seletiva de *issues*. Em segundo lugar, metodologicamente, por apontar que as posições políticas podem ser mensuradas nos programas de governo, a partir de tais ênfases. Por fim, com relação à perspectiva comparada, pelo fato de constituir um amplo banco de dados das plataformas eleitorais em diversos países e períodos.

Todavia, é preciso considerar as limitações da codificação proposta, que tem levado a classificações ideológicas deficientes, senão esdrúxulas, dos partidos brasileiros (MADEIRA; TAROUCO, 2013). Uma primeira questão é o fato de que, embora a categorização de *quasi-sentences* deva ser exclusiva, o esquema oferecido não é exaustivo, o que torna a classificação demasiadamente dependente da avaliação de cada especialista. Isso não constitui, per se, uma falha técnica, visto que pode ser contornado, por exemplo, com a classificação de mais de um profissional. Além disso, não há métodos puramente objetivos, eles sempre envolvem algum tipo de escolha. Desse ponto de vista, essa possível arbitrariedade deve ser considerada apenas como uma limitação mais evidente – e, neste sentido, positiva porque clara – do que escolhas envolvidas em outras técnicas.

Há, no entanto, uma considerável rigidez no esquema utilizado pelo *MARPOR*. Por um lado, isso garante a comparação entre diversos países, com ampla abrangência longitudinal. Mas, por outro, faz desconsiderar uma série de fatores como os contextos locais e elementos mais conjunturais de determinados períodos. Pragmaticamente, a consequência disso é que o pressuposto que caracteriza direita e esquerda na Europa se estende para todo e qualquer outro lugar, em qualquer época.

6 Fundado em 1979 com o objetivo de mensurar posições políticas a partir de programas de governo, o hoje denominado *Manifesto Research on Political Representation (MARPOR)* analisa documentos de mais de 61 democracias, 769 eleições e 1.197 partidos, incluindo os brasileiros. Disponível em <https://manifesto-project.wzb.eu/>. Último acesso em 25/06/2020.

Quadro 2 – Categorias de codificação dos programas de governo do MARPOR

Domínio 1: Relações Externas
101 Relações externas especiais: Positivo
102 Relações externas especiais: Negativo
103 Anti-imperialismo: Positivo
  103.1 Anti-imperialismo estatal
  103.2 Influência externa financeira
104 Forças Armadas: Positivo
105 Forças Armadas: Negativo
106 Paz
107 Internacionalismo: Positivo
108 Integração europeia/latino-americana: Positivo
109 Internacionalismo: Negativo
110 Integração europeia/latino-americana: Negativo

Domínio 2: Liberdade e Democracia
201 Liberdade e Direitos Humanos: Positivo
  201.1 Liberdade
  201.2 Direitos Humanos
202 Democracia
  202.1 Geral: Positiva
  202.2 Geral: Negativa
  202.3 Democracia representativa: Positiva
  202.4 Democracia direta: Positiva
203 Constitucionalismo: Positivo
204 Constitucionalismo: Negativo

Domínio 3: Sistema Político
301 Descentralização: Positivo
302 Centralização: Positivo
303 Eficiência governamental e administrativa: positivo
304 Corrupção política: Negativo
305 Autoridade política: Positivo
  305.1 Autoridade política: competência partidária
  305.2 Autoridade política: competência pessoal
  305.3 Autoridade política: governo forte
  305.4 Elites pré-democráticas: Positivo
  305.5 Elites pré-democráticas: Negativo
  305.6 Reabilitação e compensação

Domínio 4: Economia
401 Economia de livre mercado: Positivo
402 Incentivos: Positivo
403 Regulação do mercado: Positivo
404 Planejamento econômico: Positivo
405 Corporativismo: Positivo
406 Protecionismo: Positivo
407 Protecionismo: Negativo
408 Objetivos econômicos
409 Gerenciamento keynesiano de demanda: Positivo
410 Crescimento econômico
411 Tecnologia e infraestrutura: Positivo
412 Controle da economia: Positivo
413 Nacionalismo: Positivo
414 Ortodoxia econômica: Positivo
415 Análise marxista: Positivo
416 Anticrescimento econômico: Positivo
  416.1 Anticrescimento econômico: Positivo
  416.2 Sustentabilidade: Positivo

Domínio 5: Bem-estar e Qualidade de vida
501 Proteção ambiental
502 Cultura: Positivo
503 Igualdade: Positivo
504 Expansão do Estado de Bem-Estar
505 Limitação do Estado de Bem-Estar
506 Expansão da educação
507 Limitação da educação

Domínio 6: Tecido da Sociedade
601 Estilo de vida nacional: Positivo
  601.1 Geral
  601.2 Imigração: Negativo
602 Estilo de vida nacional: Negativo
  602.1 Geral
  602.2 Imigração: Negativo
603 Moralidade tradicional: Positivo
604 Moralidade tradicional: Negativo
605 Lei e ordem
  605.1 Lei e ordem: Positivo
  605.2 Lei e ordem: Negativo
606 Espírito cívico: Positivo
  606.1 Geral
  606.2 Ativismo bottom-up
607 Multiculturalismo: Positivo
  607.1 Geral
  607.2 Integração de imigrantes: Diversidade
  607.3 Direitos indígenas: Positivo
608 Multiculturalismo: Negativo
  608.1 Geral
  608.2 Integração de imigrantes: Assimilação
  608.3 Direitos indígenas: Negativo

Domínio 7: Grupos Sociais
701 Classes trabalhadoras: Positivo
702 Classes trabalhadoras: Negativo
703 Agricultura e agricultores
  703.1 Agricultura e agricultores: Positivo
  703.2 Agricultura e agricultores: Negativo
704 Classe média e grupos profissionais: Positivo
705 Grupos minoritários: Positivo
706 Grupos demográficos não-econômicos: Positivo

000 Não se aplica

Fonte: VOLKENS et al. (2016).

Não é verdade que a codificação do *MARPOR* leva, necessariamente, à constituição de um espaço político unidimensional, visto que o projeto possui cinco diferentes dimensões programáticas: (1) *planeco*, que mede a intervenção do Estado na economia; (2) *markeco*, que, mensura a liberalização econômica; (3) *welfare*, relacionado às menções sobre o Estado de Bem-Estar Social; (4) *intpeace*, voltado para a paz e relações internacionais; (5) *rile*, a escala esquerda-direita. Ou seja, embora tenha sido aplicado no Brasil apenas a partir do espectro ideológico, seria possível agregá-lo a outras dimensões.

Isso, no entanto, não resolveria o problema do que o MARPOR substantivamente considera como esquerda e direita e a sua aplicabilidade a diversos países e períodos. A escala *rile,* construída com a partir de definições de Laver e Budge (1992), não inclui todas as categorias do esquema. Através dela, a posição ideológica de um partido é dada pela soma dos percentuais de ênfases de direita menos os percentuais de ênfases das categorias de esquerda (Quadro 3).

Quadro 3 – Escala rile MARPOR

Direita	**Esquerda**
104 Forças Armadas: Positivo	103 Anti-imperialismo: Positivo
201 Liberdade e Direitos Humanos	105 Forças Armadas: Negativo
203 Constitucionalismo: Positivo	106 Paz: Positivo
305 Autoridade política	107 Internacionalismo: Positivo
401 Economia de livre mercado	202 Democracia
402 Incentivos: Positivo	403 Regulação do mercado: Positivo
407 Protecionismo: Negativo	404 Planejamento econômico: Positivo
414 Ortodoxia econômica: Positivo	406 Protecionismo: Positivo
505 Limitação do Estado de Bem-Estar	412 Controle da economia: Positivo
601 Estilo de vida nacional: Positivo	413 Nacionalização: Positivo
603 Moralidade tradicional: Positivo	504 Expansão do Estado de Bem-Estar
605 Lei e ordem	506 Expansão da educação
606 Espírito cívico: Positivo	701 Classes trabalhadoras: Positivo

Fonte: VOLKENS et al. (2016).

É importante observar que, nessa concepção, a direita inclui quase todos os domínios, com exceção dos grupos sociais, privilegiando categorias relacionadas à economia e ao tecido da sociedade. A esquerda, porém, não abrange ênfases dessa última dimensão, desconsiderando temas-chave do comportamento social, comumente utilizados para a classificação ideológica, como a legalização do aborto, legalização da maconha, entre outros. Mais importante do que isso, a esquerda não inclui a categoria que, para Bobbio (1995), é a que a define – a igualdade. A categorização, que não é isenta de controvérsias, também exclui a questão do multiculturalismo e considera propostas relativas a Direitos Humanos como *issues* de direita.

Portanto, os problemas em utilizar a escala *rile* para classificar os partidos brasileiros não se restringe ao fato de que o esquema de codificação *MARPOR* unifica os conceitos de esquerda e direta entre países e períodos diversos. Há, também, uma limitação de caráter substantivo no que se refere às categorias tomadas como referências para a mensuração do posicionamento ideológico. Embora a literatura reconheça essas inadequações e proponha adaptações, é essa a escala que tem sido utilizada para classificar ideologicamente os partidos por meio dos programas de governo no Brasil. Além disso, ela tem sido adotada como dimensão programática exclusiva. Nesse sentido, a já mencionada ausência de partidos brasileiros na direita pode estar associada a essas limitações.

Em resumo, não é possível aplicar a codificação ideológica proposta pelo MARPOR ao caso brasileiro – ao menos, não sem alterações. Essa conclusão foi apontada por Tarouco e Madeira (2013), que relembraram que, no Brasil, não seria exatamente a questão da igualdade que dividiria esquerda e direita, mas sim os meios para alcançá-la. Para a direita, isso de daria através do reforço da ordem via Estado, enquanto, para a esquerda, seria por meio da demanda de políticas estatais distributivas, simultânea à contestação da repressão estatal aos movimentos sociais e políticos igualitaristas. Nesse sentido, eles reforçam a necessidade de adequação da escala ao Brasil, considerando tais ambiguidades no que se refere às expectativas relativas ao papel do Estado.

> A escala do MRG, apresentada acima, inclui algumas categorias que não se aplicam ao sentido da distinção esquerda e direita no Brasil. Por conta da experiência histórica da ditadura militar e da transição, a esquerda incorpora reinvindicações que o MRG identifica como de direta: liberdade, direitos humanos e constitucionalismo, por exemplo. Da mesma forma, as defesas da paz como meta geral e do internacionalismo, que nunca chegaram a constituir uma bandeira das esquerdas brasileiras, constam, naquela escala, entre as categorias indicativas de posicionamento à esquerda (TAROUCO; MADEIRA, 2013, p. 158).

Além disso, os autores acrescentam que a defesa de democracia, identificada como de esquerda, é comum a todos os partidos – ou pelo menos era em 2013 – e que categorias indicadas pelos *MARPOR* como de direita, a exemplo de nacionalismo, moralidade tradicional e lei e ordem estão mais relacionadas ao conservadorismo.

A identificação desse desajuste entre a escala *rile* e do que, de fato, distinguiria a esquerda da direita no Brasil deram origem adaptações dessa categorização. Tarouco e Madeira (2013) consideraram que, dentro do esquema de codificação do *MARPOR*, as categorias indicativas da direita seriam: (1) Forças Armadas: Positivo; (2) Economia de livre mercado; (3) Incentivos; (4) Ortodoxia econômica; (5) Limitação do Estado de Bem-Estar; (6) Classe média e grupos profissionais. A posição final dos partidos na escala foi dada pelo total de menções nessas categorias menos o total obtido nas categorias consideradas de esquerda: (1) Regulação do mercado; (2) Planejamento econômico; (3) Controle da economia; (4) Análise marxista; (5) Expansão do Estado de Bem-Estar; (5) Classes trabalhadoras: Positivo. Essa nova medida, no entanto, não resultou classificação muito distinta da anterior: apenas o programa do PFL, de 2005 e de 1995, e do PSDB, de 2001, foram categorizados como de direita (TAROUCO; MADEIRA, 2013)[7].

Oliveira (2014), por sua vez, propôs uma classificação ideológica dos partidos brasileiros através da análise dos programas de governo das eleições presidenciais em 1998 e 2014, mobilizando variáveis do Banco de Asserções Programáticas (BAP). A processo de codificação é semelhante ao esquema do *MARPOR*, porém com diferentes categorias e distintos agrupamentos de variáveis. O autor também estruturou sua proposta em sete grandes eixos: (1) Política Social; (2) Política Comercial; (3) Política Fiscal; (4) Política de Gasto Público; (5) Política Financeira; (6) Política Monetária; (7) Política de Regulação[8]. A posição ideológica dos partidos foi dada pela frequência de asserções nos documentos analisados. Nesse estudo, o Brasil também possuía campo de direita restrito.

Em estudo anterior, analisei 3.410 plataformas eleitorais, registradas por candidatos a prefeito nas eleições de 2016 em 1.335 municípios brasileiros (SALLES; GUARNIERI, 2019). Nele, evidenciei que os candidatos competem programaticamente; isto é, apresentam propostas distintas de seus oponentes. No entanto,

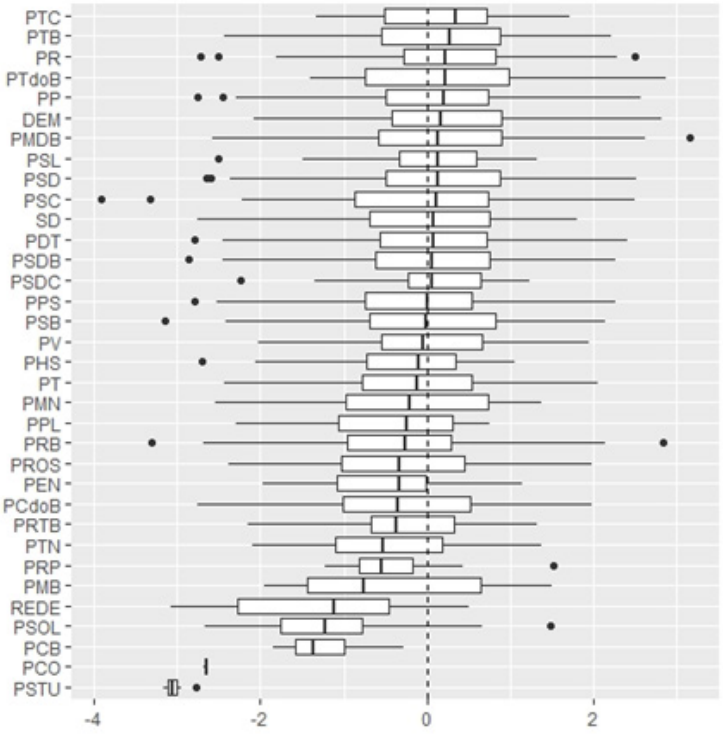
7 A classificação completa, entre a extrema esquerda (-100) e a extrema direita (+100), é a seguinte: PDT 1979 (-23,4), PTB 1979 (-22,9), PTB 2001 (-20,2), PT 1980 (-13,8), PT 1990 (-12,4), PDT 1994 (-12,0), PDS 1979 (-7,2), PP 2003 (-5,2), PPB 1995 (-5,2), PFL fund (-4,9), PSDB 1988 (-3,3), PMDB 1994 (-1,3), PMDB 1994 (-1,3), PFL 2005 (+6,0), PSDB 2001 (+6,5), PFL 1995 (+9,8) (TAROUCO; MADEIRA, 2013).

8 Para conferir cada categoria indicativa de esquerda e direita nesses eixos, consultar Oliveira (2014, p. 76).

salientei que ideologia não consiste em condição para a existência de uma competição em torno de *issues*, de modo que fatores ideológicos são programáticos, mas esses nem sempre são ideológicos.

Apesar disso, o posicionamento espacial dos partidos a partir da frequência de palavras utilizadas em seus programas de governo não permite desconsiderar a variável ideologia (Gráfico 1).

Gráfico 1 – Distribuição de propostas por partido no total de municípios



Fonte: SALLES; GUARNIERI, 2019, p.17.

Considerando a mediana da posição dos partidos no conjunto de municípios da amostra, teríamos a seguinte classificação: PSTU, PCO, PCB, PSOL, REDE, PMB, PRP, PTN, PRTB, PCdoB, PEN, PROS, PRB, PPL, PMN e PT, na esquerda; PV, PSB, PPS, PSDC, PSDB, PDT e SD, no centro; PSC, PSD, PSL, PMDB, DEM, PP, PTdoB, PR, PTB, PTC, na direita. Obviamente, há inconsistências nessa possível classificação, como, por exemplo, o PRB e o PRTB na esquerda e o PDT no centro. Todavia, ela não exclui o campo da direita do espectro programático brasileiro.

Apesar disso, é preciso notar que, embora a mediana esteja posicionada em um dos polos espaciais, não há concentração da distribuição das posições de um mesmo partido na maior parte dos casos, com exceção do PRP, PMB, REDE, PSOL, PCB, PCO e PSTU. Isso pode se relacionar com o fato de a comparação ter sido estabelecida no total de municípios desconsiderando qualquer outra varável sociodemográfica, econômica ou mesmo política da organização e da força dos partidos nessas localidades.

Em resumo, todos esses apontamentos mostram que, embora alguns estudos tenham se voltado para a análise dos programas de governo, ainda há questões em aberto no que se refere à ideologia: as plataformas trazem poucos elementos ideológicos ou estamos encontrando dificuldades para mensurá-los?

4. Os programas de governo enquanto dado

Há diversas técnicas de análise de texto que possibilitam a análise dos programas de governo, visando ou não a classificação ideológica. O primeiro deles, já amplamente mencionado, é a codificação manual, com destaque pelo esquema de categorias e domínios elaborados pelo *MARPOR*. Pelos motivos já explicitados, me dedico a apresentar agora técnicas alternativas e automatizadas de análise de texto.

É fundamental ter em mente os quatro princípios da análise quantitativa de texto elencadas por Grimmer e Stewart (2013). (1) Modelos automatizados podem ser considerados errados em sua totalidade, porém úteis. Isso porque nem sempre são capazes de mensurar as estruturas complexas da linguagem e seu sentido não literal, o que exige que sejam avaliados com relação a utilidade da tarefa científica e social a que se propõem. (2) Esses métodos não são capazes de substituir, sobremaneira, as habilidades humanas, mas ampliá-las. Nesse sentido, permanece a tarefa do pesquisador em guiar o processo e tomar as decisões necessárias para a aplicação dos modelos, além de interpretar seus resultados. (3) Não há um método globalmente mais eficiente para a análise de texto automatizada. Diferentes problemas e conjuntos de dados conduzem a diferentes objetivos, que podem incluir: identificação de palavras que distinguem a linguagem do grupo (LAVER, BENOIT; GARRY, 2003; MONROE; COLARESI; QUINN, 2008); proporção de documentos que se enquadra a categorias predeterminadas (HOPKINS; KING, 2010); posições espaciais dos atores obtidas através dos textos (LAVER; BENOIT; GARRY, 2003; SLAPIN; PROKSCH, 2008). (4) Qualquer que seja o modelo adotado pelo pesquisador é fundamental que seja devidamente validado.

As técnicas de análise automatizada de texto podem ser divididas em dois grupos: as de categorização e as de escalonamento (GRIMMER; STEWART, 2013). As primeiras classificam os documentos em categorias, que podem ser pré-definidas ou latentes. Isto é, pode-se criar dicionários e, a partir disso, codificar os documentos, em uma espécie de técnica MARPOR automatizada; ou é possível deixar que essas categorias sejam extraídas automaticamente dos documentos, no que se chama *topic models* ou modelos de tópicos[9]. Já as técnicas de escalonamento oferecem posições espaciais mensuradas a partir da análise dos documentos. Porque estou interessada em uma classificação ideológica que, como já apresentado, é intrinsecamente espacial, opto, neste artigo, por um modelo de escalonamento.

Embora Grimmer e Stewart (2013) tenham garantido que não há método globalmente eficiente para a análise automatizada de texto, Lowe (2016) argumenta que existe essencialmente apenas uma maneira de escalonar dados textuais[10]. Para ele, a lógica comum para o dimensionamento de itens contáveis incorpora, quer de maneira explícita ou implícita, o fundamento da *teoria da saliência* (BUDGE; FARLIE, 1983). Especificamente, isso significa que todos os métodos de escalonamento político a partir de dados textuais refletem a ideia de que as posições contidas nos documentos são reconstruções de baixa dimensionalidade, delimitadas a partir de padrões de *ênfases seletivas* (LOWE, 2016).

No que se refere à análise automatizada de texto, há duas técnicas principais para escalonamento: *wordscores* (LAVER; BENOIT; GARRY, 2003) e o *wordfish* (SLAPIN; PROKSCH, 2008). A primeira, é supervisionada, ou seja, exige um conjunto de textos de referência *(training set)* e, por isso, foi desconsiderada neste trabalho. Afinal, não

9 Para aprofundar o conhecimento sobre cada um dos métodos disponíveis de análise automatizada de texto, conferir Izumi e Moreira (2018).

10 Para ver outras técnicas além do *wordscores* e do *wordfish*, consultar Lowe (2016).

é consensual o conteúdo substantivo de esquerda e direita. Aqui, emprego o *wordfish*, que atribui posições espaciais de maneira não supervisionada, a partir da frequência de palavras nos documentos. No modelo, os autores assumem que os atores se localizam em um espaço político de baixa dimensão e que as posições ali estabelecidas impactam na taxa em que as palavras são utilizadas em seus textos. Ele funciona da seguinte maneira: admite-se que a distribuição de cada palavra p de um documento i (y_{ip}) origina-se de um processo de Poisson – ou seja, possui apenas um parâmetro, que representa simultaneamente a média e a variância –,

$$y_{ip} Poisson(\lambda_{ip})\, \lambda_{ip} = exp(\alpha_I + \psi_j + \beta_j * \omega_I)$$

em que λ_{ip} é a contagem da palavra p no manifesto do partido i; α é um conjunto de efeitos fixos por documento; ψ é um conjunto de efeitos fixos por palavras; β é uma estimativa do peso específico de uma palavra, capturando a importância da palavra p na discriminação entre posições partidárias; ω é a estimativa da posição política a partir do documento i. No modelo, o efeito fixo por palavras controla termos mais usados do que outros, bem como o efeito fixo por documento equilibra programas maiores do que outros.

Como qualquer modelo, o *wordfish* possui suas limitações. A principal é a unidimensionalidade. Contudo, sua aplicação se adequa ao *corpus* deste estudo, já que consiste em documentos nos quais candidatos e partidos se posicionam sobre uma variada gama de assuntos. Outro problema é que a análise produzida se baseia, sobretudo, na frequência de palavras. Dessa forma, não se pode assumir resultados no que se refere aos sentidos dos textos. Como bem destacaram Grimmer e Stewart (2013), ainda fogem do escopo das análises automatizadas o controle de ironia e, neste trabalho, dos sinônimos, por exemplo. Logo, se um candidato propõe *construção de moradias* e outro a *construção de casas*, as últimas palavras serão apontadas como divergências, mesmo possuindo o mesmo sentido[11]. Além disso, é fundamental salientar que se trata de um modelo comparativo, de modo que só é atribuída determinada posição a um candidato a partir da frequência das palavras que ele utiliza em relação a seus concorrentes.

Neste artigo, utilizo o *wordfish* para analisar 889 programas de governo registrados por candidatos a cargos executivos no Brasil de 2010 a 2018[12] nos três níveis de disputa. Embora tenha elaborado que a competição programática nem sempre se organiza ideologicamente, minha hipótese é que está relaciona com o nível de disputa. Não porque algumas são mais visíveis, ou politicamente "corretas", ou institucionalizada, mas porque a natureza da competição estrutura, de algum modo, a natureza das políticas propostas. Com isso, quero dizer que é mais difícil para um prefeito articular ideologicamente seu programa de governo enquanto deve propor ações muito específicas para as cidades. Eles não possuem, por exemplo, competência para implementar grandes políticas mais ou menos liberais ou conservadoras. Essas agendas mais amplas, no entanto, encontram mais espaço nas esferas superiores, sobretudo nas disputas presidenciais. Por isso, espero que a posição espacial dos programas de governo dos candidatos à Presidência se distribua em uma escala esquerda-direita, e que essa dimensão dê espaço a outras, à medida que alcança estados e municípios.

11 Apesar disso e dos controles que o próprio *wordfish* possui, é preciso considerar que as palavras, através de sua semântica e etimologia, não possuem significados arbitrários. Ou seja, ainda que *casas* e *moradias* possam ser utilizadas com o mesmo sentido, isso não se estende de maneira aleatória ao universo linguístico, de modo que essas palavras não podem significar *árvores* ou *carros*.

12 A lei 12.034/2009 exige que todos os candidatos a cargos executivos registrem propostas de governo no TSE, no momento do pedido de registro da candidatura. A lei não versa sobre o tamanho ou tipo de documento a ser apresentado. Todos os programas estão disponíveis no site <http://divulgacandcontas.tse.jus.br/divulga/#/>. Último acesso em 28/06/2020.

Na prática, esse exercício será realizado a partir de uma comparação da posição obtida pelos partidos considerando o conjunto de eleições em cada nível com a classificação da literatura. Desconheço grandes divergências na classificação dos partidos feitas até o momento, seja através de *surveys* com especialistas ou com políticos ou mensuradas através do comportamento parlamentar. Neste artigo, utilizo como referência a classificação elaborada por Codato, Berlatto e Bolognesi (2018), pela razão prática de abranger um volume maior de partidos, incluindo os mais novos. Considero casos convergentes com a classificação dos autores como ideologicamente alinhados. Para as disputas nas quais a posição dos partidos não coincide com a classificação da literatura, concluo que outras dimensões organizam a competição programática.

É importante ressaltar que, antes de serem aplicados ao *wordfish*, os programas foram pré-processados. Foi padronizado todo o *encoding* e o formato dos arquivos; a ordem das palavras foi desconsiderada *(bag of words)*; foram retirados dos textos: acentos, numerais, pontuação, palavras com conteúdo desnecessário, como artigos, numerais, preposições e conjunções, e palavras que que, em conjunto, tiveram frequência menor do que o dobro do total de documentos de um modelo. Para evitar que palavras semelhantes fossem contadas separadamente, diminuindo a frequência de um tema, utilizei o *stemming*, processo que retira afixos das palavras, reduzindo-as ao seu tronco (*stem*)[13]. Esses processos resultaram em uma *matriz termos-documentos (tdm)* em que cada documento *i* representa um vetor que contabiliza a frequência com que cada *stem* (S) único ocorre, $S_i = S_{i1}, S_{i2}, \ldots, S_{in}$. Em seguida, essa matriz foi aplicada ao *wordfish*. Na próxima seção, apresento e analiso os resultados obtidos.

5. Competição programática no Brasil (2010-2018): uma disputa ideológica?

Neste artigo, trabalho com três diferentes *corpora*, organizados a partir do nível da disputa a que os programas se referem, totalizando 889 documentos[14]. O primeiro *corpus* diz respeito às plataformas de três eleições presidenciais – 2010, 2014 e 2018 – e soma 33 programas. O segundo se refere às disputas para governador nos mesmos anos e totaliza 468 plataformas. Por fim, o terceiro *corpus* agrega 388 documentos registrados por candidatos a prefeito nas capitais brasileiras em 2012 e em 2016[15] (Tabela 1).

13 Para exemplificar, palavras como "trabalho", "trabalhadora" e "trabalhadores" são todas consideradas como "trabalh". Sobre esse processo, também é importante considerar que ele nem sempre coincide com o radical morfológico das palavras.

14 Desconsiderei mudanças de nomes dos partidos no período, visando facilitar a comparação e a melhor visualização dos gráficos.

15 Porque a lei que exige e disponibiliza os programas é de 2009, só foram consideradas eleições posteriores a essa data.

Tabela 1 – Distribuição de documentos dos corpora

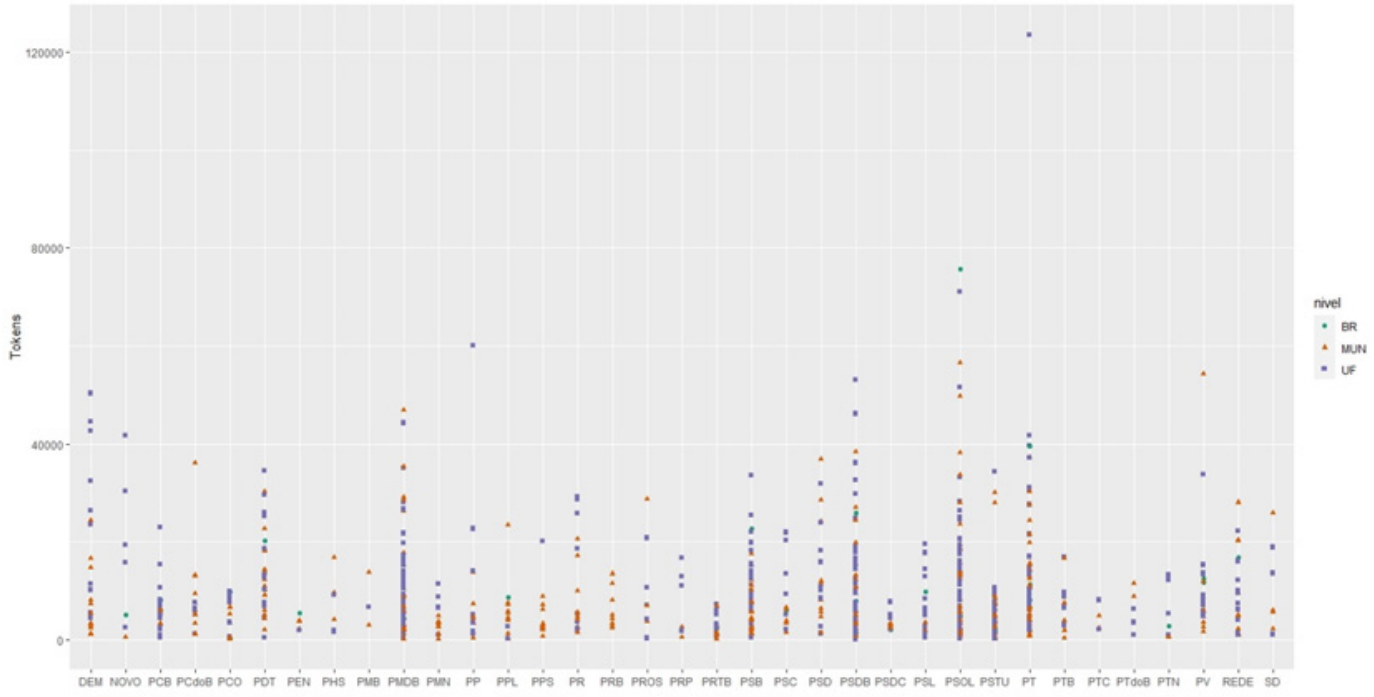
	Brasil	Estado	Município	Total
DEM	-	12	12	24
NOVO	1	5	1	7
PCB	2	21	5	28
PCdoB	-	4	11	15
PCO	2	16	7	25
PDT	1	13	17	31
PEN	1	1	3	5
PHS	-	3	3	6
PMB	-	1	2	3
PMDB	1	44	27	72
PMN	-	4	9	13
PP	-	9	10	19
PPL	1	3	13	17
PPS	-	2	10	12
PR	-	8	10	18
PRB	-	-	9	9
PROS	-	5	3	8
PRP	-	5	3	8
PRTB	2	14	12	28
PSB	1	28	22	51
PSC	1	6	5	12
PSD	-	10	10	20
PSDB	3	39	30	72
PSDC	3	5	4	12
PSL	1	16	2	19
PSOL	3	67	46	116
PSTU	3	41	32	76
PT	3	42	35	80
PTB	-	7	5	12
PTC	-	2	2	4
PTdoB	-	3	2	5
PTN	1	4	2	7
PV	2	14	10	26
REDE	1	11	10	22
SD	-	3	4	7
Total	33	468	388	889

Fonte: elaboração própria.

Como já mencionado, a lei que exige o registro dos programas de governo não especifica nenhuma característica com relação a seus conteúdos. Observando o tamanho desses documentos, também não é possível relacionar o tamanho dessas plataformas com os partidos, pois há variação no número de tokens16 incluídos nos documentos tanto entre agremiações quanto entre níveis de disputa (Gráfico 2). O maior programa foi registrado pelo PT do Acre nas eleições de 2014 e possuía 123.571 tokens, enquanto o menor foi elaborado pelo PSDB do Distrito Federal na mesma disputa e tinha apenas 67 tokens.

16 Unidade de texto analisada. Neste trabalho, o número de tokens equivale à quantidade de stems.

Gráfico 2 – Distribuição dos programas de acordo com número de tokens



Fonte: elaboração própria.

Considerando não apenas o tamanho dos documentos, mas também os termos mais frequentes, é possível observar algumas variações no que se refere às esferas e à ideologia (Tabela 2). Por exemplo, *deus* aparece apenas entre as mais recorrentes das plataformas de direita, assim como *segunranc*. Nas dos partidos de centro, isso acontece com os termos *invest* e *sustent* e, nas de esquerda, com *popul* e *cultur*. *Direit* aparece nos programas de centro e de esquerda, mas com maior frequência nos últimos. Com exceção desses tokens, restam somente mudanças na recorrência com que são mencionados: *trabalh* ocupa o 1º lugar entre os partidos de esquerda, enquanto fica em 8º entre os de direita e em 11º entre os de centro. O mesmo ocorre com *social*, que é mais frequente, respectivamente, nas plataformas das legendas de esquerda, direita e centro. Por fim, *desenvolv* é mais mencionado por candidatos de centro e possui recorrência equivalente entre os de direita e os de esquerda. Essas diferenças, no entanto, não parecem ser determinantes no que se refere aos conteúdos previstos como pertencentes a cada uma dessas categorias.

Já no *corpus* referente às eleições para o cargo de governador, surgem novos termos em comparação aos mais frequentes nas plataformas presidenciais, como *escol*, *apoi*, *recurs*, *atend* e *educ* (Tabela 3). A maior parte dessas palavras aparece nos programas dos partidos de centro, de modo que a semelhança entre os termos mais frequentes nas duas arenas de disputa é maior entre as legendas de esquerda e de direita. Assim, *segunrac* segue como uma especificidade do ranking da direita, bem como *cultur* permanece no da esquerda. *Escol* é mais mencionada pelos de direita e não aparece entre os mais recorrentes nos programas de esquerda, que, por sua vez, falam em *educ*. Diferentemente das plataformas presidenciais, *popul* surge, nos estados, para agremiações de todo o espectro ideológico, embora, junto com *trabalh*, seja mais frequentes entre as de esquerda.

Tabela 2 – Termos mais frequentes no corpus das eleições presidenciais

	Direita	Centro	Esquerda	Total
1º	tod	tod	**trabalh**	**trabalh**
2º	acima	dev	govern	tod
3º	govern	**desenvol**	**popul**	govern
4º	estad	ser	tod	nacional
5º	nacional	nacional	nacional	desenvolv
6º	**desenvolv**	govern	**desenvolv**	**social**
7º	dev	estad	**social**	popul
8º	**trabalh**	form	**direit**	dev
9º	**social**	garant	garant	direit
10º	ser	pod	estad	estad
11º	melhor	**trabalh**	dev	garant
12º	todos	nov	nov	ser
13º	**deus**	apoi	ser	nov
14º	setor	**social**	sistem	form
15º	federal	maior	form	pod
16º	tudo	**invest**	pod	sistem
17º	bolsonaro2018	qualidade	part	maior
18º	nov	**sustent**	grand	part
19º	pod	**direit**	**cultur**	**invest**
20º	**seguranc**	sistem	maior	setor

Fonte: elaboração própria.

Tabela 3 – Termos mais frequentes no *corpus* das eleições estaduais

	Direita	Centro	Esquerda	Total
1º	estad	estad	estad	estad
2º	govern	govern	govern	govern
3º	desenvolv	desenvolv	**trabalh**	desenvolv
4º	tod	tod	**popul**	tod
5º	melhor	nov	tod	trabalh
6º	social	melhor	desenvolv	social
7º	nov	**social**	**social**	popul
8º	**escol**	estadual	garant	Bnov
9º	**popul**	sistem	**direit**	estadual
10º	ser	implant	estadual	ser
11º	sistem	invest	form	melhor
12º	**trabalh**	ser	ser	sistem
13º	estadual	**popul**	**cultur**	garant
14º	projet	projet	part	form
15º	outr	**atend**	**educ**	outro
16º	**seguran**	**recurs**	grand	escol
17º	**recurs**	**apoi**	nov	educ
18º	form	**trabalh**	pod	projet
19º	implant	produt	sistem	pod
20º	produt	**escol**	outr	recurs

Fonte: elaboração própria.

Nas eleições municipais, *cultur* passa a figurar no ranking de partidos de todo o espectro ideológico, mas com maior frequência entre os de esquerda (Tabela 4). Nesses, a particularidade passa a ser o termo *particip* e a manutenção de *direit*, agora ausente dos mais recorrentes da direita e do centro. *Escol* ingressa no ranking da esquerda. Além dessas palavras, *urban* se insere entre os mais recorrentes para agremiações de todo o espectro, ressaltando, além do 1º lugar, que o debate agora é sobre as cidades.

Tabela 4 – Termos mais frequentes no corpus das eleições municipais

	Direita	Centro	Esquerda	Total
1º	cidad	cidad	cidad	cidad
2º	municipal	municipal	**popul**	municipal
3º	desenvolv	desenvolv	municipal	popul
4º	tod	tod	**trabalh**	tod
5º	**popul**	**popul**	**social**	social
6º	**social**	nov	tod	desenvol
7º	**escol**	melhor	govern	trabalh
8º	**atend**	**social**	desenvolv	govern
9º	melhor	implant	**cultur**	escol
10º	implant	govern	garant	cultur
11º	projet	**atend**	**educ**	atend
12º	nov	**escol**	**escol**	garant
13º	cri	**urban**	**direit**	nov
14º	**urban**	plan	ser	urban
15º	intergr	**cultur**	**particip**	educ
16º	ser	projet	**urban**	ser
17º	govern	ser	dev	melhor
18º	**cultur**	sistem	form	implant
19º	red	**trabalh**	outr	projet
20º	promov	**educ**	pod	cri

Fonte: elaboração própria.

Apesar de algumas variações, a descrição dos termos mais frequentes não é de tudo útil para pensar a classificação ideológica dos partidos, por isso a importância de analisá-los por meio do modelo de escalonamento, *wordfish*. A maior diferença, é que agora são considerados todos os termos de um programa e as suas palavras têm peso a partir da frequência, como já apresentado. O ponto de partido é a disputa presidencial, visto que foram os documentos utilizados pela literatura já apresentada e, também, são a arena nas quais grandes pacotes de políticas são passíveis de serem propostos e implementados.

Como já mencionado, o *wordfish* atribui posições a partir da frequência e não do sentido das palavras. Para alguns, isso pode colocar em xeque o pressuposto deste trabalho de não identificar disputa programática em situações nas quais não há divergência entre os termos mobilizados. Afinal, candidatos poderiam usar termos semelhantes, para não dizer idênticos, com sentidos distintos. Seria o caso, por exemplo, de um que se dedicasse a propor políticas de trabalho voluntário e criação de postos de trabalho na iniciativa privada; e de outro que privilegiasse tópicos sobre trabalho relacionando-os ao funcionalismo público. Nessa situação, *trabalh* seria contado nos dois documentos. Embora essa limitação seja verdadeira, não acredito que seja incontornável, pois, se o *wordfish* não captura o sentido das palavras, a inclusão de todos os termos oferece algum grau

de diferenciação da política. No exemplo mencionado, tokens relacionados a iniciativa privada e a voluntário entrariam na mensuração da posição do primeiro partido, assim como funcionalismo público seria considerada na do segundo.

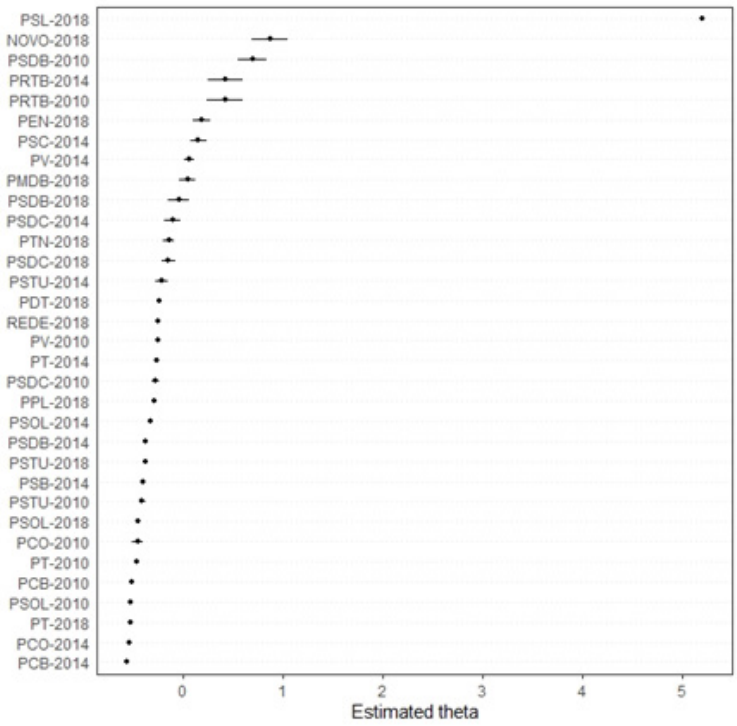
Observado o Gráfico 3, é possível perceber que, diferentemente dos resultados obtidos através do uso da codificação *MARPOR*, o *wordfish* escalona as legendas mais ou menos próximas ao centro do espectro (ideológico?), embora a maioria delas esteja à esquerda,. Por essa classificação, o partido mais à direita é o PSL, em 2018, e o PCB, em 2014, o mais à esquerda. De uma maneira geral, com exceção do programa de 2014 do PSTU, que aparece perto do centro, pode-se afirmar que o posicionamento dos partidos de esquerda é consistente com a classificação da literatura, uma vez que PT, PSOL, PSTU, PCO e PCB estão próximos e em um dos polos da escala. Isso não significa que não há inconsistências, como o fato de a plataforma do PSDB de 2014 estar entre a do PSOL (2014) e a do PSTU (2014); ou a do PSDC de 2010 se encontrar mais próxima da do PT (2014) e da do PPL (2018). No entanto, o espectro programático das eleições presidenciais parece ser fortemente organizado pela ideologia.

5.1. Competição programática nos executivos subnacionais (2010-2018)

Esse panorama, no entanto, sofre alterações quando consideradas as eleições para o executivo estadual e municipal. Em alguns estados, a configuração é mais semelhante com o cenário nacional. Já em outras, é bastante diferente, como é o caso do Acre (Gráfico 4). Considerando a disputa para governador, sequer poderia afirmar que há uma competição programática efetiva, uma vez que o posicionamento dos partidos e quase sobreposto. Ou seja, eles usaram praticamente os mesmos termos em seus programas de governo. Há diferenças, como o PSDB e o PT e 2014, mas a mais significativa é, indiscutivelmente, o posicionamento mais extremo do programa do PSL em 2018[17]. As plataformas apresentam maior divergência entre si nas disputas pela prefeitura de Rio Branco, mas, mesmo assim, o PSOL (2012) está localizado no centro, junto à maior parte das legendas. Os polos são protagonizados pelo PT na esquerda e PSC na direita. Assim, a competição programática no estado parece se organizar mais ideologicamente no nível municipal do que no estadual. Afinal, nessa arena, a disputa seria muito pouco programática, porque não divergente.

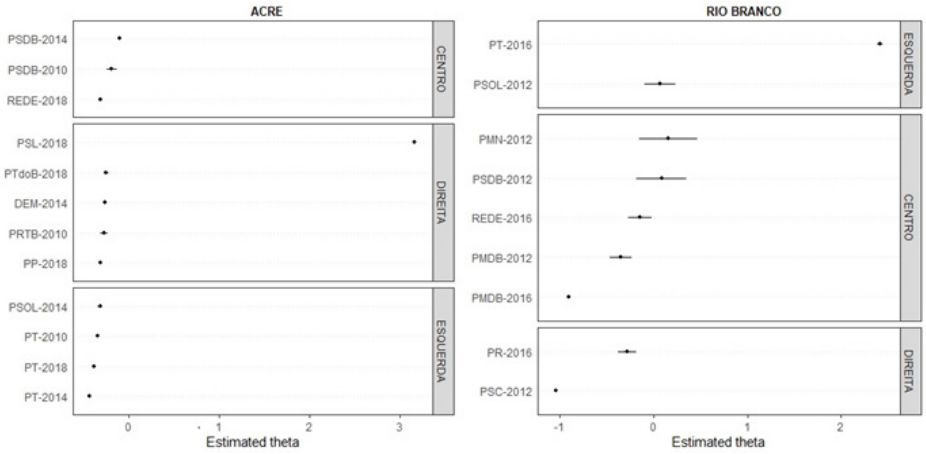
17 É importante ressaltar que o sentido (positivo ou negativo) da posição é dado arbitrariamente pelo modelo. Desse modo, estar o polo negativo não representa necessariamente estar à esquerda, esta é uma interpretação do analista.

Gráfico 3 – Competição programática presidencial (2010-2018)



Fonte: elaboração própria.

Gráfico 4 – Competição programática no Acre (2010-2018)

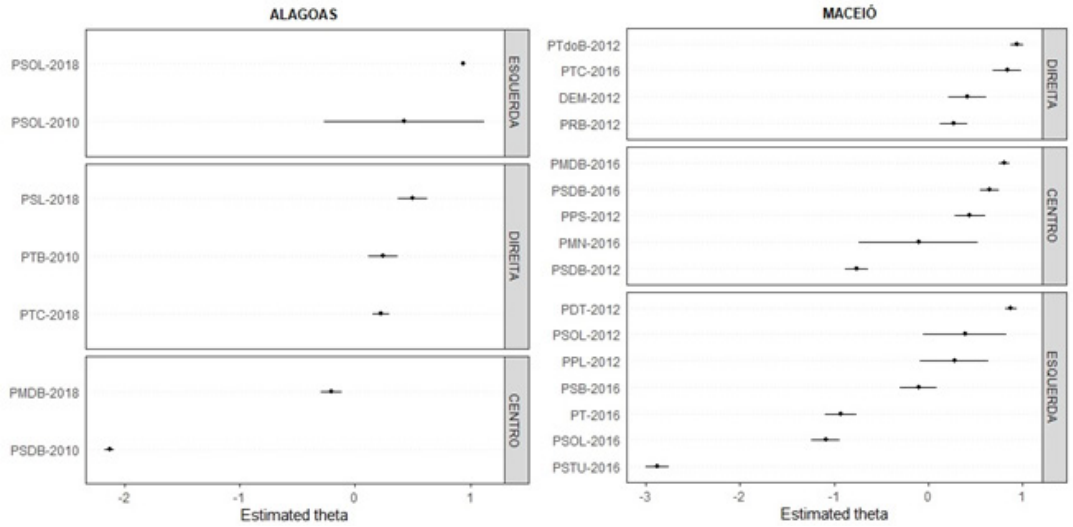


Fonte: elaboração própria.

Em Alagoas, há mais divergência e, portanto, maior competição entre as plataformas de candidatos ao executivo estadual do que no Acre (Gráfico 5). O PSOL ocupa a posição mais extrema à esquerda e o PSDB, à direita. PSL, PTB e PTC apresentaram propostas semelhantes. No entanto, o fato de o PSL, em 2018, ter registrado programa em localização próxima ao do PSOL, em 2010, indica limitações da interpretação dessa escala como um espectro precisamente ideológico. Essas questões também se reproduzem nas disputas municipais da capital alagoana. De um lado, PTdoB, PTC, PMDB e PSDB se localizam mais à direita, enquanto PSTU, PSOL e PT, se concentram mais à esquerda. Da mesma maneira que na eleição estadual, este cenário é, em alguma medida turvo, na medida que posições semelhantes foram atribuídas a PPL, PSOL, PPS e PRB. Em geral, a competição pro-

gramática parece mais organizada ideologicamente em Alagoas do que no Acre. Todavia, as inconsistências podem sugerir que talvez existam outras dimensões programáticas relevantes nessas disputas.

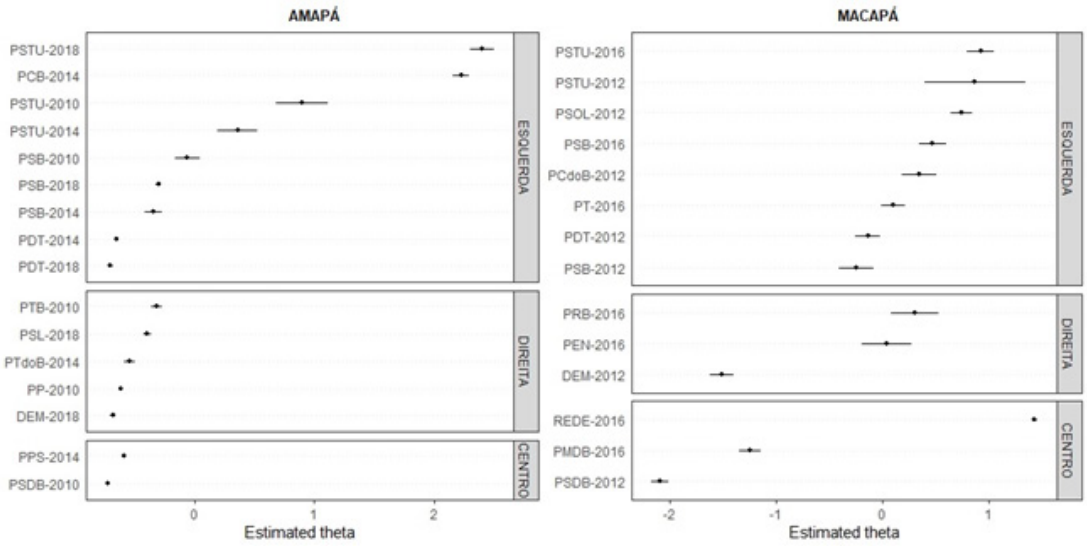
Gráfico 5 – Competição programática em Alagoas (2010-2018)



Fonte: elaboração própria.

No Amapá, as diferenças ideológicas aparentam estar mais bem delineadas na esquerda, já que PSTU, PCB e PSB (2010) possuem posições bem demarcadas e próximas entre si (Gráfico 6). Diferentemente do Acre, aqui a disputa se estrutura de maneira mais polarizada e, mais à direita, se encontram PSDB, PDT, e DEM. Há inconsistências, sobretudo se consideramos a proximidade do PSL com o PSB, em 2018 e em 2014. A delimitação mais evidente do campo da esquerda também se reproduz na disputa pela prefeitura de Macapá, protagonizada por PT, PCdoB, PSB (2016), PSOL e PSTU. A posição mais à esquerda, porém, é a do programa da REDE, em 2016. À direita estão, respectivamente, PSDB, DEM e PMDB, ao mesmo tempo em que PEN, PRB, PSB (2012) e PDT ocupam posições mais centrais.

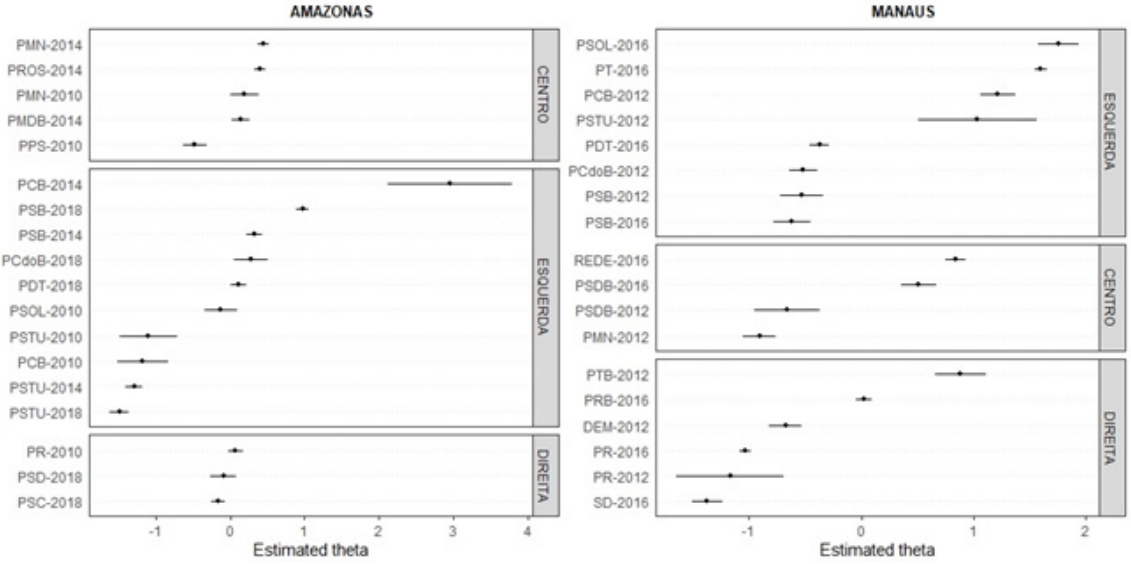
Gráfico 6 – Competição programática no Amapá (2010-2018)



Fonte: elaboração própria.

De forma distinta das configurações vistas em outros estados, no Amazonas a competição programática parece se estruturar de maneira mais divergente quando comparamos as duas arenas de disputa (Gráfico 7). Se nas eleições para o executivo estadual, a maioria dos partidos se localiza no centro do espectro programático, na competição para o executivo municipal a escala é mais polarizada. Assim, na disputa estadual, PSC, PSD, PR, PSOL, PDT, PSB (2014), PMDB, PROS e PMN ocupam posições central, enquanto PSTU e PCB estão à esquerda e PSB (2018) e PCB (2014) à direita. Essa configuração parece indicar que tal competição programática poderia ser melhor compreendida por outras dimensões, somadas ou não à ideológica. Isso, de alguma maneira, se estende às eleições municipais, visto que, embora as posições estejam menos centrais, há polarização em todos os campos ideológicos. Assim, a esquerda aparece dividida entre: PSOL, PT, PCB e PSTU, por um lado, e PDT, PCdoB e PSB, por outro. O centro, entre REDE e PSDB (2016) e PSDB (2012) e PMN; e a direita entre PTB, por um lado, e SD, PR e DEM, por outro, com PRB ocupando posição central.

Gráfico 7 – Competição programática no Amazonas (2010-2018)



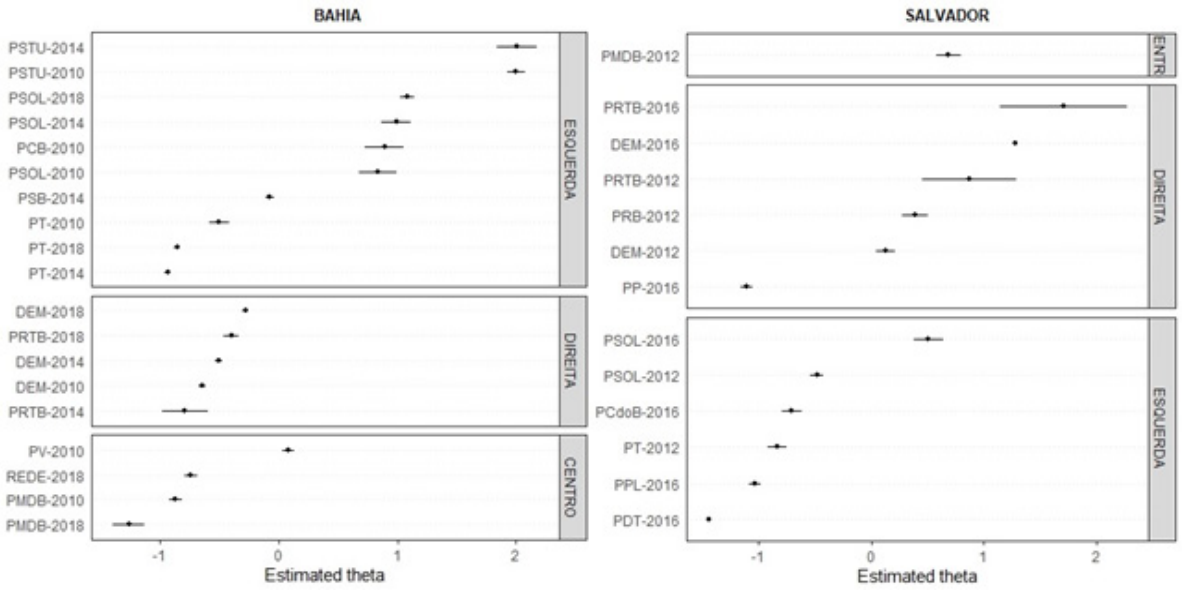
Fonte: elaboração própria.

Na Bahia, a posição dos partidos de esquerda é bem demarcada tanto nas disputas estaduais quanto nas municipais (Gráfico 8). Nas primeiras, o PT constitui uma exceção, já que se posiciona mais próximo dos partidos de centro e direita – o que acontece em 2016 com o PSOL nas eleições para prefeito. No nível estadual, as agremiações de centro se localizam mais aos extremos do que as de direita, mas, em geral, possuem posições semelhantes. Na arena local, por sua vez, o espaço central não é ocupado e, excluindo os casos do PP e do PSOL em 2016, a competição programática parece ideologicamente alinhada, visto que PST, PPL, PT, PCdoB e PSOL (2012) se encontram à esquerda e DEM, PRB, PRTB e PMDB, à direita.

Programas registrados nas eleições de 2018 ocupam as posições mais extremas no espectro programático estadual do Ceará, com o PSL em um dos polos e o PSTU no outro (Gráfico 9). PMDB e PSDB (2010) se localizam no centro, com movimento do PSDB mais à direita em 2018. Há, no entanto, inconsistências no campo da esquerda, já que PSTU e PSOL se posicionam próximos ao valor -1 e PT e PSB mais perto de 1. Nesse sentido, a esquerda seria polarizada ao longo de um espaço

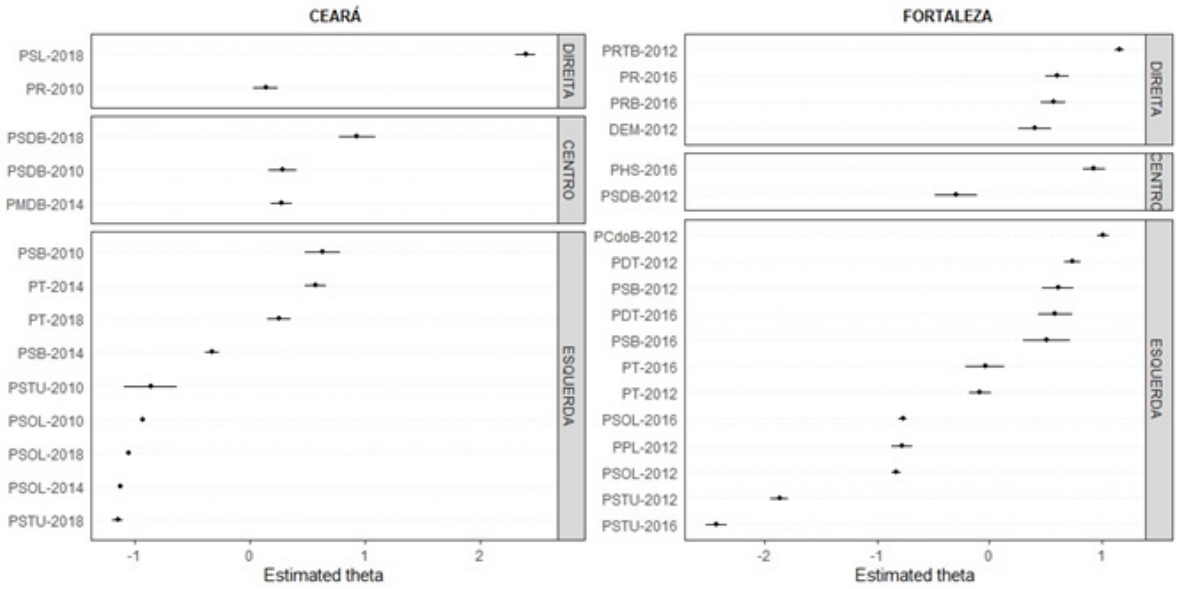
programático que parece incluir outras dimensões além – ou em vez – da ideológica. Esse cenário se radicaliza quando consideradas as eleições para a prefeitura de Fortaleza, com maior dispersão da esquerda em todo o espectro programático. Se ele consistisse em uma escala ideológica, o PT ocuparia posição central e somente PSOL, PPL e PSTU estariam localizados na esquerda. O PCdoB, por exemplo, estaria mais à direita do que DEM, PRB e PR. Assim, outras dimensões programáticas parecem articular essas disputas.

Gráfico 8 – Competição programática na Bahia (2010-2018)



Fonte: elaboração própria.

Gráfico 9 – Competição programática no Ceará (2010-2018)

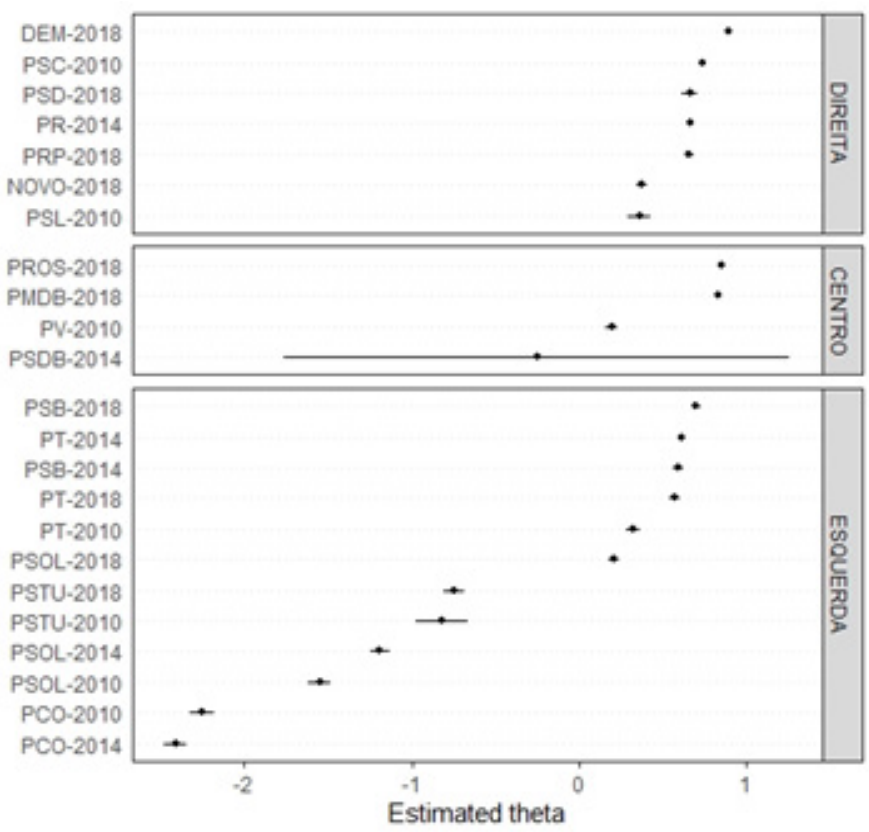


Fonte: elaboração própria.

A competição programática para o executivo do Distrito Federal também parece não se organizar de maneira ideológica, visto que não há posições delimitadas para partidos considerados de esquerda, centro e direita (Gráfico 10). Em outras palavras, há posições semelhantes para diferentes partidos, independentemente de suas classificações ideológicas. Com localização mais extrema e delimitada estão os pequenos partidos de esquerda, PSTU, PSOL e PCO. A reiteração desses posicionamentos ao longo dos estados e municípios parece sugerir que esses partidos podem exercer uma influência centrípeta sobre os demais de esquerda, dado que o caráter comparativo do modelo.

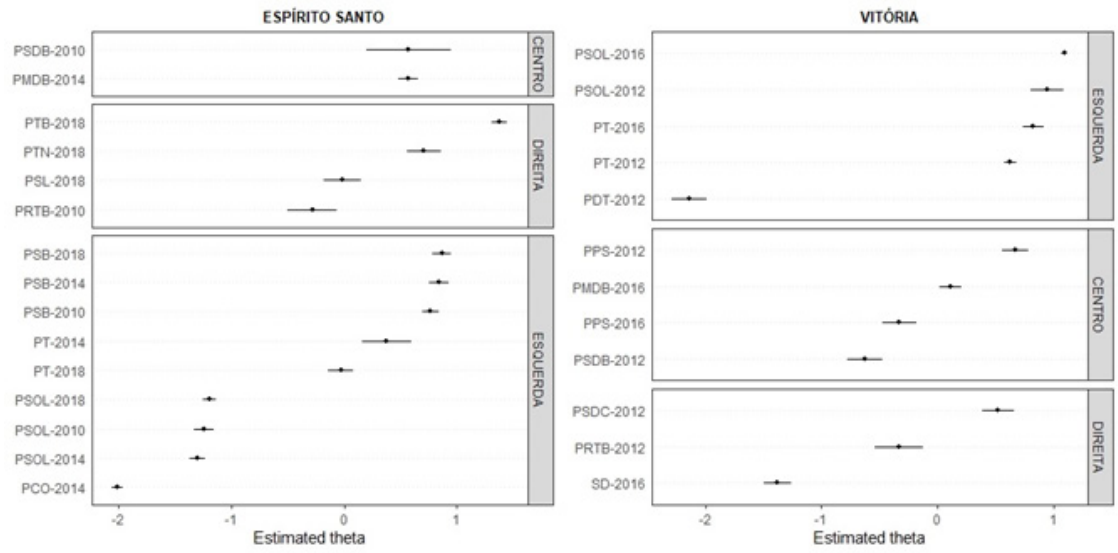
No Espírito Santo, as posições mais extremas são consistentes com a classificação ideológica dos partidos (Gráfico 11). No nível estadual, está, de um lado, o PTB e, de outro, o PCO. Isso não ocorre, no entanto, no nível municipal, onde legendas de esquerda ocupam ambos os polos, PSOL no positivo e PDT no negativo. Apesar disso, a localização das agremiações de esquerda é mais concentrada quando considerados os programas registrados por candidatos à prefeitura de Vitória do que ao governo estadual. Nessa esfera, não haveria grande divergência entre PSDB, PMDB e PSB, por exemplo, e PSL e PRTB estão mais próximos do PT do que do PTN. Esses dados parecem indicar que, também neste estado, há outras dimensões programáticas organizando a competição.

Gráfico 10 – Competição programática no Distrito Federal (2010-2018)



Fonte: elaboração própria.

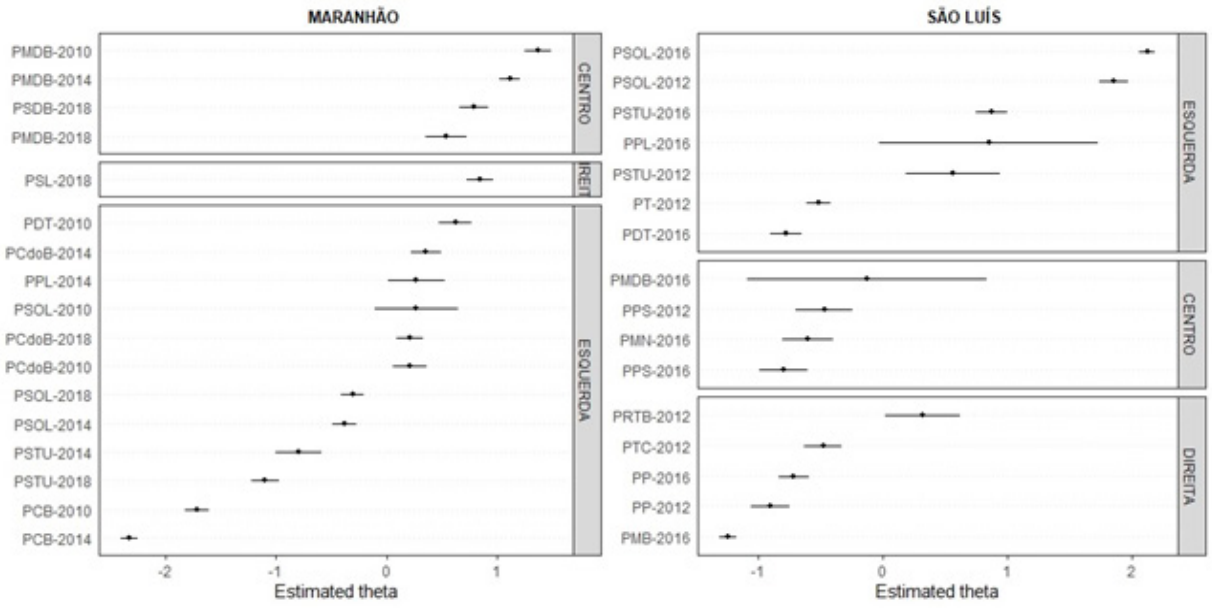
Gráfico 11 – Competição programática no Espírito Santo (2010-2018)



Fonte: elaboração própria.

À primeira vista, pode parecer que o caso do Maranhão é parecido com o cenário estadual de Goiás, já que as posições também se distribuem quase em forma de reta inclinada (Gráfico 13). Se observamos o ponto médio, no entanto, percebemos que ele não divide os partidos entre esquerda e direita, separando apenas as legendas mais extremas, PSOL, PSTU e PCB. O PSL também ocupa posição mais central do que o PMDB no total das disputas, porém o maior problema parece ser o fato de que PCdoB, PPL, PSOL e PDT também dividem de alguma maneira esse mesmo espaço. A competição no nível local, por sua vez, parece apresentar menos inconsistência do que se refere à relação entre ideologia e a localização dos partidos. Assim, embora PRTB e PT estejam mais ao centro, assim como o PDT mais à direita, há um posicionamento mais concentrado de legendas de uma mesma classificação ideológica.

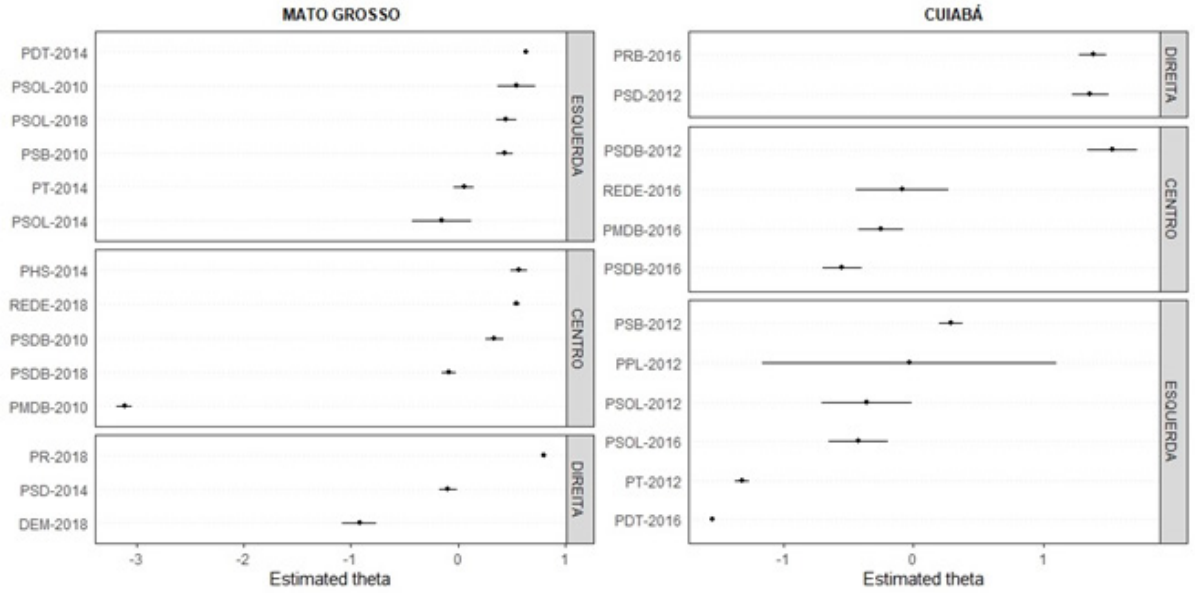
Gráfico 13 – Competição programática no Maranhão (2010-2018)



Fonte: elaboração própria.

A disposição dos partidos na disputa para o governo estadual do Mato Grosso se assemelha à visualizada no Acre, dado que há significativa sobreposição, indicando pouca divergência – e, portanto, competição programática (Gráfico 14). Os candidatos a prefeito, por sua vez, distinguem mais suas propostas entre si. No entanto, a disputa é travada, principalmente, em torno do centro, onde se localizam PSOL, PPL, PSB, PSDB (2016), PMDB e REDE. À esquerda, estão somente PT e PDT e, à direita, PSDB, PSD e PRB. Em comparação com outras capitais, o cenário em Cuiabá não constitui uma completa desordem ideológica, mas parece estar conectado com outras dimensões programáticas.

Gráfico 14 – Competição programática no Mato Grosso (2010-2018)

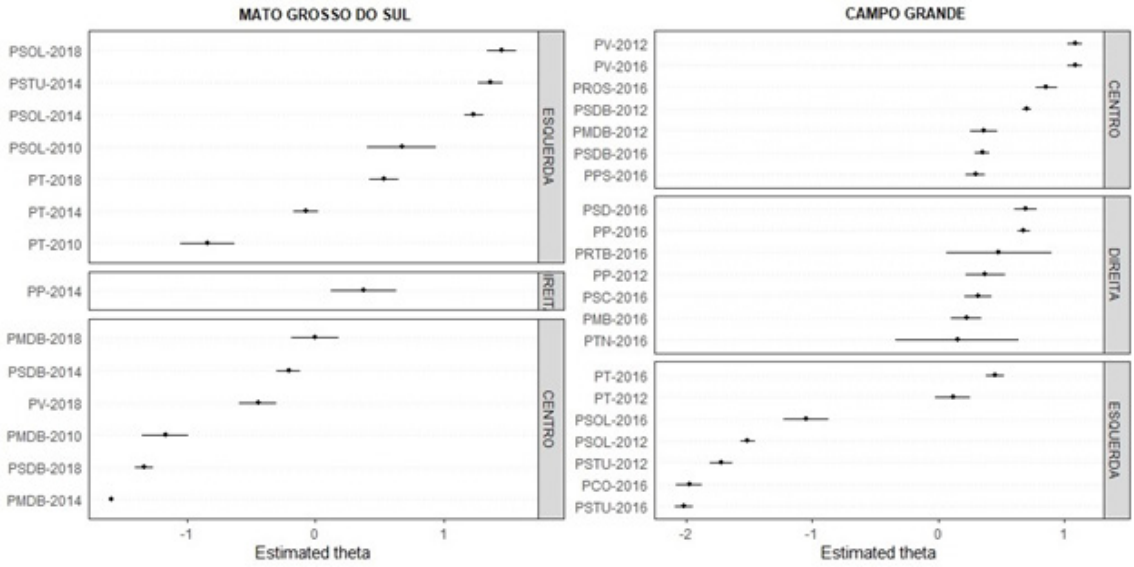


Fonte: elaboração própria.

No Mato Grosso do Sul, por sua vez, a disputa é mais ideologicamente alinhada, ainda que com inconsistências (Gráfico 15). Na arena estadual, a principal delas consiste no fato de pelo menos a metade dos partidos de centro ocuparem o campo da direta, enquanto a legenda de direita, PP, aparecer na centro-esquerda. Na esquerda, o PT, em 2014, tem com posição central e, em 2010, à direita. Em 2018, o partido está mais próximo dos demais, PSOL e PSTU.

O cenário estadual se reproduz nas disputas para a prefeitura de Campo Grande. Nos dois anos considerados, o PT aparece com posição central (2012) e mais à direita (2016), e os partidos de centro em posição mais extrema do que os de direita. As outras agremiações parecem ideologicamente organizadas, com PSOL, PSTU e PCO à esquerda, PMDB (2012), PSDB e PPS no centro, e PTN, PMB, PSC, PP, PRTB e PSD à direita.

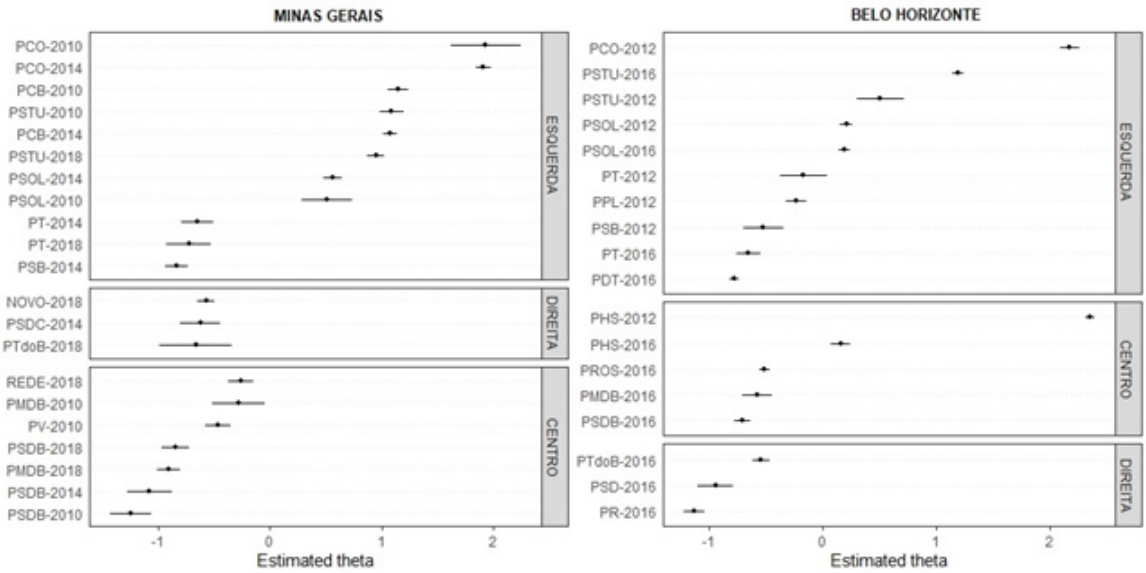
Gráfico 15 – Competição programática no Mato Grosso do Sul (2010-2018)



Fonte: elaboração própria.

Essa configuração se assemelha à encontrada nas eleições para o executivo estadual em Minas Gerais, onde PSDB e PMDB (2018, 2014) aparecem mais à direita do que o NOVO, o PSDC e o PTdoB, assim como PT e PSB (Gráfico 16). Novamente, o caso sugere que os partidos de esquerda mais extremos e menos competitivos exercem força centrípeta sobre os demais no mesmo campo ideológico. Nas disputas pela prefeitura de Belo Horizonte, isso se acentua, visto que inclui mais agremiações: PT, PPL, PSB e PDT figuram entre o centro e a direita. Embora, nesse nível, os partidos de centro não apresentem posições mais extremas que os de direita, o PHS surge tão à esquerda quanto o PCO. Esses cenários podem ser interpretados como um caso em que a dimensão ideológica se junta a outras, uma vez que a competição programática nem é totalmente organizada nem, tampouco, totalmente desalinhada.

Gráfico 16 – Competição programática em Minas Gerais (2010-2018)

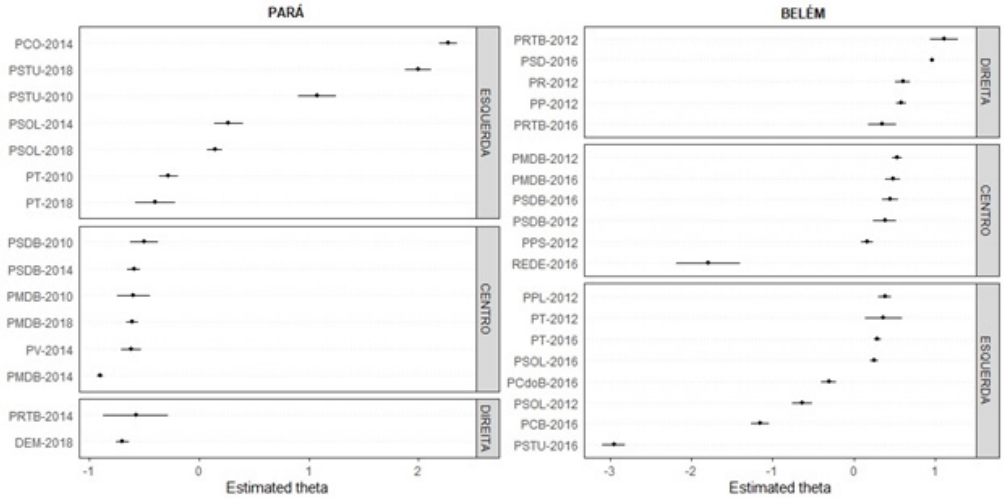


Fonte: elaboração própria.

No Pará, a ideologia não parece ser a dimensão primordial na organização da competição programática nem no nível estadual nem no local (Gráfico 17). No primeiro, PRTB, PV, PMDB (2010, 2018) e PSDB apresentam praticamente a mesma posição, que é também muito próxima do DEM e do PT em 2018. O PMDB, em 2014, é o partido que possui posicionamento mais extremo à direita e o PCO, no mesmo ano, é o mais extremo no polo oposto. Nas eleições para a prefeitura de Belém, o PRTB, em 2012, e o PSTU, em 2016, ocupam essas respectivas posições. No caso da disputa municipal, a direita é mais extrema do que o centro, porém, mesmo entre parte da esquerda, os espaços ocupados são semelhantes, reforçando o pouco alinhamento ideológico da disputa programática.

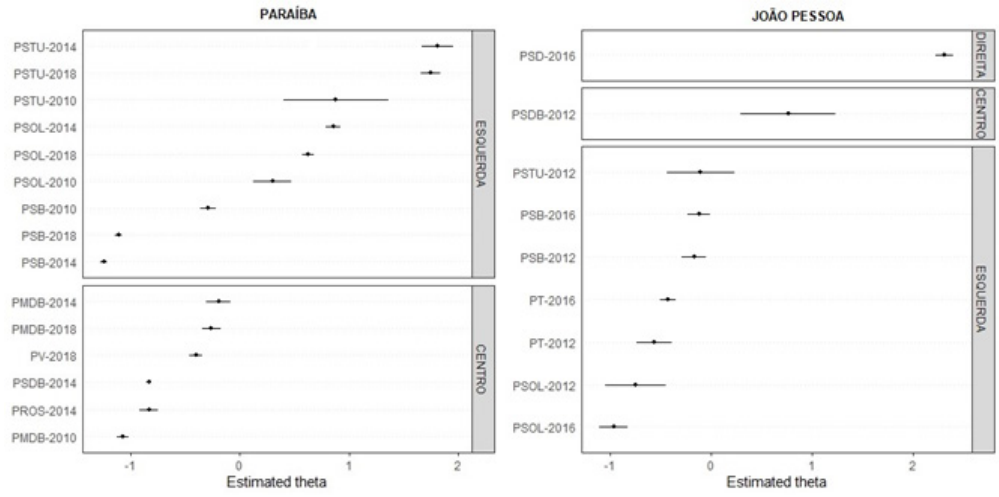
A competição programática na Paraíba aparenta ser parcialmente organizada pela ideologia, ao menos no nível estadual (Gráfico 18). Isso porque PSOL e PSTU ocupam a esquerda e PMDB (2018, 2014) e PV, o centro. Já o PSB apresenta posição central em 2010 e mais à direita em 2014 e 2018, onde também estão localizados PSDB, PROS e PMDB (2010). Nas disputas pela prefeitura de João Pessoa, a competição programática é organizada pela ideologia de forma mais consistente, com PSOL, PT, PSTU e PSB à esquerda, PSDB no centro e PSD na direita.

Gráfico 17 – Competição programática no Pará (2010-2018)



Fonte: elaboração própria.

Gráfico 18 – Competição programática na Paraíba (2010-2018)

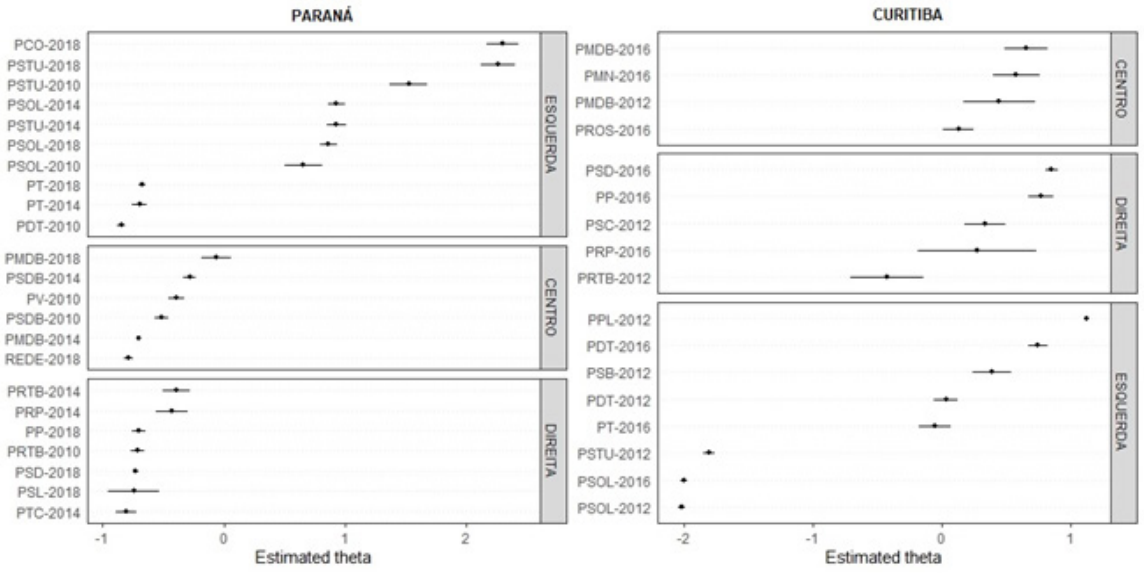


Fonte: elaboração própria.

A competição programática para o cargo de governador do Paraná situa de maneira consistente os partidos de direita (Gráfico 19). As legendas de centro, aparecem em posições mais extremas do que centrais, com exceção do PSDB (2014) e do PMDB (2018). O maior problema parece ser a localização do PT e do PDT também nesse espaço, possivelmente influenciadas por posicionamento mais acentuado do PCO, PSTU e PSOL, como visto em outros estados. No que se refere às eleições municipais, embora haja aglomeração das legendas entre os valores 0 e 1, elas estão mais dispersas entre si quando comparadas à arena estadual. Isso não garante, no entanto, alinhamento ideológico, uma vez que os partidos de centro estão à direita do PSD, PP, PSC e PRP e as posições extremas são ocupadas pela esquerda: um dos polos, pelo PSOL, e, outro, pelo PPL.

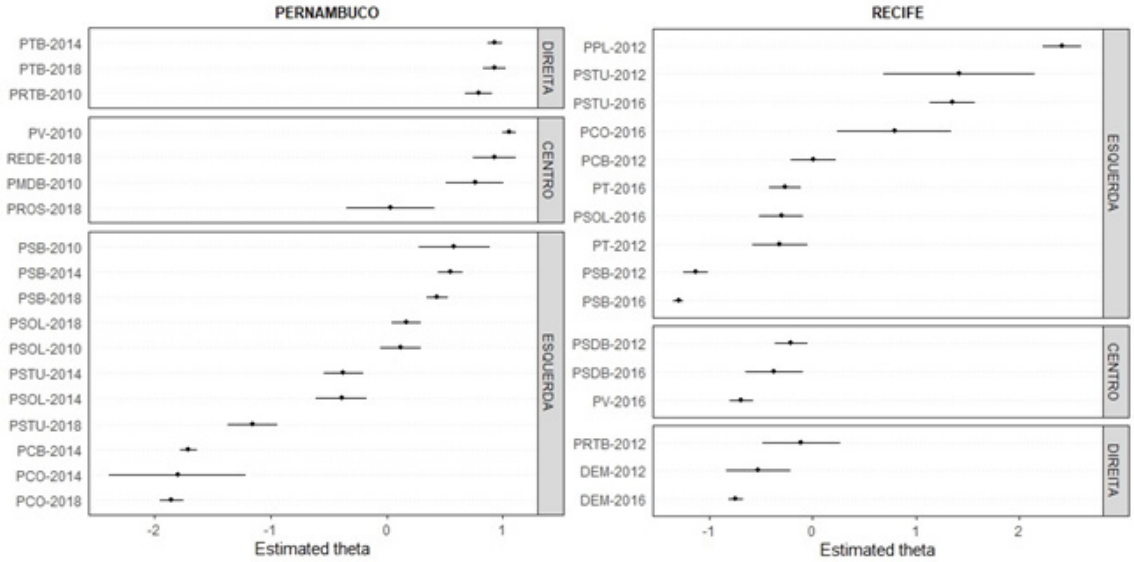
A dispersão da esquerda também se reproduz nas disputas travadas nos âmbitos estadual e municipal em Pernambuco (Gráfico 20). Nos dois casos, parte das legendas se posiciona à esquerda – PSTU, PSOL (2014), PCB e PCO, no nível estadual, e PCO, PSTU e PPL, no municipal – enquanto os demais se espalham entre o centro e a direita. O caso de Pernambuco parece um pouco mais acentuado que os demais, pois mesmo partidos como PSOL (2010, 2018), na arena estadual, e PCB, em Recife, se situam de maneira mais central. Em ambos os casos, as agremiações de centro e de direita possuem posicionamento semelhante, entre 0 e 1 do espectro programático. Tais configurações apontam mais para o impacto de outras dimensões que não a ideologia na organização do espaço programático de disputa.

Gráfico 19 – Competição programática no Paraná (2010-2018)



Fonte: elaboração própria.

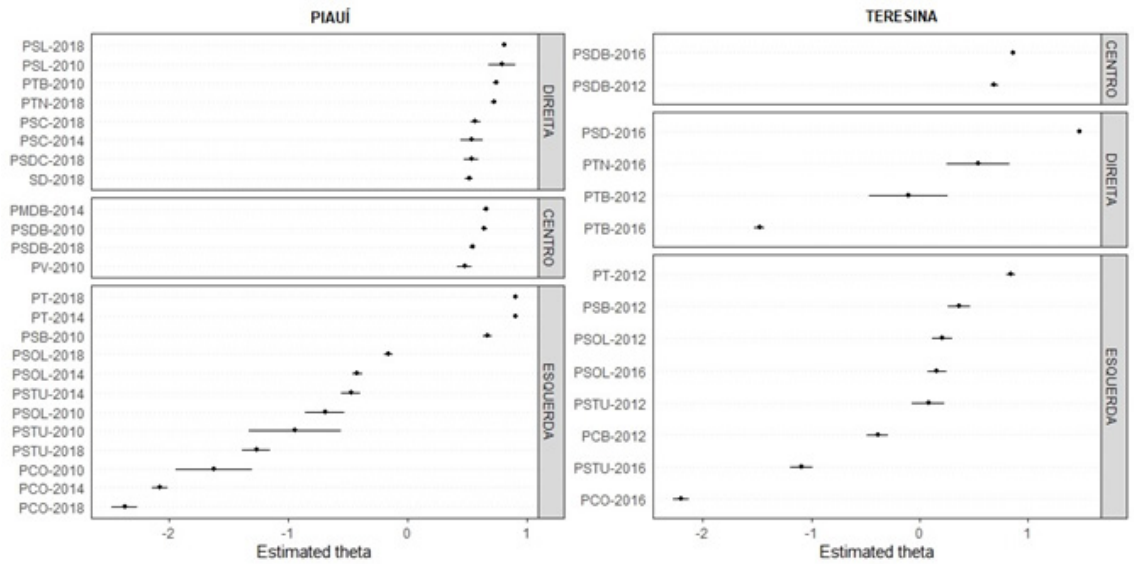
Gráfico 20 – Competição programática em Pernambuco (2010-2018)



Fonte: elaboração própria.

No Piauí, os cenários traçados a partir do posicionamento dos partidos no espectro programático é distinto (Gráfico 21). Enquanto nas eleições estaduais há concentração dos partidos de direita em um polo oposto de onde se localizam as agremiações de esquerda, as posições de legendas de ambos os campos são mais dispersas nas disputas pela prefeitura de Teresina. Há, porém, uma reincidência nos dois casos: os partidos de centro aparecem à direita, sem ocupar, no entanto, a posição mais extrema. O polo da direita é protagonizado pelo PSD e o da esquerda, pelo PCO. Apesar disso, PSTU, PSOL e PTB se situam em espaços centrais, indicando a fraca organização ideológica da competição programática municipal. Com exceção das posições do PSB e PT, a ideologia parece ter maior capacidade de alinhamento nas eleições para governador.

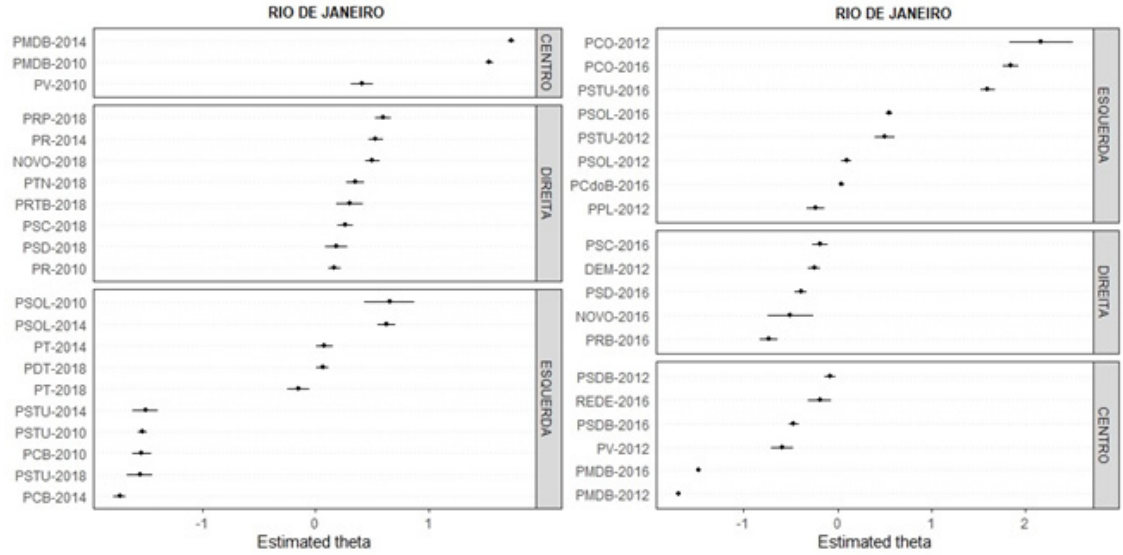
Gráfico 21 – Competição programática no Piauí (2010-2018)



Fonte: elaboração própria.

A configuração das disputas no Rio de Janeiro chama a atenção pela posição extrema à direita do PMDB, tanto na área estadual quanto local (Gráfico 22). Ela se distingue não somente pelo partido ser classificado no centro, mas, principalmente, porque posicionamentos mais acentuados costumam ser atribuídos a legendas menos competitivas. Não é apenas essa a particularidade encontrada. Nas eleições para o governo do estado, PSOL aparece na direita e PT e PDT no centro. Nas disputas pela prefeitura do Rio, os partidos de esquerda não se dispersam pela direita, mas PPL possui localização central, muito próxima do PSC, DEM, PSDB e REDE. Esses panoramas mostram que a dimensão ideológica é insuficiente para explicar o espaço programático de disputa, sobretudo no nível estadual.

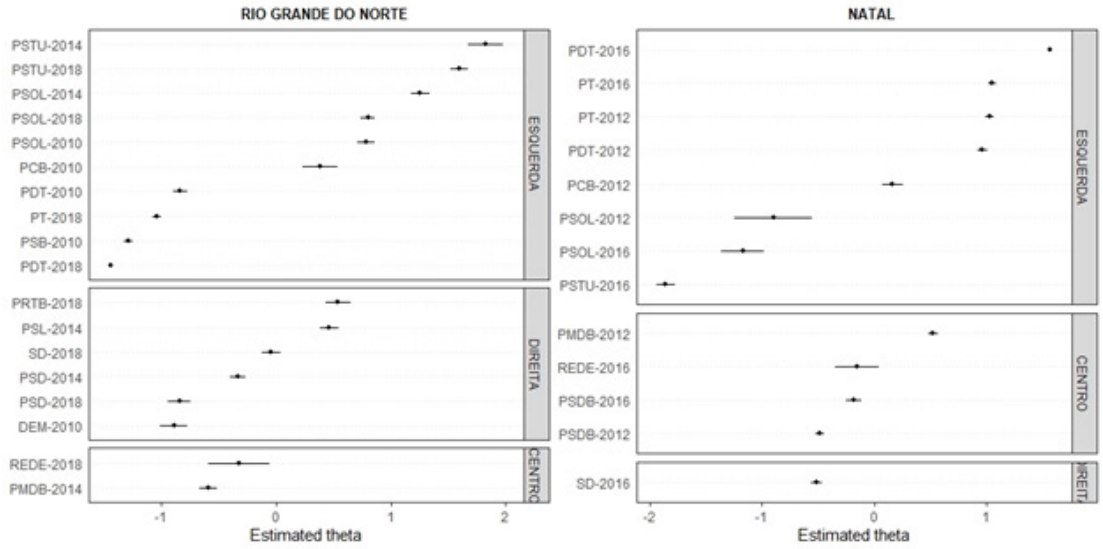
Gráfico 22 – Competição programática no Rio de Janeiro (2010-2018)



Fonte: elaboração própria.

Esse resultado parece se estender ao caso do Rio Grande do Norte, com a diferença que a configuração entre os dois níveis é mais alinhada do que a observada no Rio de Janeiro (Gráfico 23). Aqui, partidos de esquerda ocupam as posições mais extremas nos dois polos. No âmbito estadual, PSTU, PSOL e PCB se situam em polarização ao PDT, PT e PSB e, no municipal, a oposição se daria entre PSOL e PSTU, de um lado, e PDT e PT, de outro, com o PCB ocupando posição central. Se não é a primeira vez que essa configuração polarizada se estrutura no campo da esquerda, é no Rio Grande do Norte que ela ocorre simultaneamente com oposição entre os partidos de direita, no plano estadual, e entre os de centro, no municipal. Esses arranjos não tornam difícil descartar a hipótese de que a competição programática nesses locais se estrutura de maneira ideológica.

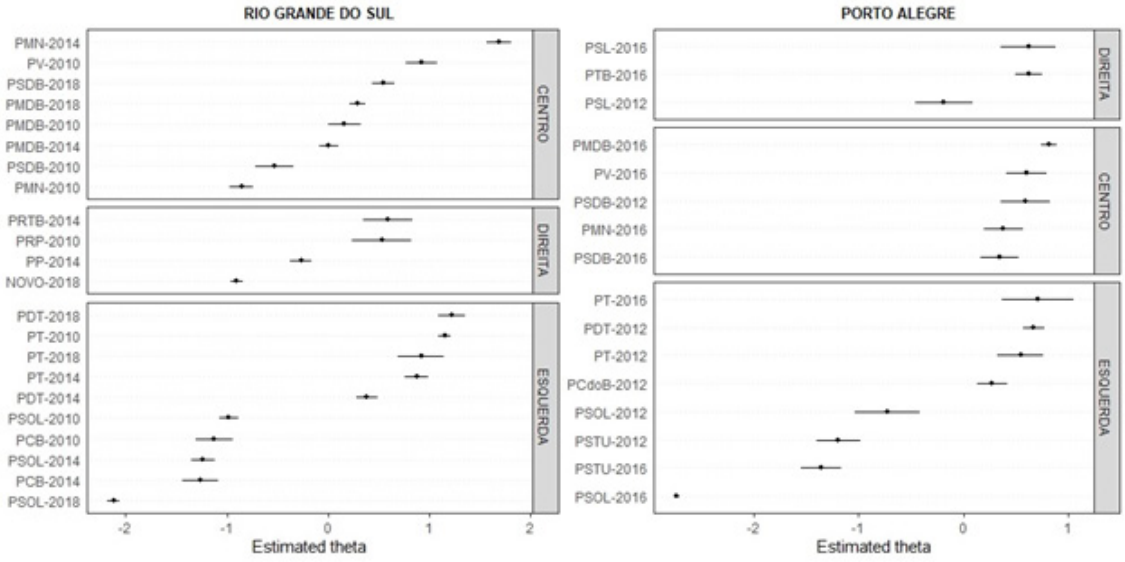
Gráfico 23 – Competição programática no Rio Grande do Norte (2010-2018)



Fonte: elaboração própria.

O cenário de polarização entre os partidos de um mesmo campo se reproduz nas disputas para governador do Rio Grande do Sul, nas quais tanto legendas de centro quanto de esquerda e de direita aparecem em oposição entre si (Gráfico 24). Isso também acontece nas eleições para a prefeitura de Porto Alegre, mas apenas entre os partidos de esquerda. Nesse caso, PSOL e PSTU se posicionam à esquerda, enquanto PCdoB, PDT e PT aparecem na direita, com posições semelhantes ao PSL e ao PTB. Partidos de centro também se localizam mais à direita do que próximos ao ponto médio do espectro programático. Assim, a ideologia, exclusivamente, não é capaz de explicar nenhuma das duas configurações apresentadas.

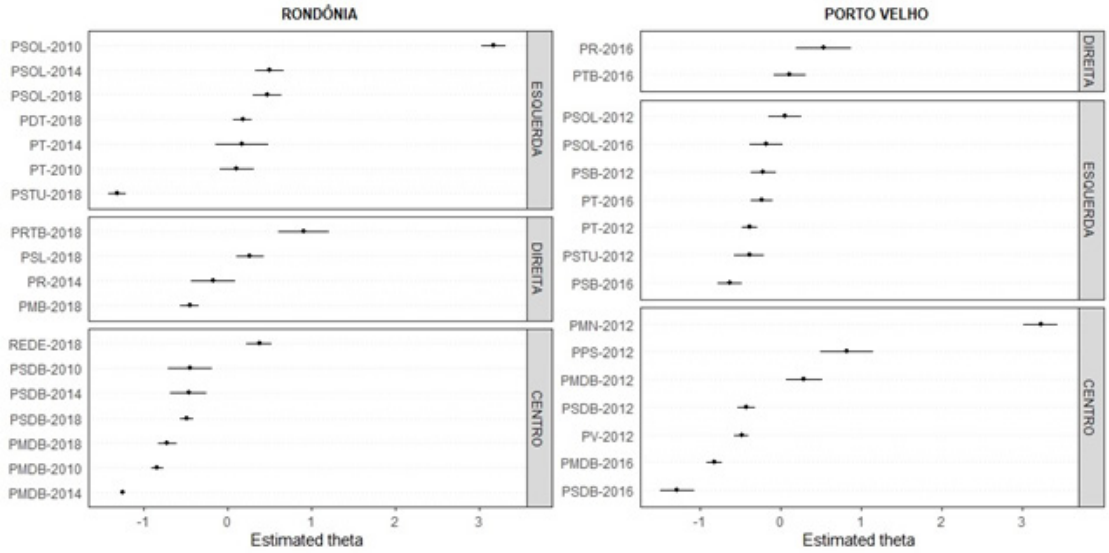
Gráfico 24 – Competição programática no Rio Grande do Sul (2010-2018)



Fonte: elaboração própria.

Em Rondônia, o cenário não é tão distinto. Nas duas esferas consideradas, as posições extremas de ambos os polos são ocupadas por partidos de um mesmo campo ideológico (Gráfico 25). No nível estadual, isso se trava entre os partidos de esquerda, especificamente PSOL e PSTU. Todavia, no municipal, a polarização acontece entre os de centro, com PMN de um lado e PSDB (2016), de outro. Nas eleições estaduais, partidos de direita e de centro também se dispersam entre os dois polos, como visto no Rio Grande do Sul. Já nas disputas para a prefeitura de Porto Velho, a esquerda se dispõe em forma de reta muito pouco inclinada, indicando concentração das legendas em um espaço mais próximo do centro. Na direita, PTB tem localização central, enquanto o PR se posiciona mais ao extremo. Novamente, essas não parecem ser competições programáticas organizadas ideologicamente.

Gráfico 25 – Competição programática em Rondônia (2010-2018)



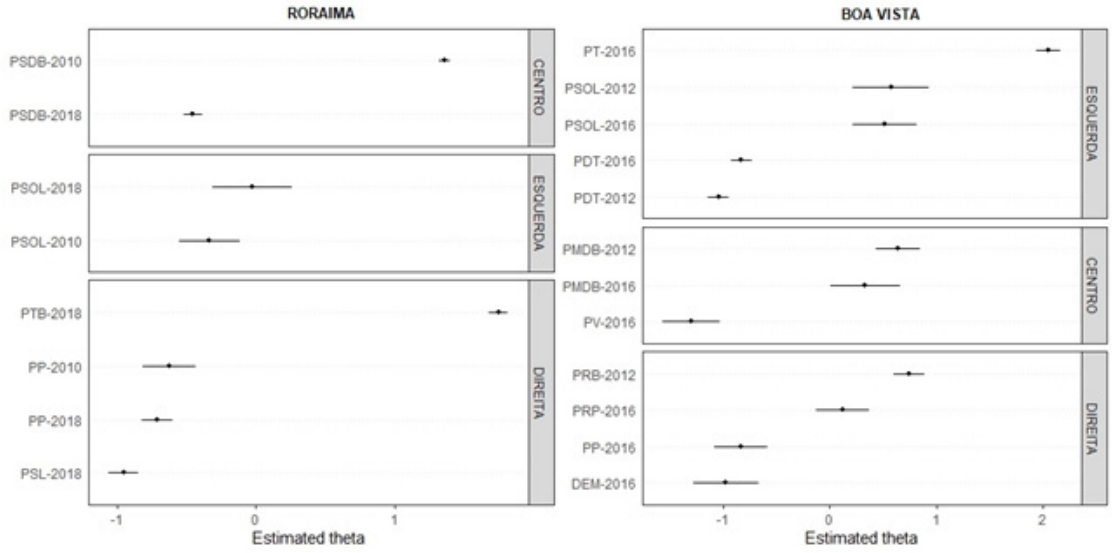
Fonte: elaboração própria.

Esse resultado se repete nas disputas em Roraima, onde PSOL ocupa as posições mais centrais no âmbito estadual e há dispersão entre partidos de um mesmo campo ideológico no nível municipal (Gráfico 26). Nas eleições para governador, o PSDB polariza entre programas registrados em 2010 e em 2018, assim como há oposição na direita entre PTB, de um lado, e PP e PSL, de outro. Inclusive, é importante notar que são especificamente PTB e PSL que possuem as posições mais extremas de todo o espectro programático – que não pode ser considerado ideológico – na esfera estadual. No nível local, tais posicionamentos são ocupados por PT e PV. Mas também há uma série de inconsistências na configuração apresentada nas disputas pela prefeitura de Boa Vista, entre as quais destaco a dificuldade em identificar quais seriam os polos de esquerda e de direita, dada a dispersão das legendas em cada campo.

Em Santa Catarina, o cenário é distinto – todavia sem resultar no alinhamento ideológico da competição programática (Gráfico 27). Na esquerda, estão PCO, PSTU, PCB e PSOL (2018, 2014), na arena estadual, e apenas o PSTU, na municipal. O partido mais próximo do centro nos dois casos é o PSOL, em 2010 e em 2012. Todos as demais agremiações se situam na direita, o que inclui PPL, PT, PEN, DEM, PSD, PRP, PSL, PP, PV, REDE, PSDB e PMDB, no nível estadual, e PSOL, PPL, PSB, PCdoB, PSD, PEN, PP e PMDB, no municipal.

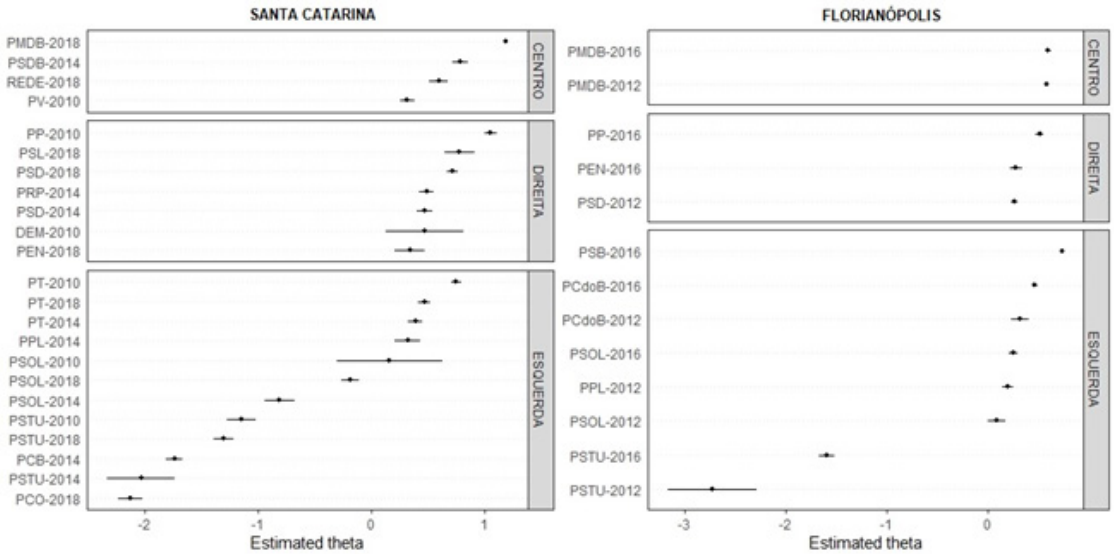
A distribuição das posições dos partidos de esquerda ao longo dos dois polos programáticos também acontece no estado de São Paulo (Gráfico 28). Além disso, agremiações de centro e de direita apresentam localização semelhante entre si, como visto em outros estados. Nesse panorama, apenas PCO, PSTU, PCB e PSOL estão à esquerda, com os demais partidos se dispersando entre o centro e a direita. As duas posições mais extremas de todo o espectro também se enquadram na esquerda. A disputa para a prefeitura de São Paulo parece muito alinhada a essa configuração, com a diferença que, nessa arena, partidos de centro também se posicionam à esquerda, especificamente o PMDB (2012). No nível local, as posições extremas não se concentraram na esquerda, dividindo-se ente PCO e REDE.

Gráfico 26 – Competição programática em Roraima (2010-2018)



Fonte: elaboração própria.

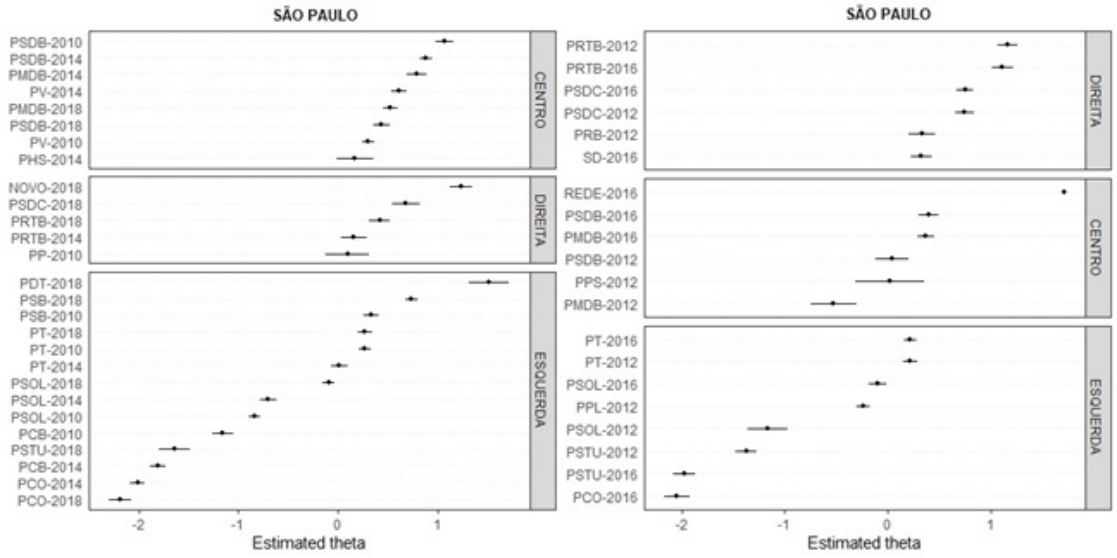
Gráfico 27 – Competição programática em Santa Catarina (2010-2018)



Fonte: elaboração própria.

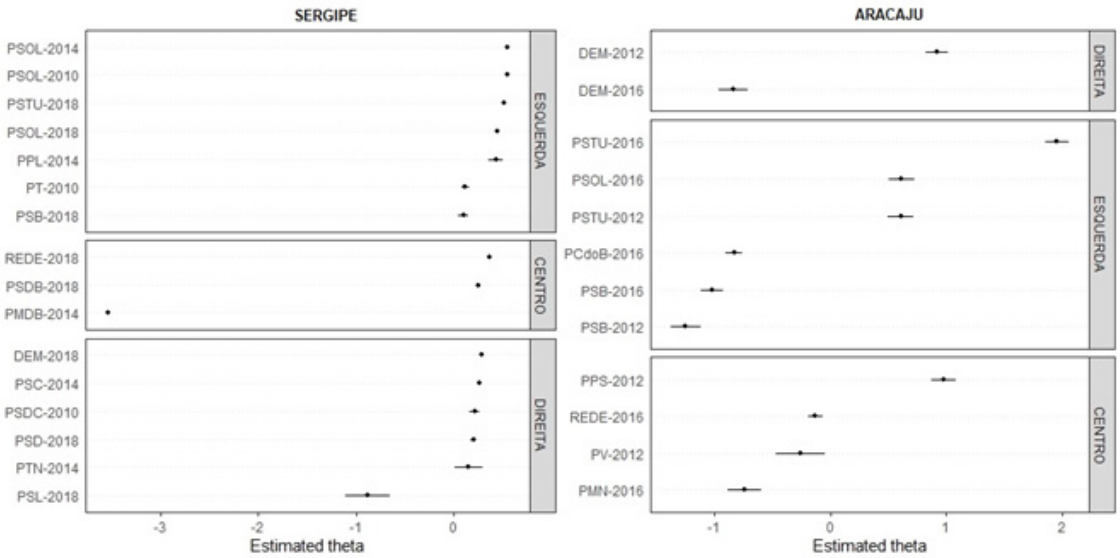
Sergipe é um dos casos em que há pouca divergência – e, portanto, competição programática – no nível estadual (Gráfico 29). Isso pode ser identificado com a posição da maioria dos partidos quase em uma reta vertical próxima ao centro. Nesse cenário, os programas um pouco mais distintos são os do PT e PSB e, de maneira mais acentuada, do PSL e do PMDB. As posições obtidas nas disputas pela prefeitura de Aracaju, por sua vez, apresentam maior divergência. Todavia, a dimensão ideológica parece ser insuficiente para explicar essa configuração, já que partidos de todas as classificações ideológicas se encontram dispersos no espectro programático.

Gráfico 28 – Competição programática em São Paulo (2010-2018)



Fonte: elaboração própria.

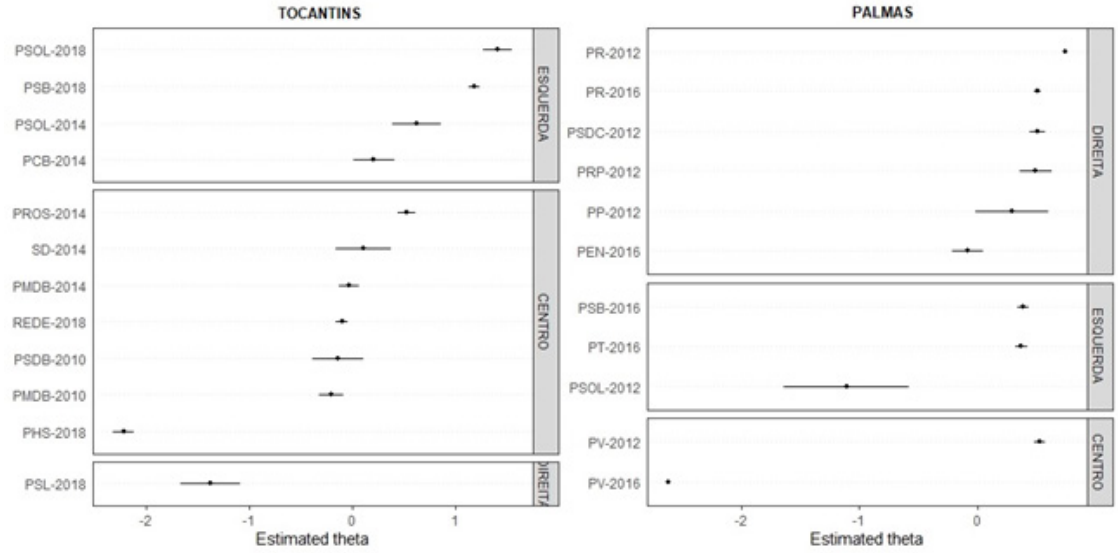
Gráfico 29 – Competição programática em Sergipe (2010-2018)



Fonte: elaboração própria.

As eleições para o governo estadual em Tocantins foram as que, junto a Goiás, apresentaram maior alinhamento ideológico (Gráfico 30). Nesse sentido, é possível observar que PSOL, PSB e PCB se localizam na esquerda; PMDB, PSDB, REDE e SD no centro; PSL na direita. Apenas duas posições podem ser consideradas mais inconsistentes com relação à classificação ideológica: a do PHS, mais à direita que o PSL, e a do PROS, na centro-esquerda. É importante observar que o PCB, apesar que situado na esquerda, possui localização muito próxima ao centro. Nível municipal, a divergência entre os partidos se reduz, fazendo com que se posicionem quase em um único polo, entre os valores 0 e 1. Com exceção do PV (2016) e do PSOL, que apresentam localização mais extrema em um mesmo lado, PR, PSDC, PRP, PP, PEN, PSB, PT e PV (2012) compartilham o mesmo espaço programático que, por esse motivo, parece se organizar a partir de outras dimensões que não a ideológica.

Gráfico 30 – Competição programática no Tocantins (2010-2018)



Fonte: elaboração própria.

O quadro abaixo resume os resultados obtidos, agrupados em três categorias. A primeira se refere à organização *ideológica* da competição programática, ou seja, aos casos em que as posições espaciais foram consistentes com a classificação dada pela literatura. A categoria oposta é a que chamo de *não ideológica*. Com isso, não quero dizer que a ideologia não tem relação com esses arranjos, mas que, quando analisados através do método proposto e como dimensão exclusiva, a localização dos partidos é significativamente distinta daquelas atribuídas pela literatura. Isto é, essa variável não tem força explicativa. Por fim, uma categoria intermediária, de competições programáticas *parcialmente ideológicas*, nas quais a ideologia parece determinar a posição dos partidos, porém não exclusivamente. Nesse sentido, seria necessário incluir outras dimensões à análise. Além disso, também classifiquei os casos com relação ao alinhamento entre níveis, ou seja, se as configurações obtidas nas arenas estadual e municipal eram mais ou menos distintas.

Quadro 4 – Resumo da configuração ideológica da competição programática

Competição programática estadual			Competição programática municipal			Relação entre níveis	
Ideológica	Parcialmente ideológica	Não ideológica	Ideológica	Parcialmente ideológica	Não ideológica	Alinhada	Não alinhada
GO	AM	AC	BA	AC	AL	AL	AC
TO	AP	AL	PB	MA	AM	AP	AM
	BA	DF		MT	AP	CE	BA
	CE	ES		MS	CE	ES	GO
	MA	MT		MG	ES	MA	MT
	MS	PA		PE	GO	MS	PR
	MG	PE		RJ	PA	MG	PI
	PB	RJ		RS	PR	PA	RS
	PR	RN		SP	PI	PB	RR
	PI	RS			RN	PE	SC
	SP	RO			RO	RJ	SE
		RR			RR	RN	TO
		SC			SC	RO	
		SE			SE	SP	
					TO		

Fonte: elaboração própria.

De maneira geral, os resultados obtidos mostram que poucas disputas, apenas dois casos e cada nível, são ideologicamente organizadas. Na arena estadual, no entanto, se esses cenários são somados aos de competições programáticas parcialmente ideológicas, eles constituem quase metade dos casos, totalizando 13 estados. Os demais (14) não apresentaram congruência entre a posição atribuída aos partidos e suas classificações ideológicas. No nível local, esse número é um pouco maior, 15 em 26; nas outras 9 capitais, a competição programática foi organizada pela ideologia apenas de modo parcial. Com relação ao alinhamento entre as duas arenas de disputa, 14 em 26 estados apresentaram configurações semelhantes da competição programática.

Esses apontamentos confirmam que não é o fato de os partidos distinguirem o conteúdo de seus programas de governo que essa divergência estará ideologicamente organizada. Nas eleições presidenciais, a congruência ideológica se deu de maneira mais acentuada do que nas demais esferas, com um pouco mais de força no nível estadual do que no municipal. Isso sugere que a competição, na maioria das vezes, não pode ser explicada de maneira unidimensional, em uma escala esquerda-direita, mas que outras dimensões parecem impactar. Nesse sentido, uma delas pode ser a competitividade da disputa ou dos próprios partidos, como foi apontado pela *teoria da saliência*, visto que, na maioria dos casos, pequenos partidos de esquerda, como PSTU, PCO, PCB e PSOL pareceram exercer uma força centrípeta entre seus colegas de campo ideológico. Assim, as posições mais centrais na esquerda se referiam a legendas, em geral, mais competitivas, como PSB e PT.

No que se refere ao método, o *wordfish* parece ter levado a classificações mais consistentes do que as obtidas através da codificação MARPOR, mesmo em casos em que ela foi adaptada ao contexto brasileiro. No entanto, se por essa técnica a maioria dos partidos foi considerados de esquerda ou centro-esquerda, o *wordfish*, em alguns casos, como o de Palmas, por exemplo, posicionou os partidos mais na centro-direita. Embora seja evidente a necessidade de novos exercícios, incluindo novas dimensões, e talvez elaborados através de técnicas mais

sofisticadas, que avancem para além da frequência de palavras isoladas (*bag-of-words*), o *wordfish* atua como um eficiente ponto de partida, indicando caminhos para que isso se realize.

6. Considerações finais

Os programas de governo são, por natureza, instrumentos de posições políticas, já que constituem em uma relação de problemas (e soluções) da ordem do dia, que devem ser enfrentados politicamente. Apesar disso, o uso desses objetos para a classificação ideológica dos partidos não está isento de desafios e contradições, como procurei demonstrar ao longo deste trabalho.

De modo geral, há duas questões principais relacionadas ao próprio conceito de ideologia que acabam impactando na sua mensuração através das plataformas eleitorais. A primeira se relaciona com a dimensionalidade do espaço político. Uma vez que esquerda e direita conformam uma mesma dimensão, elas são suficientes para explicar as posições dos partidos? Nesse contexto, há um dilema persistente entre reduzir o espaço político para apreendê-lo, sem que essa redução elimine características essenciais, comprometendo a análise. O segundo ponto se refere ao conteúdo substantivo das categorias esquerda e direita e como ele podem variar em diferentes contextos.

No Brasil, essas questões ainda têm avançado pouco no que se refere à relação entre ideologia e programas de governo. Como apresentado, além de serem escassos os estudos que se voltaram para esse tópico, os que o fizeram empregaram, em sua maioria, a codificação manual *MARPOR*. A combinação entre categorias fixas e o conteúdo substantivo da escala utilizada por eles tem levado a uma classificação problemática das agremiações brasileiras, com predominância quase exclusiva dos partidos na esquerda.

Tal classificação parece evidenciar, também, um problema metodológico. Por isso, neste trabalho, propus um exercício de análise dos programas de governo a partir de uma técnica automatizada de análise de texto que atribui posições espaciais a partir da frequência das palavras usadas pelos candidatos. Embora essa não seja a única técnica possível e possua limitações, sua utilização se justifica pelo fato de ideologia ser, intrinsecamente, um conceito espacial.

A análise dos programas de governo de candidatos a cargos executivos no Brasil desde 2010, nas diferentes arenas de disputa, ofereceu um panorama mais complexo – ainda que imperfeito – quando comparado aos estudos que empregaram a codificação MARPOR. Por trás, há um pressuposto de ordem teórica, de que a competição programática não necessariamente se converte em uma disputa ideológica, embora uma disputa ideológica seja sempre programática. Em outras palavras, se os candidatos competem programaticamente, eles elaboram propostas distintas de seus oponentes. Porém, elas podem se organizar ou não ideologicamente, como poderiam se organizar ou não partidariamente, dado um conjunto de disputas, e assim por diante. Os achados deste artigo reforçam esse argumento. Para as eleições presidenciais consideradas, houve organização ideológica da disputa. Todavia, essa força estruturadora perdeu parcialmente o impacto no nível estadual e, um pouco mais, na arena local. Isso significa que, nessas esferas, nem sempre a posição atribuída aos partidos através de suas plataformas eleitorais convergiu com a classificação ideológica da literatura.

Embora esse achado não seja trivial, há muitas questões que precisam ser mais profundamente observadas, como as outras dimensões e variáveis que poderiam determinar a posição dos partidos no espectro programático. Nesse sentido, os resultados deste estudo apontam para a necessidade de relacioná-las com dados eleitorais, como o grau de competitividade das disputas e

das legendas. Afinal, como salientado, partidos menos competitivos parecem exercer uma força centrípeta sobre os demais em um mesmo campo ideológico.

Da perspectiva do método, esforços futuros devem ser capazes de discutir os limites do uso do *wordfish*, tanto pela sua unidimensionalidade quanto por considerar apenas a frequência de palavras isoladamente, sem ponderar outras características, como ordem e sentido. Dada a importância do tema, sobretudo em contextos de proeminência da nova direita, estão postos os principais e próximos desafios de uma agenda interessada em compreender a relação entre competição programática e ideologia no Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

ALDRICH, John. (1983), "A downsian spatial model with party activism". *American Political Science Review*, vol. 77, no. 4: 974–990.

ALESINA, Alberto. (1988), "Credibility and policy convergence in a two-party system with rational voters". *The American Economic Review*, vol. 78, no. 4: 796–805.

ALLAN, James; SCRUGGS, Lyle. (2004), "Political partisanship and welfare state reform in advanced industrial societies". *American Journal of Political Science*, vol. 48, no. 3: 496-512.

SALLES, Nara. (2020), "Do paradoxo à competição: o lugar da dimensão programática nas disputas eleitorais". Revista Brasileira de Ciência Política, no. 32: 93-134.

SALLES, Nara; GUARNIERI, Fernando. (2019), "Estratégia eleitoral nos municípios brasielrios: componente programático e alinhamento partidário". Revista de Sociologia e Política, vol. 27, n. 72: e001.

BABIRESKI, Flávia. (2014), "As diferenças entre a direita do Brasil, Chile e Uruguai: análise dos programas e manifestos partidários". *Paraná Eleitoral*, vol. 3, no. 1: 171–198.

BENOIT, Kenneth; LAVER, Michael. (2006), *Party Policy in Modern Democracies*. New York, Routledge.

BOBBIO, Norberto. (1995). *Direita e esquerda: razões e significados de uma distinção política*. São Paulo, Ed. UNESP.

BUDGE, Ian; FARLIE, Dennis. (1983), *Explaining and predicting elections: issue effects and party strategies in twenty-three democracies*. London/Boston, Unwin Hyman.

CHAPPELL JR., Henry.; KEECH, William. (1986), "Policy motivation and party differences in a dynamic spatial model of party competition". *American Political Science Review*, vol. 80, no. 3: 881–899.

CODATO, Adriano; BERLATTO, Fábia; BOLOGNESI, Bruno. (2018), "Tipologia dos políticos de direita no Brasil: uma classificação empírica". *Análise Social*, no. 229: 870-897.

DE SIO, Lorenzo; WEBER, Till. (2014) "Issue yield: A model of party strategy in multidimensional space". *American Political Science Review*, vol. 108, no. 4: 870–885.

DOWNS, Anthony. (1957), *An economic theory of democracy*. New York, Harper.

GRIMMER, Justin; STEWART, Brandon. (2013), "Text as data: the promise and pitfalls of automatic content analysis methods for political texts". *Political Analysis*, vol. 21, no. 3: 267–297.

GROFMAN, Bernard. (2004), "Downs and two-party convergence". *Annual Review of Political Science*, vol. 7: 25-46.

HOPKINS, Daniel; KING, Gary. (2010), "A method of automated nonparametric content analysis for social science". *American Journal of Political Science*, vol. 54, no. 1: 229-247.

IZUMI, Maurício; MOREIRA, Davi. (2018), "O texto como dado: desafios e oportunidade para as ciências sociais". *BIB*, vol. 2, no. 86: 138-174.

JORGE, Vladimyr; SILVA, Mayra; FARIA, Alessandra; FERREIRA, Ana Rita. (2018), "Análise dos programas eleitorais dos candidatos a presidente em 2014: o posicionamento ideológico do PT e do PSDB". *Revista de Sociologia e Política*, vol. 26, no. 67: 1–20.

KLINGEMANN, Hans-Dieter; VOLKENS, Andrea; BUDGE, Ian; BARA, Judith; MCDONALD, Michael. (2006), *Mapping Policy Preferences II: Estimates for Parties, Electors,*

and Governments in Eastern Europe, European Union, and OECD 1990-2003. Oxford, Oxford

University Press.

LAVER, Michael; BUDGE, Ian. (1992), *Party policy and coalition policy in Europe*. London, Macmillan.

LAVER, Michael; BENOIT, Kenneth; GARRY, John. (2003), "Extracting policy positions from political texts using word as data". *Political Science Review*, vol. 97, no. 2: 311–331.

LEHMANN, Pola; MATTHIEß, Theres; MERZ, Nicolas; REGEL, Sven; WERNER, Annika. (2019), *Manifesto Corpus*. Version: 2019b. Berlin: WZB Berlin Social Science Center.

LOWE, Will. (2016), "Scaling things we can count". Working paper, *American Political Science Association meeting.*

MADEIRA, Rafael; VIEIRA, Soraia; TAROUCO, Gabriela. (2017), "Agendas, preferências e competição: PT e PSDB em disputas presidenciais". *Caderno CRH*, vol. 30, no. 80: 257–273.

MONROE, Burt; COLARESI, Michael; QUINN, Kevin. (2008), "Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict". *Political Analysis*, vol. 16, no. 4: 372-403.

MONROE, Burt; SCHRODT, Philip. (2008), "Introduction to the special issue: The statistical analysis of political text". *Political Analysis*, vol. 16, no. 4: 351–355.

OLIVEIRA, Augusto. (2011), "Os discursos programáticos dos partidos políticos na América Latina: identidade de classe e política econômica". *Mediações*, vol. 16, no. 1: 201–221.

OLIVEIRA, Augusto. (2014), *Representação Programática em 16 Democracias Presidencialistas: América Latina, 2000-2010*. 226 p. Tese (Doutorado em Ciência Política) – Universidade Federal do Rio Grande do Sul, Porto Alegre.

PETROCIK, John. (1996), "Issue ownership in presidential elections, with a 1980 case study". *American Journal of Political Science*, vol. 40, no. 3: 825–850.

REPASS, David. (1971), "Issue salience and party choice". *American Political Science Review*, vol. 65, no. 2: 389–400.

RIKER, William. (1982), "The two-party system and Duverger's Law: An essay on the history of political science." *American Political Science Review*, vol. 76, no. 4: 753–766.

ROBERTSON, David. (1976), *A Theory of Party Competition*. New York, Wiley.

ROEMER, John. (2001), *Political competition: theory and applications*. Cambridge, Harvard University Press.

SARTORI, Giovanni. (1976), *Parties and Party Systems: a framework for analysis*. Colchester: ECPR Press.

SCHMITTER, Philippe; KARL, Terry. (1991), "What democracy is... and is not". *Journal of Democracy*, vol. 3, no. 2: 75–88.

SLAPIN, Jonathan; PROKSCH, Sven-Oliver. (2008), "A scaling model for estimating time-series party positions from texts". *American Journal of Political Science*, vol. 52, no. 3: 705-722.

STOKES, Donald. (1963), "Spatial models of party competition". American Journal of Political Science, vol. 34, no. 2: 565–598.

TAROUCO, Gabriela. (2011), "Brazilian parties according to their manifestos: Political identity and programmatic emphases". *Brazilian Political Science Review*, vol. 5, no. 1: 54–76.

TAROUCO, Gabriela; MADEIRA, Rafael. (2013), "Partidos, programas e o debate sobre esquerda e direita no Brasil". *Revista de Sociologia e Política*, vol. 21, no. 45: 149–165.

VOLKENS, Andrea; LEHMANN, Pola; MATTHIEß, Theres; MERZ, Nicolas; REGEL, Sven; WERNER, Annika. (2016), *The Manifesto Project Dataset - Codebook.* Manifesto Project

(MRG/CMP/MARPOR). Version 2016a. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

WITTMAN, Donald. (1983), "Candidate motivation: a synthesis of alternative theories". *American Political Science Review*, vol. 77, no. 1: 142–157.

Ideology and Parties in Brazil: considerations and methods using party

Nara Salles - State University of Rio de Janeiro

Resumo

O objetivo deste artigo é discutir a classificação ideológica dos partidos brasileiros a partir de seus programas de governo. Para isso, organizo a exposição a partir de dois eixos. O primeiro, de caráter teórico, debate se e de que maneira esses documentos podem ser utilizados como objetos de mensuração de posições ideológicas, sobretudo a partir das contribuições da teoria da saliência da competição partidária. Nesse âmbito, apresento uma síntese das classificações já elaboradas para o caso brasileiro por meio dos programas de governo. Os limites dessas categorizações dão origem ao segundo eixo da discussão proposta, de caráter metodológico, e se volta para a apresentação de técnicas de análise automatizada de texto, alternativas ao esquema de codificação manual elaborado pelo Manifesto Project/MARPOR. Dando corpo às questões levantadas ao longo do artigo, realizo um exercício de análise de 889 plataformas eleitorais registradas por candidatos a cargos executivos no Brasil nos três níveis de disputa desde 2010. Para isso, utilizo a técnica de escalonamento, wordfish, e concluo que, embora haja casos nas diversas arenas em que a competição programática é organizada ideologicamente, há outras dimensões que podem estruturar essas disputas.

Palavras-chave: Análise automatizada de texto; Ideologia; Partidos; Programas de governo; Wordfish.

Abstract

This article aims to discuss the ideological classification of Brazilian parties, based on their government programs. For this, I organize the exhibition based on two axes. The first, from a theoretical feature, debates whether and how these documents can be used as objects of measurement of ideological positions, especially from the contributions of the salience theory of party competition. In this context, I present a summary of the classifications already developed for the Brazilian case through government programs. The limits of these categorizations give rise to the second axis of discussion, from a methodological feature, and turns to the presentation of different automated text analysis techniques, alternatives to the manual coding scheme developed by the Manifesto Project/MARPOR. Giving substance to the questions raised throughout the article, I carry out an analysis exercise of 889 electoral platforms registered by candidates for executive positions in Brazil at the three levels of dispute since 2010. For this, I use the scaling technique, wordfish, concluding that, although there are cases in the various arenas in which the programmatic competition is organized ideologically, there are other dimensions that can structure these disputes.

Keywords: Automated text analysis; Government programs; Ideology; Parties; Wordfish.

1. Introduction

Left and right are important classificatory concepts and instruments in the political and social sphere. Although ideology has gone through a reevaluation that signaled its weakening after the fall of the Berlin wall and the rise of post-materialistic issues, many countries, such as Brazil, saw emerging over the past few years the strength of the (new) right, indicating that these are still crucial definitions to understand political agendas and positions. But is it possible to ideologically classify parties by the electoral manifestos presented by their candidates?

To the *salience theory of party competition*, parties run with selective emphases on different political proposals compiled on their manifestos or electoral platforms (BUDGE; FARLIE, 1983; ROBERTSON, 1976). If the analysis of these documents enables knowledge of these emphases, are they sufficient to classify parties on a left-right scale? Which categorizations have been proposed thus far for the Brazilian case, using electoral manifestos as objects? What do these classifications say?

This paper aims to elaborate on and discuss each of these questions. To do so, a few challenges are posed related to the ideological classification from the perspective of programmatic competition, that is, the divergence among the political proposals from the candidates. These issues are not new; however, they persist in the debate between electoral competition and ideology.

First, if we take left and right as categories defined a priori, especially historically, we run the risk of building a rigid perspective, that inhibits the identification of multidimensional political positions and one-off strategies, emergent from specific elections and contexts. On the other hand, if we abandon the ideological spectrum as analytical instrument, we can incur in the error of ruling out more general political preferences, hindering comprehension and predictions of new debates and phenomena, such as the election of Bolsonaro.

From a methodological perspective, the ideological classification of Brazilian parties using manifestos has led to curious results, such as the near inexistence of right-wing parties. One explanation for this may be related to the method used by the studies, the coding by the *Manifesto Project/MARPOR*. Among a series of limitations, this technique works with fixed left and right categories, applicable to the more than 45 democracies the project focuses on since 1979. That is, it neglects the relative configurations these concepts may take on, not only in space, but in time. How is it possible to move forwards?

After presenting and discussing these questions, I propose as an exercise the analysis of 889 electoral manifestos registered by candidates to executive office in Brazil since 2010, in the three different levels of competition. To do so, I use the automated text analysis technique *wordfish*, which attributes spatial positions to parties given the words used in their electoral platforms (SLAPIN; PROKSCH, 2008). My main argument is that not always is a programmatic competition organized ideologically and, therefore, this becomes an empirical research problem.

In this paper, I investigate if that dimension is the one that structures elections for the presidency, state governments, and mayoral seats in the capitals, comparing the positions attributed to parties with the ideological classification highlighted by the literature. In this analysis, I am interested not only in the position of the parties in a programmatic spectrum and, possibly, ideological, but also the variation or constancy of this arrangement in the different levels of competition.

The sections of this paper are organized thusly: in the first, I discuss left and right in the perspective of com-

petition, reflecting if the ideological classification can be constructed from manifestos. In the second, I synthesize the results studies that have proposed to classify Brazilian parties through these documents have come to and their implications. In the third, I present alternative techniques to the coding by the *Manifesto Project/MARPOR* with automated text analysis and justify the use of *wordfish*. In the fourth, I describe and analyze the spatial positions attributed to parties in each election when it comes to the ideological dimension. Finally, I present a final discussion, highlighting the main points for a research agenda that articulates ideology an electoral competition in Brazil.

2. Ideology, electoral competition, and the place of election manifestos

Before starting the specific debate on how Brazilian parties have been ideologically classified from their electoral manifestos[1], there is an issue that cannot be side-stepped: is it possible to measure ideological positions through these documents? This question directly leads us to the nature of these platforms and the contests they are inserted in.

There is no democracy without competition (SCHMITTER; KARL, 1991), but there is no single way of understanding it. The most explored and enduring path in electoral studies, investigates the configurations of these contests from, in particular, the number of competitors and their relative positions, in order to understand their institutional and social conditioning factors (DOWNS, 1957; RIKER, 1982; ROEMER, 2001). Another approach, partly derived from the previous one, seeks to identify the content of the competition, the issues around which it is structured, what are the political preferences of the actors involved and how they are mobilized strategically (BUDGE; FARLIE, 1983; ROBERTSON, 1976).

It is not that ideology would not have an impact in determining the relative positions which interest spatial models. The problem is that the classic Downsian proximity model would lead to what I call *programmatic paradox* (SALLES, 2020). Ideology is, for Downs (1957), an informational shortcut to identify party preferences with smaller cost to the electorate. However, because parties have electoral maximization as their single goal, and knowing that voters make their decision based on the logic of proximity, the centripetal movement, aiming for the median voter, would be inevitable. In other words, because proximity among the political positions is what defines an election and all want to win, they would have incentives to converge[2].

From this point of view, it would make little sense to investigate ideological positions from electoral manifestos. After all, the centripetal movement imposed by the proximity model would make them converge more rather than differentiate. Nevertheless, the literature has highlighted that, in the real world, the Downsian model does not work so perfectly – at least, not without violating at least one of its assumptions. Grofman (2004) synthesizes each of them, presenting alternative approaches. For this paper, it is enough to adopt the idea that, although parties compete for votes, this does not constitute an exclusive goal or even an end in itself (ALESINA, 1988; CHAPPELL JR.; KEECH, 1986; SARTORI, 1976; WITTMAN, 1983).

1 Translation note: in the original, the author opted to reinforce the use of electoral manifestos, rather than party manifestos. The specific expression used would have caused confusion in English, as "government programs" are usually used by established governments as work plans, rather than their electoral platforms. Unless otherwise noted, manifestos will here be used as the electoral documents parties and candidates run on, rather than the documents parties register as their official position statement.

2 I do not intend to defend that there is no difference in Down's (1957) proposal when it comes to biparty or multiparty systems and the possession or not of complete information, because there is. These specificities do not invalidate the conclusion that it, roughly, points to party convergence.

Robertson (1976) was a pioneer in dealing with these issues from the perspective of manifestos. Inspired by analyses in journalistic reports, he classified British parties based on the salience given in their electoral manifestos between 1922 and 1974, measuring the amount of times a given area was mentioned in relation to the total number of proposals in each document. As a result, he pointed to a general characteristic: what guided the parties was the selective emphasis of different policies and not a direct conflict. That is, Robertson (1976) identified that, instead of presenting distinct positions on the same theme, competitors highlighted issues that were more favorable to themselves and avoided mentioning policies that were salient to their opponents.

Later, his findings were corroborated, deepened, and systematized by Budge e Farlie (1983), giving rise to the *salience theory of party competition* which, basically, has the fundamental concept that parties compete through the selective emphasis of issues – which I call programmatic competition. In addition, for *salience theory*, the political positions that parties assume in these contests can be measured in the manifestos through the selective emphasis given by each competitor.

Consequently, yes, it is possible to extract political positions from manifestos (BENOIT; LAVER, 2006; BUDGE; FARLIE, 1983; GRIMMER; STEWART, 2013; KLINGEMANN et al., 2006; LAVER, BENOIT; GARRY, 2003; MONROE; SCHRODT, 2008). However, is measuring political positions equal to measuring ideological positions? In other words: what does manifesto analysis measure? Is it, necessarily, ideology?

For *salience theory*, the selective emphasis on issues – and, therefore, the content of these documents – would be determined by three factors: (1) the success or consensus regarding policies; (2) issue ownership; (3) the competitiveness of the election. The first was presented by Robertson (1976), when arguing that there are issues that will be common to all parties, because it is a consensus in the population that they are relevant and necessary. Thus, he states:

> The only time when parties not only need not to be different, but cannot afford to be, is where one set of policies appears to be working [...]. That it is foolish to advocate the same policies as one's opponent when they fail, and foolish to differ when they succeed is a rule [...]". (ROBERTSON, 1976, p.130).

If good performance and popular support regarding policies act as a centripetal incentive on party positions, the idea of issue ownership shows that parties have specific agendas (BUDGE; FARLIE, 1983; DE SIO; WEBER, 2014; PETROCIK, 1996; REPASS, 1971). This means that parties associate previously, historically, and for long periods with certain issues and, that although such positions may be moderate, they must present some consistency, being responsive to the activity expectations of their most faithful voters (ALDRICH, 1983). In the context of the discussion proposed in this work, would issue ownership be related to ideology?

Finally, the last factor related to the emphases given by parties in their manifestos: the degree of electoral competitiveness. Becoming a salient issue attracts a liquid vote entry – for a given competitor and their opponent. Therefore, it is not strategic to mention policies emphasized by rivals, given the risk of favoring them. Issues may become a common field for emphases however, when elections are too close, in contests where the margin of victory is small. Although parties prefer non-competitive positionings closer to their preferences, the position occupied by them will depend on their electoral strength: where there is high voting, there is greater incentive for extreme positioning (BUDGE; FARLIE, 1983; ROBERTSON, 1976).

It is key to observe, in this perspective, that the manifestos are, by nature, instruments of political positioning but not only that. They are *strategic* objects of electoral competition. Because preferences are given endogenously – and not deterministically as Downs (1957) postulated –, the salience structure is not rigid, but dependent on the evaluation of leaders on which dimensions are more salient for the electorate at each election. This means that political positionings expressed in the platforms may not always be a portrait of the parties' ideological position. Consequently, manifestos may include contextual aspects of reality, personal elements on the candidates and their rivals, among others. It is specifically due to this that they are considered endogenous spaces of political competition.

The matter of strategy brings about a first challenge to think about ideology from manifestos. If issues are mobilized strategically, given the context of the agenda for each election, can any and all content proposed in the platforms be used for the parties' ideological classification? If so, which technique would ensure that changes in the parameters of the left, center, and right categories and their variants are incorporated in the analysis? These questions lead us to the more classic dilemma when it comes to ideology and political texts: can positionings be distributed in a single dimension?

This debate is not new, and it is one of the main critiques directed at the classic Downsian model, for example. Stokes (1963), in one of the first and most resounding criticisms of the author's work, argues that the electorate distributes their preferences with relation not to one, but several issues and, moreover, that they vary in relevance overtime and cannot be subsumed in a single axis, right-left. Hence the aforementioned importance of parties strategically identifying and mobilizing latent dimensions in the competition "[...] the skills of political leaders who must maneuver for public support in a democracy consist partly in knowing what issues dimensions are salient to the electorate or can be made salient by suitable propaganda". (STOKES, 1963, p. 372).

The problem is that, if competition cannot be understood unidimensionally, the left end right categories are so. If we return to the beginning of the use of these labels as synthesizers of political positionings, it is not difficult to realize that they are, by nature, unidimensional and spatial categories (BOBBIO, 1995). After all, in the French Revolution, the ones who sat at the left in the Estates-General meeting were the ones with an affinity for egalitarianism and social reforms, while the ones who sat on the right were closer to the aristocracy and conservatism. However, from the late 18th century to present day, what do the left and right categories indicate? Are they still useful to classify the political and social world?

If we have as reference the classic contribution by Bobbio (1995), the content of this categorization is centered, mainly, around the issue of equality. While the right recognized inequality as natural, limiting political action, the left was guided by the goal to overcome such order. Considering social spending, for instance, left-wing governments prefer redistributive measures, meanwhile center and right ones have a greater affinity with adjustment policies (ALLAN; SCRUGGS, 2004). This binarism, however, is mediated by variables such as liberty, tradition, hierarchy, participation, order, and autonomy (BOBBIO, 1995). Wouldn't those mediations be new dimensions of the political space related, but not reduced to, ideology?

Benoit and Laver (2012) seem convinced that it is so: "Even this drastically oversimplified notion of a 'left-right dimension' refers to two potentially separable issues: one concerns economic policy, a second con-

cerns state regulation of social behavior". (BENOIT; LAVER, 2012, p. 195). However, they also seem to believe that, if in theory there are more potential dimensions of distinction among political preferences and positions than ideological, in practice they may be correlated. This means that knowing someone's position on a given issue may be useful in predicting their preference regarding other themes. Thus, from an analytical standpoint, the following question emerges: how to use the left and right categories without significantly reducing the political space, but also without adding so many dimensions that it becomes impossible to comprehend it? This question is added to others such as: can the meaning of left and right be established a priori? Do these definitions vary according to the context and time of each election?

This discussion is vital, from theoretical and methodological perspectives, and is directly related to the ideological classification attributed to Brazilian parties through the analysis of manifestos. As I demonstrate, even the adapted use of a 30-year-old, preestablished codification for the analysis of any democracy has led to intriguing results. Then, if manifestos can be used to measure ideological positions, this classification is not assured by the simple use of the objects, but also demands the employ of adequate techniques.

3. Electoral manifestos and the ideological classification of Brazilian parties

These challenges that surround obtaining ideological positions through manifestos extend to the few studies they have done so for the Brazilian case until now. Among them are highlighted, especially, the contributions by Tarouco (2011) and Tarouco and Madeira (2013) given the largest volume, variety, and breadth of the documents analyzed. More recently, new studies were added to these that incorporate other presidential elections (JORGE *et al.*, 2018), that proposed to debate the issue of polarization (MADEIRA; VIEIRA; TAROUCO, 2017), or ventured into comparative approaches among manifestos from different countries in Latin America (BABIRESKI, 2014; OLIVEIRA, 2011). Among them, there is something in common: the almost absolute prevalence of Brazilian parties in the left of the ideological spectrum (Box 1).

Box 1 – Ideological classification of Brazilian parties with the MARPOR coding

Reference	Right	Left
Lehmann *et al.* (2019b)	PSL 2018 (+19,8)	PT 1989 (-35,6)
	PSDB 1989 (+19,8)	PSOL 2006 (-35,4)
		PT 1998 (-26,9)
		PV 2010 (-26,5)
		PT 2018 (-25,4)
		PRONA 1994 (-24,8)
		PT 2010 (-24,2)
		PPS 1998 (-24,1)
		PDT 2018 (-22,5)
		PT 2002 (-19,7)
		PT 1994 (-19,6)
		PT 2014 (-18,5)
		PT 2006 (-18,3)
		PSDB 1998 (-17,2)
		PSDB 2010 (-15,7)
		PPS 2002 (-15,7)
		PSDB 2002 (-15,4)
		PSDB 2006 (-13,3)
		PSB 2002 (-11,7)
		PSDB 2014 (-11,0)
		PSB 2014 (-11,0)
		PSDB 2018 (-2,7)
		PRN 1989 (-2,6)
		PSDB 1994 (-1,8)
Jorge *et al.* (2018)		PT 2014 (-0,13)
		PSDB 2014 (-0,06)
Babireski (2014)		PP (-15,2)
		DEM (-11,2)
Tarouco e Madeira (2013)	PFL 1995 (+6,5)	PDT 1979 (-37,3)
		PTB 1979 (-36,8)
		PT 1980 (-34,4)
		PTB 2001 (-27,2)
		PMDB 1981 (-25,0)
		PT 1990 (-22,0)
		PDS 1979 (-15,1)
		PP 2003 (-14,3)
		PPB 1995 (-14,1)
		PSDB 1988 (-13,6)
		PDT 1994 (-10,3)
		PMDB 1994 (-8,5)
		PFL fund. (-5,4)
		PFL 2005 (-3,6)
		PSDB 2001 (-0,4)

Source: the author.

The classification of Brazilian parties may cause some discomfort. According to these, there are practically no right-wing parties in Brazil. Therefore, even parties consensually considered to be on the right, such as PP, PFL/DEM, and PRONA find themselves in the left of the ideological spectrum. On the other hand, parties commonly located closer to the center, such as the PMDB and PSDB, have scores very similar to PT's, for example. These results may indicate that the content of the documents analyzed discussed other factors that were not ideological – at least, not an a predominant or determining way – or some methodological insufficiency, since the use of the codification from the *Manifesto Project/MARPOR* as a technique[3], albeit adapted, is a common characteristic of these studies.

In practice, measuring is done through the manual codification of *quasi-sentences*, or sections that contain only a political proposal, from one of the 56 categories, distributed across seven domains (Box 2). Some of them, as it is possible to observe, contain subcategories and the positive or negative direction of the proposal, but that is not a rule.

There is no doubt that the effort by *MARPOR* is fundamental in many ways. The first, already mentioned, by its supposed founder, postulates that party competition is structured from the selective emphasis of issues. Secondly, methodologically, by pointing out that political positions can be measured in manifestos, from these emphases. Lastly, regarding the compared perspective, due to its being an ample dataset of electoral platforms from several countries and periods in time.

Nevertheless, it is necessary to consider the limitations of the proposed codification, which has led to deficient, if not strange, ideological classifications of Brazilian parties (MADEIRA; TAROUCO, 2013). One first issue is that, although categorization of *quasi-sentences* should be exclusive, the scheme offered is not exhaustive, which makes classification too dependent on the evaluation of each specialist. This does not constitute, per se, a technical flaw, given that it can be fixed, for example, with a classification by more than one professional. In addition, there are no purely objective methods, they always involve some kind of choice. From this perspective, this possible arbitrariness must be considered just as a more evident limitation – and, consequently, positive because it is clear – then the ones involved in other techniques.

There is, however, considerable rigidity in the scheme used by *MARPOR*. On the one hand, this ensures comparability among several countries, with ample longitudinal breadth. However, on the other, it excludes a series of factors such as local contexts and more contextual elements of certain time periods. Pragmatically, the consequence is that the assumption that characterizes right and left in Europe extend to any and all places, at any time.

3 Founded in 1979 with the goal of measuring political positions from manifestoes, the now named Manifesto Research on Political Representation (MARPOR) analyzes documents from over 61 democracies, 769 elections, and 1,197 parties, including Brazilian ones. Available at <https://manifesto- project.wzb.eu/>. Last accessed on 25/06/2020.

Box 2 – Coding categories for manifestos from MARPOR

Domínio 1: External Relations
101 Foreign special relationships: Positive
102 Foreign special relationships: Negative
103 Anti-imperialism: Positive
  103.1 State centered anti-imperialism
  103.2 Foreign financial influence
104 military: Positive
105 military: Negative
106 Peace
107 Internationalism: Positive
108 European Community/Union or Latin America Integration: Positive
109 Internationalism: Negative
110 European Community/Union or Latin America Integration: Negative

Domain 2: Freedom and Democracy
201 Freedom and Human Rights: Positive
  201.1 Freedom
  201.2 Human Rights
202 Democracy
  202.1 General: Positive
  202.2 General: Negative
  202.3 Representative democracy: Positive
  202.4 Direct democracy: Positive
203 Constitutionalism: Positive
204 Constitutionalism: Negative

Domain 3: Political System
301 Decentralization: Positive
302 Centralization: Positive
303 Governmental and administrative efficiency: positivo
304 Political corruption: Negative
305 Political authority: Positive
  305.1 Political authority: party competence
  305.2 Political authority: personal competence
  305.3 Political authority: strong government
  305.4 Pre-democratic elites: Positive
  305.5 Pre-Democratic Elites: Negative
  305.6 Rehabilitation and compensation

Domain 4: Economy
401 Free market economy: Positive
402 Incentives: Positive
403 Market regulation: Positive
404 Economic planning: Positive
405 Corporatism/mixed economy: Positive
406 Protectionism: Positive
407 Protectionism: Negative
408 Economic Goals
409 Keynesian demand management: Positive
410 Economic Growth
411 Technology And Infrastructure: Positive
412 Controlled economy: Positive
413 Nationalism: Positive
414 Economic orthodoxy: Positive
415 Marxist analysis: Positive
416 Anti-growth economy and sustainability: Positive
  416.1 Anti-growth economy: Positive
  416.2 Sustainability: Positive

Domain 5: Welfare and Quality of Life
501 Environmental protection
502 Culture: Positive
503 Equality: Positive
504 Welfare state expansion
505 Welfare state limitation
506 Education expansion
507 Education limitation

Domain 6: Fabric of Society
601 National way of life: Positive
  601.1 General
  601.2 Immigration: Negative
602 National way of life: Negative
  602.1 General
  602.2 Immigration: Positive
603 Traditional morality: Positive
604 Traditional morality: Negative
605 Law and order
  605.1 Law and order: Positive
  605.2 Law and order: Negative
606 Civic mindedness: Positive
  606.1 General
  606.2 Bottom-up activism
607 Multiculturalism: Positive
  607.1 General
  607.2 Immigrant integration: Diversity
  607.3 Indigenous rights: Positive
608 Multiculturalism: Negative
  608.1 General
  608.2 Immigrant integration: Assimilation
  608.3 Indigenous rights: Negative

Domain 7: Social Groups
701 Labor groups: Positive
702 Labor groups: Negative
703 Agriculture and farmers
  703.1 Agriculture and farmers: Positive
  703.2 Agriculture and farmers: Negative
704 Middle class and professional groups: Positive
705 Underprivileged minority groups: Positive
706 Non-economic demographic groups: Positive

000 No meaningful category applies

Source: VOLKENS et al. (2016).

It is not true that the *MARPOR* coding necessarily leads to a unidimensional political space, given that the project has five different programmatic dimensions: (1) *planeco*, which measures state intervention in the economy; (2) *markeco*, which measures economic liberalization; (3) *welfare*, related to mentions of the welfare state; (4), *intpeace*, focused on peace and international relations; (5) *rile*, the left-right scale. Meaning that, although it has been applied to Brazil only for the ideological spectrum, it would be possible to aggregate it to the other dimensions.

This, however, would not solve the problem of what the *MARPOR* substantively considers as left and right and its applicability to several countries and periods. The *rile* scale, built from definitions by Laver and Budge (1992), does not include all the categories in the scheme. Using it, the ideological position of a party is given by the sum of the percentage of emphases of right minus the percentage of emphases of left categories (Box 3).

Box 3 – MARPOR rile scale

Right	Left
104 Military: Positive	103 Anti-imperialism: Positive
201 Freedom and human rights	105 Military: Negative
203 Constitutionalism: Positive	106 Peace: Positive
305 Political authority	107 Internationalism: Positive
401 Free market economy	202 Democracy
402 Incentives: Positive	403 Market regulation: Positive
407 Protectionism: Negative	404 Economic planning: Positive
414 Economic orthodoxy: Positive	406 Protectionism: Positive
505 Welfare state limitation	412 Controlled economy: Positive
601 National way of life: Positive	413 Nationalization: Positive
603 Traditional morality: Positive	504 Welfare state expansion
605 Law and order	506 Education expansion
606 Civic mindedness: Positive	701 Labor groups: Positive

Source: VOLKENS et al. (2016).

It is important to observe that, in this conception, the right includes almost all domains, with the exception of social groups, in particular categories related to the economy and the fabric of society. The left, however, does not encompass emphases in that last dimension, disregarding key themes of social behavior, commonly used for ideological classification, such as legalization of abortion, legalization of marijuana, among others. More importantly, the left does not include the category that for Bobbio (1995), is its defining one – equality. The categorization, which is not exempt from controversy, also excludes the issue of multiculturalism and considers proposals relative to Human Rights as right-wing issues.

Therefore, the issues in using the *rile* scale to classify Brazilian parties are not restricted to the fact that the *MARPOR* coding scheme unifies the concept of left and right among different countries and periods. There is also a limitation of a substantive nature when it comes to categories taken as reference to measure ideological position. Although the literature recognizes these inadequacies and proposes adaptations, this is the scale that has been used in Brazil to classify parties ideologically using manifestos. In addition, it has been adopted as an exclusive programmatic dimension. Consequently, the aforementioned absence of Brazilian parties and the right may be associated with these limitations.

In summation, it is not possible to apply the ideological codification proposed by *MARPOR* to the Brazilian case – at least, not without alterations. That conclusion

was highlighted by Tarouco and Madeira (2013), who remembered that, in Brazil, it would not be the issue of equality that would divide the left and right, but the means to achieve it. For the right, this would happen through reinforcing order via the state, meanwhile for the left, it would be through state distributive policies, concurrent with the contestation of state repression of social and political movements for equality. Thus, they reinforce the need to adjust the scale for Brazil, taking into account the ambiguities regarding the relative expectations of the state's role.

> The MRG scale, presented above, includes some categories that do not apply in the direction of the left-right distinction in Brazil. Due to the historic experience of military dictatorship and transition, the left incorporates claims that the MRG identifies as right-wing: freedom, human rights, and constitutionalism, for example. Similarly, defending peace as a general goal and internationalism, which were never part of the Brazilian left's main interests, are, in that scale, among the categories indicative of a left-wing position (TAROUCO; MADEIRA, 2013, p. 158, *our translation*).

In addition, the authors add that the defense of democracy, identified as left-wing, is common to all parties – or at least it was in 2013 – and that the categories indicated by *MARPOR* as right-wing, such as nationalism, traditional morality, and law and order are more related to conservatism.

The identification of this disconnection between the *rile* scale and what would, in fact, distinguish the left from the right in Brazil gave rise to adaptations of this categorization. Tarouco and Madeira (2013) argue that, within the *MARPOR* codification scheme, the category is indicative of the right would be: (1) Military: Positive; (2) Free market economy; (3) Incentives; (4) Economic orthodoxy; (5) Welfare state limitation; (6) Middle class and professional groups. The final position of the parties in the scale was given by the total of mentions in these categories minus the total obtained in the categories considered left-wing: (1) Market regulation; (2) Economic planning; (3) Controlled economy; (4) Marxist analysis; (5) Welfare state expansion; (5) Labor groups: Positive. This new measurement, however, did not result in a classification too different from the previous one: only the PFL's manifestos from 2005 and 1995 and the PSDB's from 2001 were categorized as right-wing (TAROUCO; MADEIRA, 2013)[4].

Oliveira (2014), on the other hand, proposed an ideological classification of Brazilian parties through the analysis of presidential elections manifestos between 1998 and 2014, mobilizing the variables in the Database of Programmatic Assertions (BAP, in Portuguese). The coding process is similar to the *MARPOR* scheme, however with different categories and distinct variable groupings. The author also structured his proposal in seven large axes: (1) Social Policy;

(2) Commercial Policy; (3) Fiscal Policy; (4) Public Spending Policy; (5) Financial Policy; (6) Monetary Policy; (7) Regulation Policy[5]. The ideological position of the parties was given by the frequency of assertions in the documents analyzed. In this study, Brazil also had a restricted right-wing field.

In a previous study, I analyzed 3,410 electoral platforms, registered by mayoral candidates in the 2016 elections in 1,335 Brazilian municipalities (SALLES; GUARNIERI, 2019). In this, I demonstrated that candidates compete programmatically, that is, they present proposals distinct from their opponents. However, I

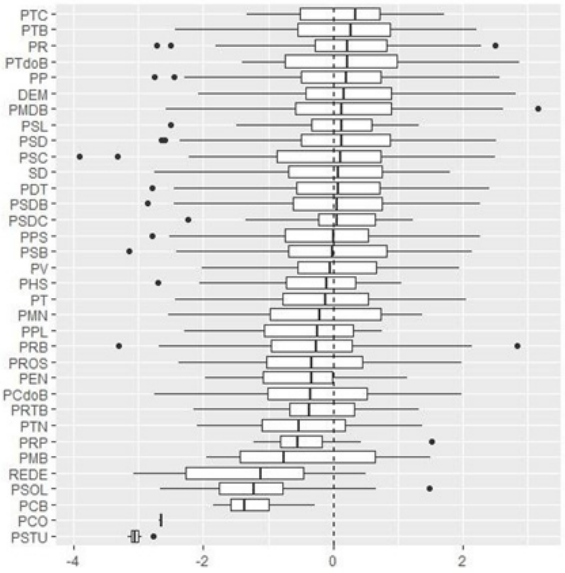
4 The complete classification, between the extreme left (-100) and the extreme right (+100) is: PDT 1979 (- 23,4), PTB 1979 (-22,9), PTB 2001 (-20,2), PT 1980 (-13,8), PT 1990 (-12,4), PDT 1994 (-12,0), PDS 1979 (-7,2), PP 2003 (-5,2), PPB 1995 (-5,2), PFL fund (-4,9), PSDB 1988 (-3,3), PMDB 1994 (-1,3), PMDB 1994 (-1,3), PFL
2005 (+6,0), PSDB 2001 (+6,5), PFL 1995 (+9,8) (TAROUCO; MADEIRA, 2013).

5 Para conferir cada categoria indicativa de esquerda e direita nesses eixos, consultar Oliveira (2014, p. 76).

highlighted that ideology is not a condition for competition around issues, in that ideological factors are programmatic, but these are not always ideological.

Despite this, the spatial positioning of parties from the frequency of words used in their manifestos does not permit to disregard the ideology variable (Graph 1).

Graph 1 – Distribution of proposals by party in the total of municipalities



Source: SALLES; GUARNIERI, 2019, p.17.

Taking into account the median position of parties in the set of municipalities in the sample, we have the following classification: PSTU, PCO, PCB, PSOL, REDE, PMB, PRP, PTN, PRTB, PCdoB, PEN, PROS, PRB, PPL, PMN, and PT, in the left; PV, PSB, PPS, PSDC, PSDB, PDT, and SD, in the center; PSC, PSD, PSL, PMDB, DEM, PP, PTdoB, PR, PTB, PTC, in the right. Obviously, there are inconsistencies in this possible classification, such as the PRB and the PRTB in the left and the PDT in the center. However, it does not exclude the right-wing field from the Brazilian programmatic spectrum.

Despite this, it is necessary to note that, although the median is positioned in one of the spatial poles, there is no concentration of the distribution of positions of one party in the majority of cases, with the exception of the PRP, PMB, REDE, PSOL, PCB, PCO, and PSTU. This can be related to the fact that the comparison was established over the total of municipalities disregarding any other sociodemographic, economic, or even political variable of the organization and strength of the parties in these locations.

In summation, all of these indications show that, although some studies have focused on the analysis of manifestos, there still unanswered questions when it comes to ideology: manifestos bring few ideological elements or are we having difficulties in measuring them?

4. Manifestos as data

There are several textual analysis techniques that enable the analysis of manifestos, with or without the goal of ideological classification. The first of which, already broadly mentioned, is manual coding, with the highlight for the scheme of categories and domains elaborated by *MARPOR*. Given the reasons already stated, I dedicate myself now to present alternative automated techniques for textual analysis.

It is fundamental to have in mind the four principles of quantitative text analysis list it by Grimmer e Stewart (2013). (1) Automated models may be considered completely wrong, but useful. This is due to the fact that they are not always capable of measuring the complex structures of language and their non-literal meaning, which requires that they be evaluated regarding the usefulness of the scientific and social task they proposed to fulfil. (2) These methods are not at all capable of replacing human skills, only enhancing them. Thus, it remains the researcher's task to guide the process and make the necessary decisions for the application of the models, in addition to interpreting their results. (3) There is no method globally more efficient for automated text analysis. Different problems and datasets lead to different goals, which may include: identifying words that distinguish the language of the group (LAVER, BENOIT; GARRY, 2003; MONROE; COLARESI; QUINN, 2008); proportion of documents that fits into predetermined categories (HOPKINS; KING, 2010); spatial positions of actors obtained through the texts (LAVER; BENOIT; GARRY, 2003; SLAPIN; PROKSCH, 2008). (4) Whatever model adopted by the researcher, it is fundamental that it is duly validated.

Automated text analysis may be divided into two groups: classification and scaling (GRIMMER; STEWART, 2013). The former classifies documents into categories, which can be pre-defined or latent. That is, one can create dictionaries and, from those, code documents, in a kind of automated *MARPOR* technique; or it is possible to let these categories be automatically extracted from the document, in what is called topic models[6]. Scaling techniques offer spatial positions measured from the analysis of the documents. Since I am interested in an ideological classification that, as stated, is intrinsically spatial, I opt in this paper for a scaling model.

Although Grimmer and Stewart (2013) guarantee that there is no globally efficient method for automated text analysis, Lowe (2016) argues that there is essentially only one way to scale textual data[7]. For him, the common logic for dimensioning countable items implies, explicitly or implicitly, the foundation of *salience theory* (BUDGE; FARLIE, 1983). Specifically, this means that all methods of political scaling from textual data reflect the idea that positions contained in documents are reconstructions of low dimension, limited by patterns of *selective emphases* (LOWE, 2016).

When it comes to automated text analysis, there are two main scaling techniques: *wordscores* (LAVER; BENOIT; GARRY, 2003) and *wordfish* (SLAPIN; PROKSCH, 2008). The first is supervised, that is, it requires a training set and, due to this, was not considered for this paper. After all, the substantive content of left and right is not consensus. Here, I employ *wordfish*, which attributes spatial positions in an unsupervised manner, from the frequency of the words in the documents. In the model, the authors assume that the actors are located in a political space of low dimension and at the positions established there impact the rate in which words are used in their texts. It works thusly: it is assumed that the distribution of each word p in a document i (y_{ip}) is originated from a Poisson process – that is, it has only one parameter, which simultaneously represents the mean and variance –,

$$y_{ip} Poisson(\lambda_{ip})\ \lambda_{ip} = exp\ (\alpha_I + \psi_j + \beta_j * \omega_I)$$

in which λ_{ip} is the count of the word p in the manifesto of party i; α is a set of fixed effects by document; ψ is a set of fixed effects by words; β is an estimation of the specific weight of a word, capturing the impor-

6 For greater understanding of each automated text analysis method available, see Izumi and Moreira (2018).

7 For other techniques beyond wordscores and wordfish, see Lowe (2016).

tance of p in differentiating among party positions; ω is the estimation of the political position from document i. In the model, the fixed effect by words controls terms more used than others, as well as the fixed effect by document balances manifestos larger than others.

As with any model, *wordfish* has its limitations. The main one is unidimensionality. Nonetheless, its application is appropriate for the corpus of this study, given that it consists of documents on which candidates and parties position themselves on a variety of issues. Another problem is that the analysis produced is based, most of all, on the frequency of words. Therefore, it is not possible to assume results regarding the meaning of the texts. As Grimmer and Stewart (2013) highlighted, it is still out of the scope of automated analysis the control of irony and, in this work, of synonyms, for example. Therefore, if a candidate proposes the *construction of homes* and another the *construction of houses*, the last words will be singled out as divergences, even when they have same meaning[8]. In addition, it is essential to point out that this is a comparative model, meaning that a position is only given to a candidate from the frequency of the words they use in relation to their opponents.

In this paper I use *wordfish* to analyze 889 manifestos registered by candidates to executive office in Brazil between 2010 and 2018[9] in the three levels of representation. Although I have discussed that programmatic competition is not always ideologically organized, my hypothesis is that it is related to the level of the contest. Not because some are more visible, politically "correct", or institutionalized, but because the nature of the competition structures, in some way, the nature of the policies proposed. With this I mean that it is harder for a mayor to ideologically articulate in their manifesto while they must propose very specific actions for cities. They do not have, for example, the prerogative to implement large policies that are more or less liberal or conservative. These broader agendas, however, find more space at higher levels, especially in presidential contests. That is why I expect that the spatial position of manifestos by presidential candidates is distributed in a left to right scale, and that that dimension gives way to others, as it reaches states and municipalities.

In practice, this exercise will be conducted from a comparison of the position obtained by parties considering the set of elections at each level with the classification by the literature. I am unaware of large divergencies in the classification of parties done thus far, whether through specialists or politicians' surveys or measured through parliamentary behavior. In this paper, I use as reference the classification elaborated by Codato, Berlatto e Bolognesi (2018), for the practical reason of encompassing a larger number of parties, including newer ones. I consider cases convergent with the authors' classification as ideologically aligned. For contests in which the position of the parties does not coincide with the literature's classification, I conclude that other dimensions organize the programmatic competition.

It is important to highlight that, before being used in *wordfish*, the manifestos were pre- processed. The encoding and format of the files was standardized; the order of the words was disregarded (bag of words); were removed from the texts: diacritics,

8 Despite this and the controls that *wordfish* itself has, it is necessary to consider that words, through semantics and etymology, do not have arbitrary meanings. That is, even if homes and houses may have been used with the same meaning, this does not extend randomly to the linguistic universe, in such a way that these words cannot mean trees or cars.

9 Law 12.034/2009 requires that all candidates to executive office register manifestos with the Superior Electoral Court (TSE), when requesting the registration of the candidacy. The law does not discuss the size or type of document to be presented. All manifestos are available on the website <http://divulgacandcontas.tse.jus.br/divulga/#/>. Last accessed on 28/06/2020.

numbers, punctuation, words with unnecessary content, such as articles, prepositions, and conjunctions, and words that, on the whole, had lower frequency then twice the total number of documents on a model. To avoid similar words being counted separately, diminishing the frequency of a theme, *stemming* was used, a process that removes suffixes from words, reducing them to their *stem*[10]. These processes resulted in a *document-term matrix* (*dtm*) in which each document *i* represents a vector that counts the frequency in which each unique *stem* (S) occurs, $S_i = S_{i1}, S_{i2}, \ldots, S_{in}$. Next, that matrix was applied to *wordfish*. Over the next section, I present and analyze the results obtained.

5. Programmatic competition in Brazil (2010- 2018): an ideological contest?

In this paper, I work with three different *corpora*, organized by the level of the contest the manifestos are referring to, totaling 889 documents[11]. The first *corpus* refers to the platforms of three presidential elections – 2010, 2014, and 2018 – and includes 33 manifestos. The second concerns gubernatorial races in the same years, totaling 468 manifestos. Lastly, the third *corpus* encompasses 388 documents registered by mayoral candidates in Brazilian capitals in 2012 and 2016[12] (Table 1).

10 As an example, words such as "trabalho", "trabalhadora" and "trabalhadores" (meaning "work", "woman worker", and "workers") are all considered "trabalh". Regarding this process, it is also important to consider that it does not always coincide with the morphological root of the words.

11 I disregarded changes in party names over the period, to make the comparison and visualization of the graphs easier.

12 Because the law that requires and makes manifestos available is from 2009, only elections after this year were considered.

Table 1 – Distribution of the corpora's documents

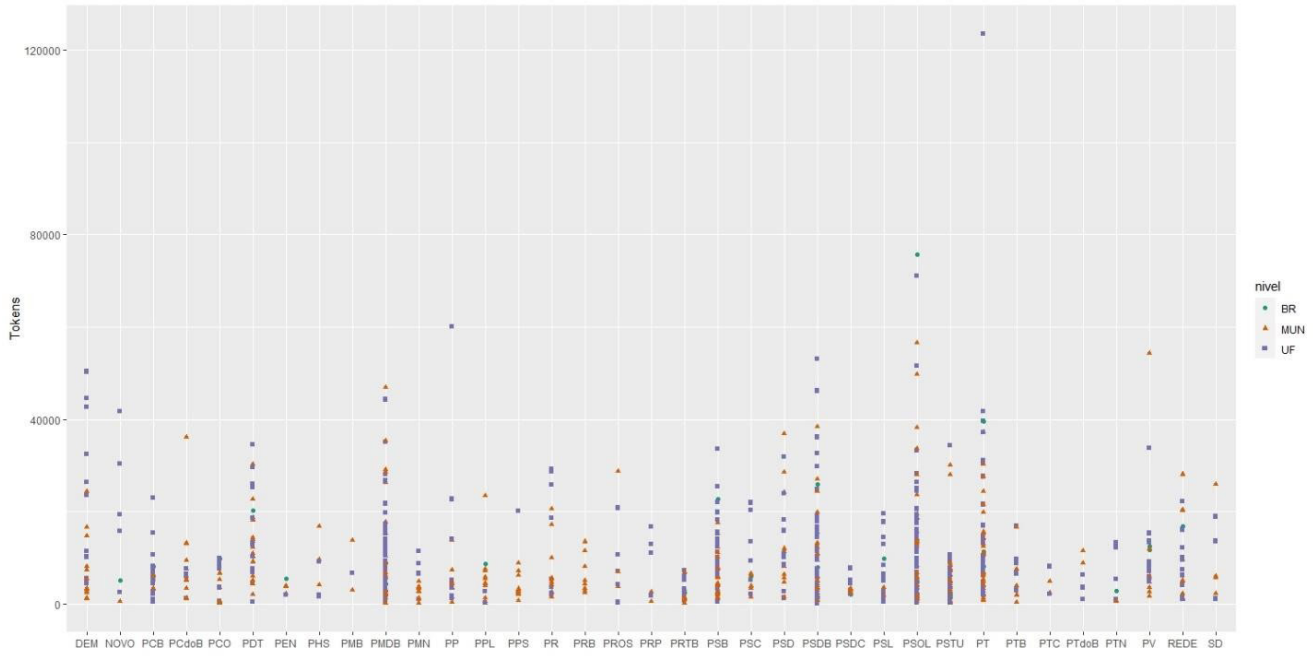
	Brazil	State	Municipality	Total
DEM	-	12	12	24
NOVO	1	5	1	7
PCB	2	21	5	28
PCdoB	-	4	11	15
PCO	2	16	7	25
PDT	1	13	17	31
PEN	1	1	3	5
PHS	-	3	3	6
PMB	-	1	2	3
PMDB	1	44	27	72
PMN	-	4	9	13
PP	-	9	10	19
PPL	1	3	13	17
PPS	-	2	10	12
PR	-	8	10	18
PRB	-	-	9	9
PROS	-	5	3	8
PRP	-	5	3	8
PRTB	2	14	12	28
PSB	1	28	22	51
PSC	1	6	5	12
PSD	-	10	10	20
PSDB	3	39	30	72
PSDC	3	5	4	12
PSL	1	16	2	19
PSOL	3	67	46	116
PSTU	3	41	32	76
PT	3	42	35	80
PTB	-	7	5	12
PTC	-	2	2	4
PTdoB	-	3	2	5
PTN	1	4	2	7
PV	2	14	10	26
REDE	1	11	10	22
SD	-	3	4	7
Total	33	468	388	889

Source: the author.

As mentioned, the law that requires the registration of manifestos does not specify any attributes regarding their content. When observing the length of these documents, it is also not possible to link the size of these platforms with the parties, since there is variation in the number of tokens[13] included in them, both between parties and levels of competition (Graph 2). The longest manifesto registered was by the PT in Acre, in the 2014 elections, with 123.571 tokens, while the shortest was elaborated by the PSDB in the Federal District in the same election, with only 67 tokens.

13 Text unit analyzed. For this paper, the number of tokens equals the amount of stems.

Graph 2 – Manifesto distribution according to the number of tokens



Source: the author.

Taking into account not only the length of the documents, but also the more frequent terms, it is also possible to observe some variations regarding spheres an ideology (Table 2). For example, *god* appears only among the most recurrent ones in right-wing platforms, as well as *secur*. In center parties, this happens with the terms *invest* and *reinforc* and, in left parties, with *popul* and *cultur*. *Right* appears in center and left-wing manifestos, but less frequently in the latter. With the exception of these tokens, there are only differences in how often they are mentioned: *work* takes the number one spot among left-wing parties, while it is 8th among right-wing and 11th among center parties. The same happens with *social*, which has a frequency rate from highest to lowest, respectively, in the manifestos of the left, right, and center. Lastly, *develop* is mentioned more often by center candidates and has an equal frequency among right and left-wing parties. These differences, however, do not seem to determine the content predicted belonging to each of these categories.[14]

In the gubernational elections *corpus*, new terms appear in comparison to the most frequent ones in presidential manifestos, such as *school*, *support*, *resource*, *attend*, and *educat* (Table 3). Most of these words appear in the manifestoes of center parties, in such a way that the similarity among the most frequent terms in the two levels of competition is greater among left and right parties. Thus, *security* remains as a specificity of the right wing's ranking, as well as *cultur* remains in the left's. *School* is mentioned more often by the right and does not appear among the most frequent in the left's manifestoes, which do talk about *educat*. Unlike the presidential manifestoes, *popul* comes up, in the states, for parties in the whole ideological spectrum, although along with *work*, is more frequent in left wing parties.

14 Translation note: these are the terms translated and stemmed; the table below shows each term in their original stemmed form in Portuguese (ones mentioned in bold).

Table 2 – Most frequent terms in the presidential elections corpus

	Right	Center	Left	Total
1º	tod	tod	trabalh	**trabalh**
2º	acima	dev	govern	tod
3º	govern	desenvol	**popul**	govern
4º	estad	ser	tod	nacional
5º	nacional	nacional	nacional	desenvolv
6º	desenvolv	govern	desenvolv	**social**
7º	dev	estad	social	popul
8º	trabalh	form	direit	dev
9º	social	garant	garant	direit
10º	ser	pod	estad	estad
11º	melhor	trabalh	dev	garant
12º	todos	nov	nov	ser
13º	deus	apoi	ser	nov
14º	setor	social	sistem	form
15º	federal	maior	form	pod
16º	tudo	invest	pod	sistem
17º	bolsonaro2018	qualidade	part	maior
18º	nov	sustent	grand	part
19º	pod	direit	cultur	**invest**
20º	seguranc	sistem	maior	setor

Source: the author.

Tabela 3 – Most frequent terms in the state elections *corpus*

	Right	Center	Left	Total
1º	estad	estad	estad	estad
2º	govern	govern	govern	govern
3º	desenvolv	desenvolv	**trabalh**	desenvolv
4º	tod	tod	**popul**	tod
5º	melhor	nov	tod	trabalh
6º	social	melhor	desenvolv	social
7º	nov	**social**	**social**	popul
8º	escol	estadual	garant	Bnov
9º	popul	sistem	**direit**	estadual
10º	ser	implant	estadual	ser
11º	sistem	invest	form	melhor
12º	trabalh	ser	ser	sistem
13º	estadual	popul	**cultur**	garant
14º	projet	projet	part	form
15º	outr	atend	**educ**	outro
16º	seguran	**recurs**	grand	escol
17º	recurs	**apoi**	nov	educ
18º	form	trabalh	pod	projet
19º	implant	produt	sistem	pod
20º	produt	escol	outr	recurs

Source: the author.

In municipal elections, cultur starts to show in the rankings of parties in the whole ideological spectrum, but more frequently in left wing parties (Table 4). In these, the particularity becomes the term *participa* and the permanence of *right*, now absent among the most frequent for right and centre. *School* enters the left's ranking. In addition to these words, *urban* is included among the most frequent for parties in the whole spectrum highlighting that, in addition to the number one spot, the debate is now about the cities.

Table 4 – Most frequent terms in the municipal elections *corpus*

	Right	Center	Left	Total
1º	cidad	cidad	cidad	cidad
2º	municipal	municipal	**popul**	municipal
3º	desenvolv	desenvolv	municipal	popul
4º	tod	tod	**trabalh**	tod
5º	popul	popul	**social**	social
6º	social	nov	tod	desenvol
7º	escol	melhor	govern	trabalh
8º	atend	**social**	desenvolv	govern
9º	melhor	implant	**cultur**	escol
10º	implant	govern	garant	cultur
11º	projet	**atend**	**educ**	atend
12º	nov	**escol**	**escol**	garant
13º	cri	urban	**direit**	nov
14º	urban	plan	ser	urban
15º	intergr	**cultur**	**particip**	educ
16º	ser	projet	**urban**	ser
17º	govern	ser	dev	melhor
18º	cultur	sistem	form	implant
19º	red	**trabalh**	outr	projet
20º	promov	**educ**	pod	cri

Source: the author.

Despite some variations, the description of the most frequent terms is not very useful to think about the parties' ideological classification, thus the importance of analysing them through a scaling model, *wordfish*. The biggest difference is that now all the terms in a manifesto are considered and its words are weighted given their frequency, as presented. The starting point is the presidential election, given the documents used by the literature already presented and, also, it being the level in which large packages of policies are likely to be proposed and implemented.

As mentioned, *wordfish* attributes positions from the frequency and not the meaning of the words. For some this can go against the assumption of this paper of not identifying a programmatic dispute in situations where there is no divergence among the terms mobilized. After all, candidates could use similar terms, not to say identical, with different meanings. That would be the case, for example, of someone who is dedicated to proposing policies on volunteer work and the creation of job openings in businesses; and another who preferred topics on work and public service. In this situation, *work* would be counted in both documents. Although that limitation is real, I do not believe is insurmountable, since although *wordfish* does not capture the meaning of the words, the inclusion of all the terms offers some degree of political differentiation. In the example mentioned, tokens related to the private businesses and volunteer work would measure the position of the first party as well as public service would be considered for the second.

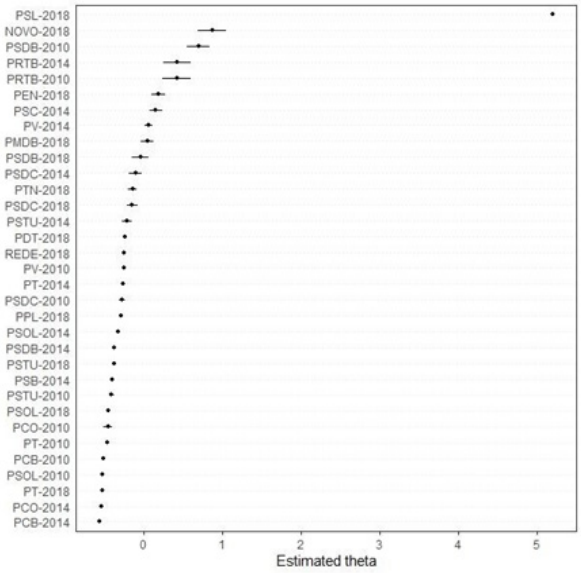
Having observed Graph 3, it is possible to see that, unlike the results obtained through the *MARPOR* codification, *wordfish* scales the parties more or less close to the (ideological?) center, although the majority of them are on the left. With this classification, the party closest to the right is the PSL, in 2018, and the PCB, in 2014, the closest to the left. Overall, with the exception of the PSTU's manifesto from 2014, which appears close to the center, it can be stated that the positioning of left-wing parties is consistent with the literature's classification, given that the PT, PSOL, PCO, and PCB are close and in one of the

scale's extremities. This does not mean that there are no inconsistencies, such as the PSDB's manifesto from 2014 being in between the PSOL's (2014) and the PSTU's (2014); or the PSDC of 2010 finding itself closer to the PT (2014) and the PPL (2018). Nevertheless, the programmatic spectrum of presidential elections seems to be strongly organized by ideology.

1.1 Programmatic competition in the subnational (2010-2018)

This landscape, however, is altered when elections for the state and municipal executive are considered. In some states, the configuration is similar to the national context. In others, it is quite different, such as Acre (Graph 4). Looking into the race for governor, it is not even possible to argue that there is an actual programmatic competition, given that the parties' positioning is almost overlayed. That is, they used practically the same terms in their manifestos. There are differences such as the PSDB and the PT in 2014, but the most significant one is, undoubtedly, the most extreme positioning of the PSL's manifesto in 2018[15]. The manifestos that present the largest divergence were in the race for the city hall of Rio Branco but, even then, the PSOL (2012) is located in the center, along with the majority of the parties. The extremities are led by the PT in the left and the PSC in the right. Therefore, the programmatic competition in the state seems to organize itself ideologically more at the municipal than at the state level. After all, in this arena, the dispute is not very programmatic, as in it is not divergent.

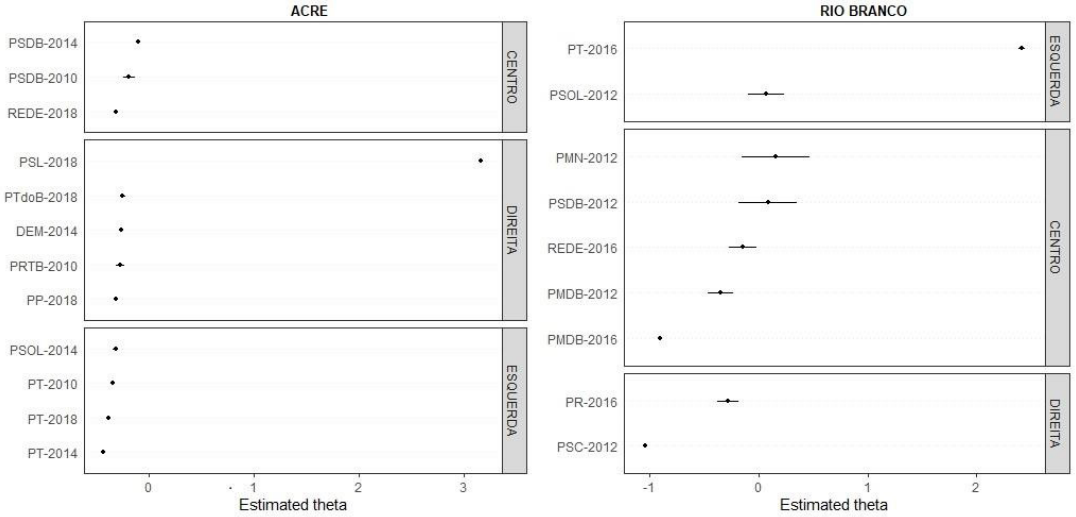
Graph 3 – Presidential programmatic competition (2010-2018)



Source: the author.

15 It is important to highlight that the meaning (positive or negative) of the position is arbitrarily given by the model. Thus, being in the negative extreme does not necessarily mean being in the left, this is something to be interpreted by the researcher.

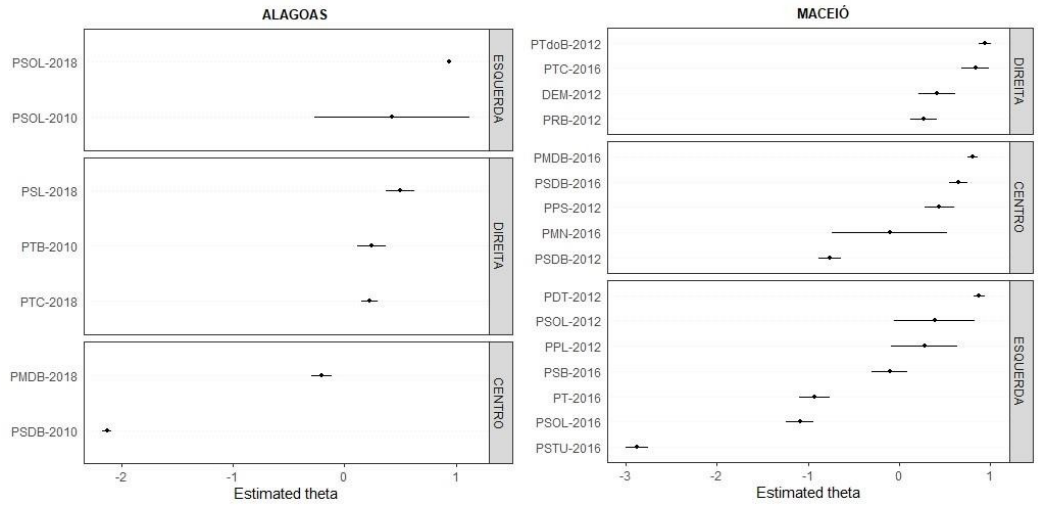
Graph 4 – Programmatic competition in Acre (2010-2018)



Source: the author.

In Alagoas, there is greater divergence and, therefore, greater competition between the manifestos of candidates to the state's executive than in Acre (Graph 5). The PSOL occupies the most leftward position and the PSDB, the most rightward. The PSL, PTB, and PTC presented similar proposals. However, the fact that the PSL, in 2018, registered a manifesto located close to the PSOL's from 2010 indicates limitations of the interpretation of this scale as a precisely ideological scale. These issues are reproduced in the municipal races of Alagoas's capital. On the one hand, PTdoB, PTC, PMDB, and the PSDB are located more to the right, while PSTU, PSOL, and the PT, are concentrated more towards the left. Similar to the state election, this context is somewhat muddy, as similar positions have been attributed to the PPL, the PSOL, the PPS, and the PRB. Overall, the programmatic competition seems more ideologically organised in Alagoas than in Acre. Nonetheless, inconsistencies may indicate that there are other programmatic dimensions relevant in these disputes.

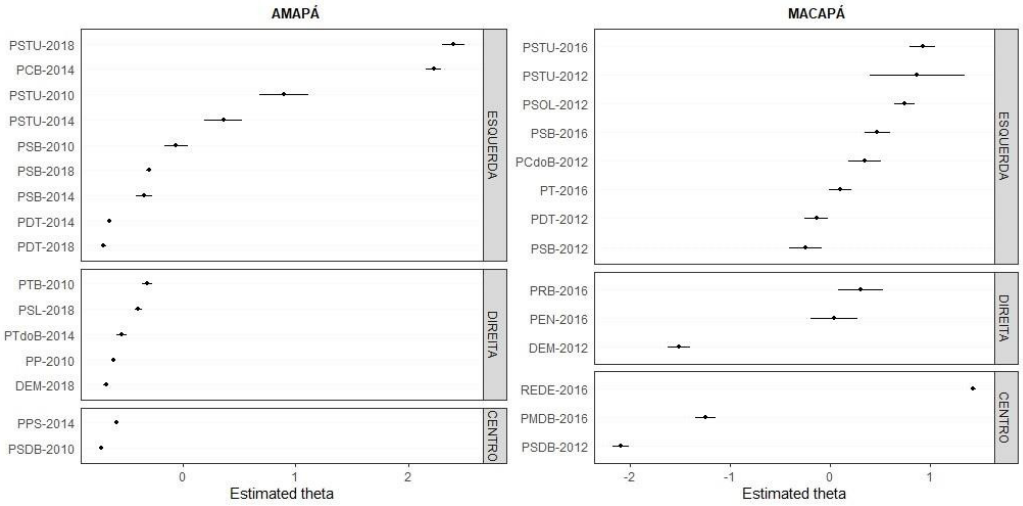
Graph 5 – Programmatic competition in Alagoas (2010-2018)



Source: the author.

In Amapá, ideological differences seem to be more defined in the left, since the PSTU, the PCB, and the PSB (2010) have well-defined positions that are close to each other (Graph 6). Unlike Acre, here the race is structured in the more polarized way and, more to the right, the PSDB, the PDT, and DEM are found. There are inconsistencies, especially if we consider the proximity of the PSL with the PSB, in 2018 and in 2014. The most evident delimitation of the left-wing field is also reproduced in the race for Macapá's City Hall, led by the PT, the PCdoB, the PSB (2016), the PSOL, and the PSTU. The most leftward position, however, is by REDE in 2016. To the right are, respectively, the PSDB, the DEM, and the PMDB, while PEN, PRB, PSB (2012), and PDT occupy more central positions.

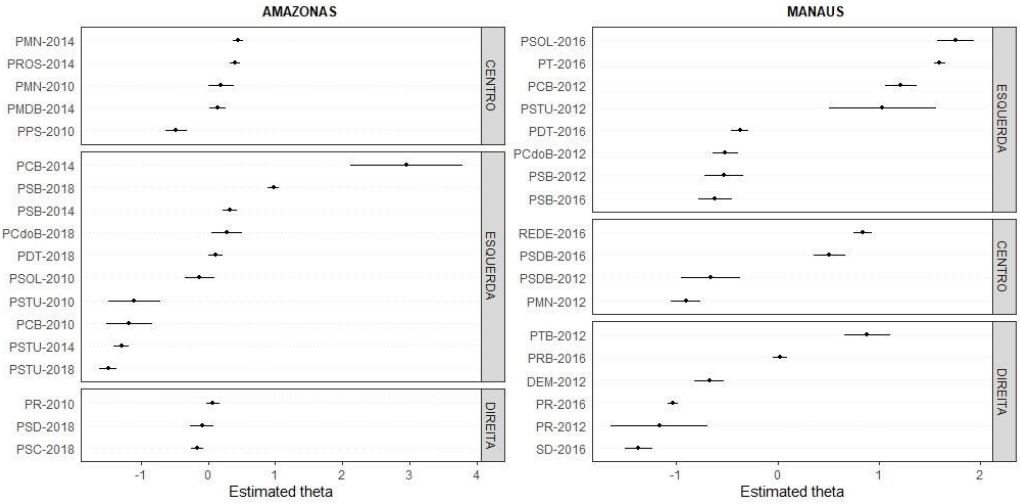
Graph 6 – Programmatic competition in Amapá (2010-2018)



Source: the author.

Differently from the configurations seen in other states, in Amazonas, the programmatic competition seems to structure itself in a more divergent manner when we compare both levels of dispute (Graph 7). In elections for the state's executive, the majority of parties is located in the center of the programmatic spectrum; in the competition for the municipal executive, the scale is more polarized. Therefore, in the state race, the PSC, PSD, PR, PSOL, PDT, PSB (2014), PMDB, PROS, and PMN occupy central positions, while the PSTU and the PCB are to the left and the PSB (2018) and the PCB (2014) to the right. This configuration seems to indicate that such a programmatic competition could be best understood by other dimensions, added or not to the ideological one. This, in some way, is extended to municipal elections, given that although the positions are less central, there is polarisation across all ideological fields. Thus, the left appears divided among: the PSOL, PT, PCB, and PSTU on one hand, and the PDT, PCdoB, and PSB, on the other. The center, between REDE and the PSDB (2016) and the PSDB (2012) and the PMN; and the right between the PTB, on the one hand, and the SD, the PR, and DEM, on the other, with the PRB occupying the central position.

Graph 7 – Programmatic competition in Amazonas (2010-2018)

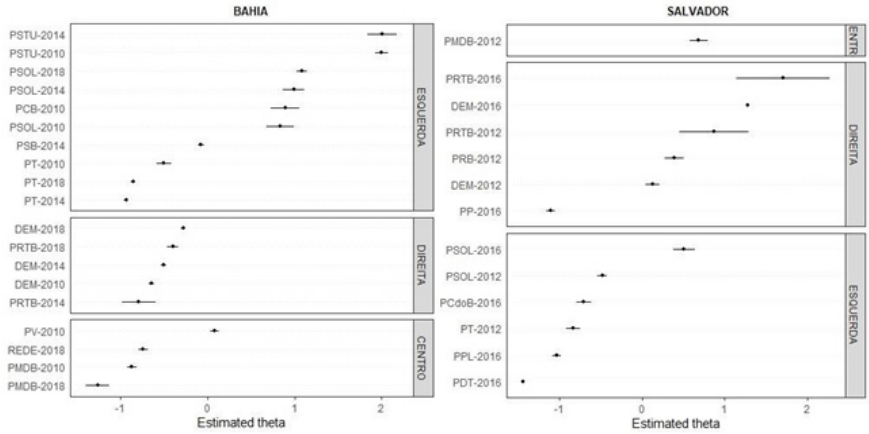


Source: the author.

In Bahia, the position of left-wing parties is well-demarcated both in state and municipal races (Graph 8). In the former, the PT is an exception since it is positioned closer to center and right-wing parties – which happens with the PSOL in mayoral elections. At the state level, center parties are located more towards extreme positions than right-wing parties but, overall, have similar positions. At the local level, the central space is not occupied and, excluding the PP and the PSOL in 2016, the programmatic competition seems aligned, given that the PST, PPL, PT, PCdoB, and the PSOL (2012) are found at the left and the DEM, PRB, PRTB, and PMDB, towards the right.

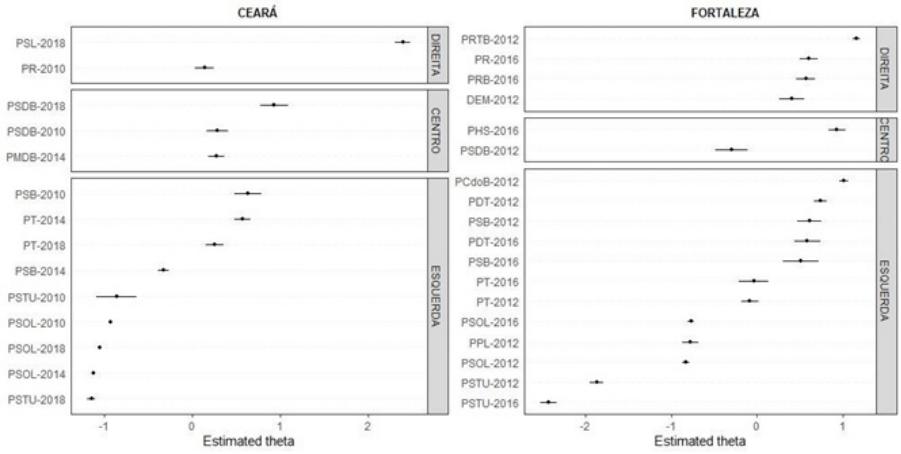
Manifestos registered in the 2018 elections occupy the most extreme positions in Ceará's state-level the programmatic spectrum, with the PSL in one of the poles and the PSTU in the other (Graph 9). The PMDB and the PSDB (2010) are located in the center, with the PSDB moving more towards the right in 2018. There are, however, inconsistencies in the left-wing since the PSTU and the PSOL are positioned closer to -1 and the PT and the PSB are closer to 1. Thus, the left would be polarized over a programmatic space that seems to include other dimensions beyond – or instead – of ideological. This scenario is radicalised when the mayoral elections for Fortaleza are considered, with a larger dispersion of the left throughout the whole programmatic spectrum. If it were an ideological scale, the PT would occupy the central position and only the PSOL, PPL, and PSTU would be located in the left. The PCdoB, for example, would be more towards the right then the DEM, PRB, and PR. therefore, other programmatic dimensions sing to articulate these races.

Graph 8 – Programmatic competition in Bahia (2010-2018)



Source: the author.

Graph 9 – Programmatic competition in Ceará (2010-2018)

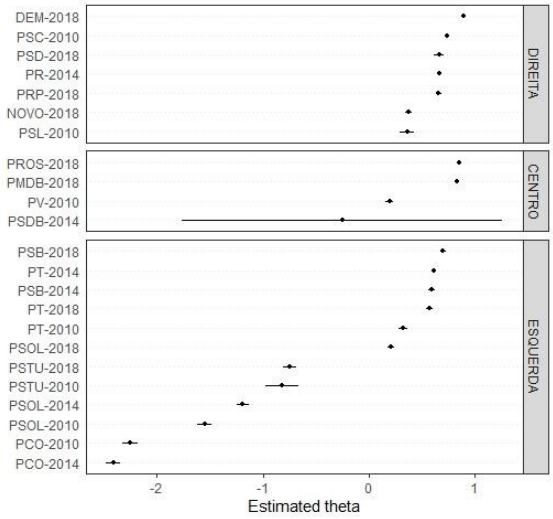


Source: the author.

The programmatic competition for the Federal District's executive also does not seem to be organised ideologically given that there are no demarcated positions for parties considered to be left-wing, center, and right-wing (Graph 10). In other words, there are similar positions for different parties, regardless of their ideological classifications, the more extreme and demarcated location being occupied by the small left-wing parties, the PSTU, PSOL, and PCO. The reiteration of these positions over states and municipalities seems to suggest that these parties may exert a centripetal influence over remaining left-wing parties, given the comparative nature of the model.

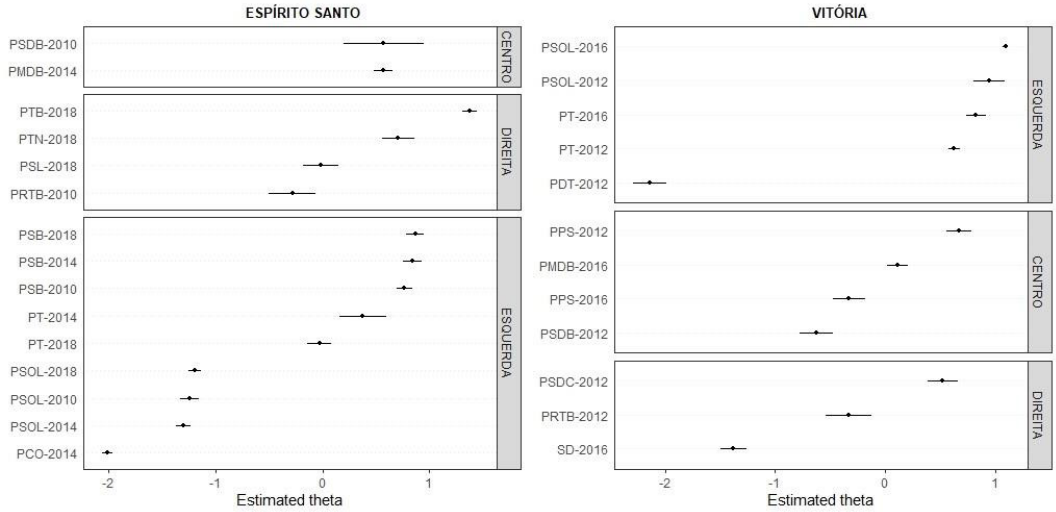
In Espírito Santo, the more extreme positions are consistent with the parties' ideological classification (Graph 11). At the state level, the PTB is on one side and the PCO is on the other. This does not happen, however, at the municipal level where left-wing parties occupy both poles, the PSOL in the positive and the PDT in the negative. Despite this, the location of left-wing parties is more concentrated when manifestos registered by candidates to the Vitória's mayoral office is considered than the state level manifestos. In this sphere, there is no large divergence between the PSDB, PMDB, and the PSB for example, and the PSL and the PRTB are closer to the PT than to the PTN. These data seem to indicate that, also in this state, there are other programmatic dimensions organizing the competition.

Graph 10 – Programmatic competition in the Federal District (2010-2018)



Source: the author.

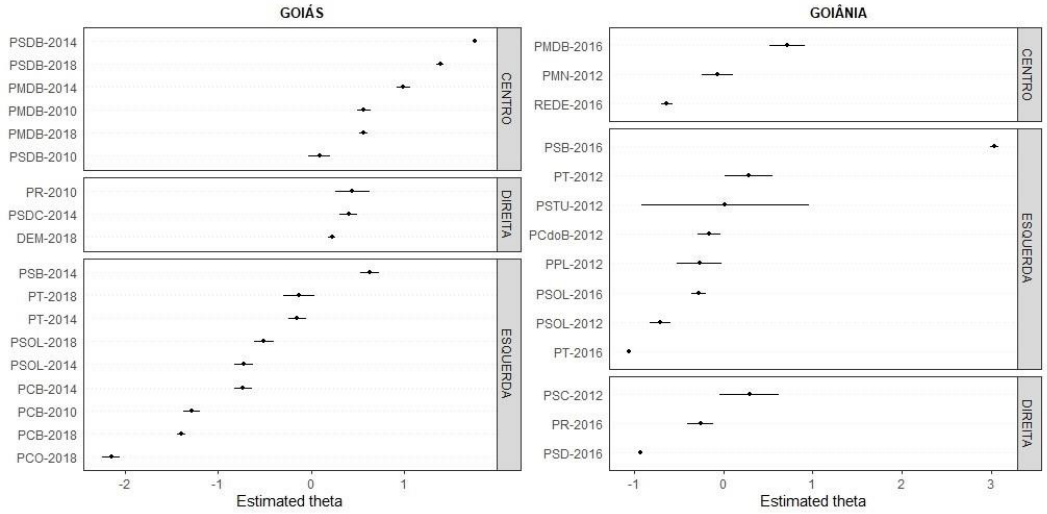
Graph 11 – Programmatic competition in Espírito Santo (2010-2018)



Source: the author.

Goiás is the first case where the ideological alignment of programmatic competition can be claimed, at least at the state level (Graph 12). Although parties considered to be in the center occupy more extreme positions then the ones in the right, there is almost no overlay of parties, in such a way that the points in the graph are shaped like an inclined line. In this classification, the PT, PSOL, PCO, and PCB are situated in the left, the PSDB (2010), PMDB (2010, 2018), PR, PSDC, and PSB are closer to the center, and the PMDB (2014) and PSDB (2014, 2018) are in the right. In the Goiânia mayoral races, there is a dispersion of the parties in all ideological groups in the programmatic spectrum, in a way that hinders the identification of left and right-wings. At this level, ideology is insufficient to explain the parties' positions in the competition.

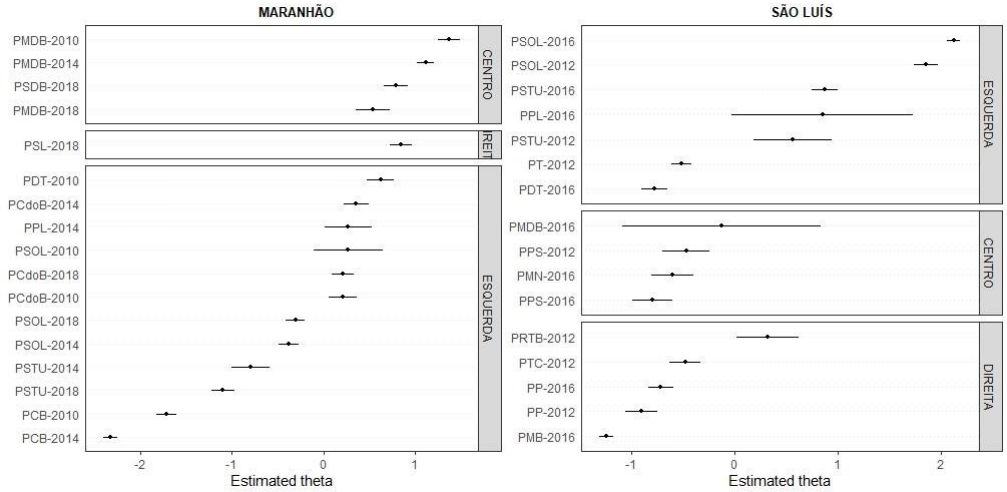
Graph 12 – Programmatic competition in Goiás (2010-2018)



Source: the author.

At first glance, it may seem that Maranhão is similar to the state scenario of Goiás, since the positions are also distributed almost in the form of an inclined line (Graph 13). If we observe the medium point, however, we can observe that it does not divide the parties between left and right, separating only the most extreme parties, the PSOL, PSTU, and the PCB. The PSL also occupies a more central position than the PMDB in the total of races, but the biggest issue seems to be that the PCdoB, PPL, PSOL, and PDT also somehow share this space. The local level competition, however, seems to be less inconsistent when it comes to the relationship between ideology and the location of parties. Thus, although the PRTB and the PT are more towards the center, and the PDT more towards the right, there is a more concentrated positioning of parties of the same ideological classification.

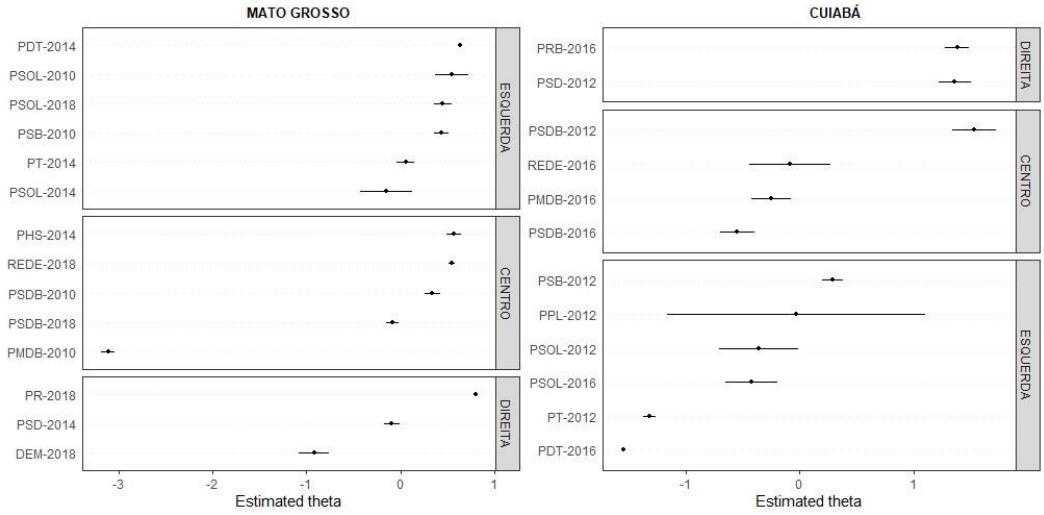
Graph 13 – Programmatic competition in Maranhão (2010-2018)



Source: the author.

The disposition of parties for Mato Grosso's state government race is similar to the one seen for Acre, given that there is significant overlay, indicating little divergence – and, therefore, programmatic competition (Graph 14). Candidates for mayor, on the other hand, differentiate their proposals more. Nevertheless, the competition is mainly around the center where the PSOL, PPL, PSB, PSDB (2016), PMDB, and the REDE are located. Towards the left are only the PT in the PDT, and towards the right, the PSDB, the PSD, and the PRB. In comparison to other capitals, the scenario in Cuiabá does not constitute complete ideological disorder but seems to be connected to other programmatic dimensions.

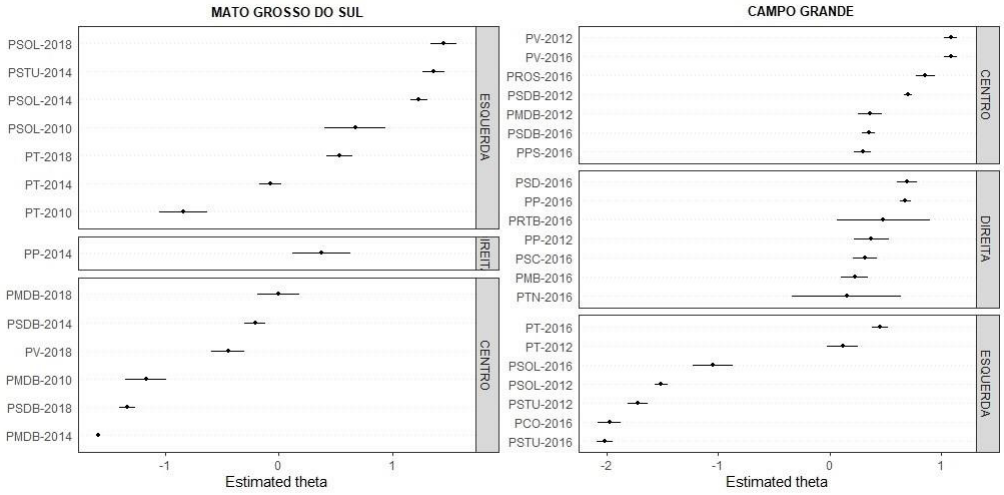
Graph 14 – Programmatic competition in Mato Grosso (2010-2018)



Source: the author.

In Mato Grosso do Sul, on the other hand, the race is more ideologically aligned, albeit with inconsistencies (Graph 15). At the state level, the main one is the fact that at least half of center parties occupy the right-wing field, while right-wing party PP appears in the center-left. In the left, the PT in 2014, is positioned in the center and in 2010, towards the right. In 2018, the party is closer to the others, the PSOL and the PSTU. The state scenario is reproduced in the races for the Campo Grande mayor's office. In the two years considered, the PT appears in the center (2012) and towards the right (2016), and center parties in a more extreme position then right-wing parties. Other parties seem ideologically organized, with the PSOL, PSTU, and PCO to the left, the PMDB (2012), PSDB, and PPS in the center, and the PTN, PMB, PSC, PP, PRTB, and the PSD in the right.

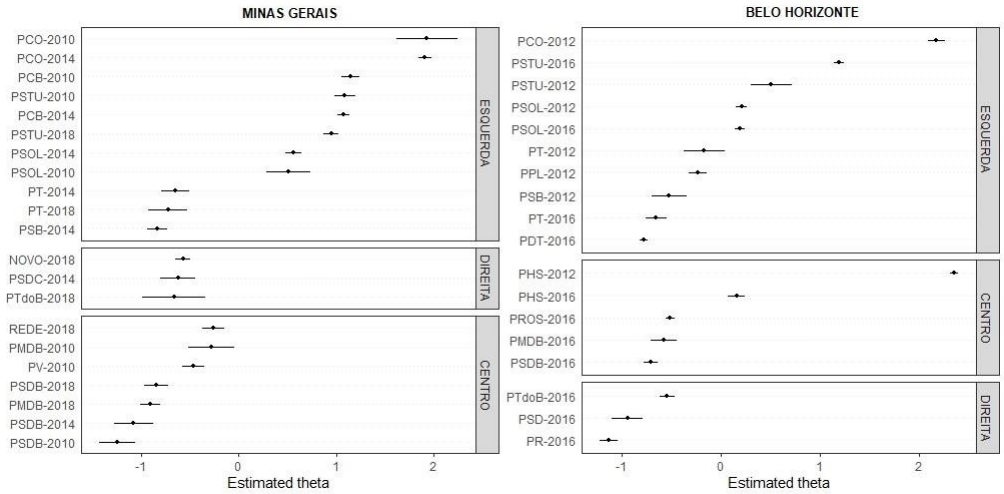
Graph 15 – Programmatic competition in Mato Grosso do Sul (2010-2018)



Source: the author.

This configuration is similar to the one found in the elections for the Minas Gerais's state executive, where the PSDB and the PMDB (2018, 2014) appear more towards the right than the NOVO, the PSDC and the PTdoB, as well as the PT and the PSB (Graph 16). Again, the case suggests that more extreme and less competitive left-wing parties exert a centripetal force over remaining ones in the same ideological field. In the races for the Belo Horizonte mayor's office this is heightened, given that it includes more parties: the PT, PPL, PSB, and the PDT appear in the center and right. Although at this level center parties do not present more extreme positions that right-wing parties, the PHS appears as towards the left as the PCO. These scenarios may be interpreted as a case where the ideological dimension is joined by others, given that the ideological dimension is not completely organized nor totally out of alignment.

Graph 16 – Programmatic competition in Minas Gerais (2010-2018)



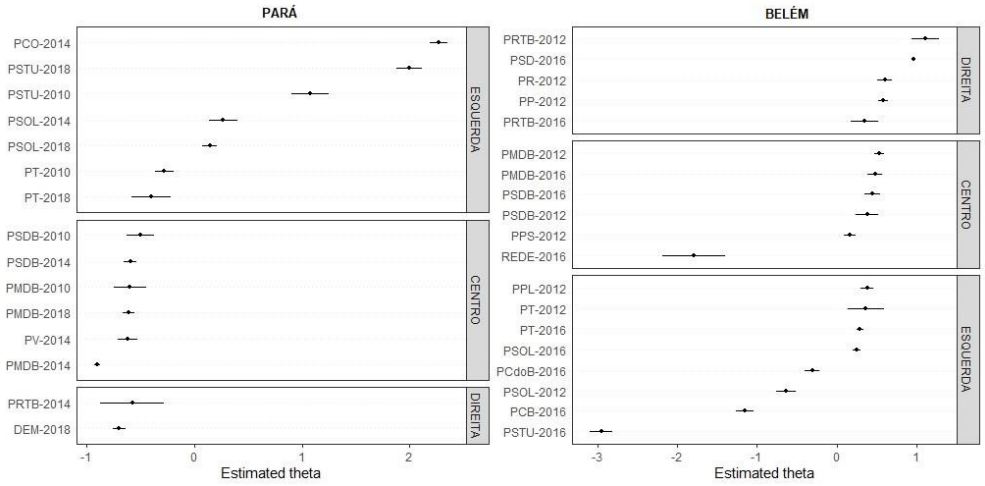
Source: the author.

In Pará, ideology does not seem to be the main dimension in organizing the programmatic competition neither at the state nor at the local levels (Graph 17). In the former, the PRTB, PV, PMDB (2010, 2018), and the PSDB hold almost the same position, which is also very close to the DEM and the PT in 2018. The PMDB in 2014 is the party that is positioned more extremely towards the right and the PCO, in the

same year, is in the most extreme end of the opposite pole. In the elections for the Belém mayor's office, the PRTB in 2012 and the PSTU in 2016 occupy these respective positions. In the municipal race, the right is more extreme than the center, however, even among part of the left, the spaces occupied are similar, reinforcing the little ideological alignment of the programmatic competition.

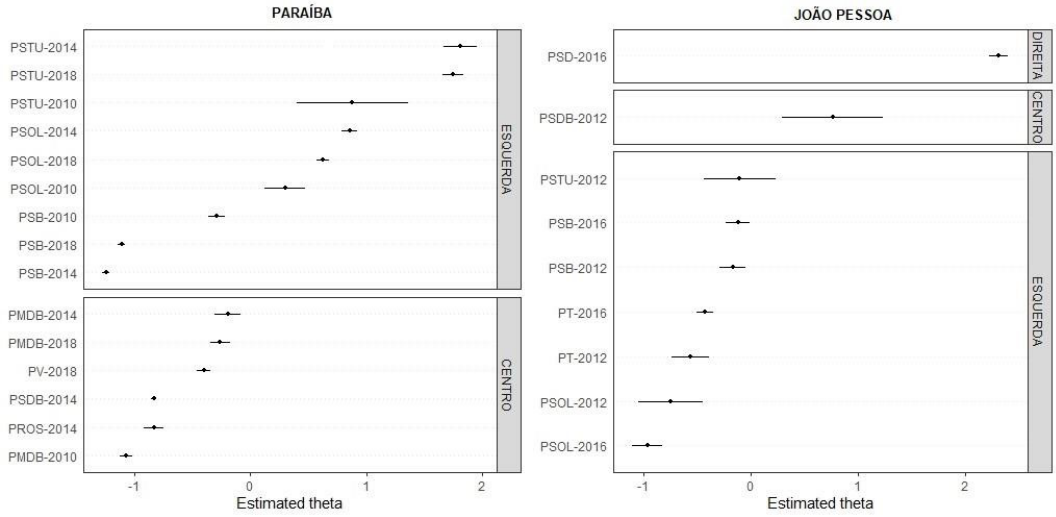
The programmatic competition in Paraíba appears to be partially organized by ideology, at least at the state level (Graph 18). This is due to the PSOL and the PSTU occupying the left and the PMDB (2018, 2014) and the PV, the center. On the other hand, the PSB is located in the center in 2010 and more towards the right in 2014 and 2018, where are also located the PSDB, PROS, and the PMDB (2010). In the races for the João Pessoa mayor's office, the programmatic competition is more consistently organized by ideology, with the PSOL, PT, PSTU, and the PSB towards the left, the PSDB in the center, and the PSD in the right.

Graph 17 – Programmatic competition in Pará (2010-2018)



Source: the author.

Graph 18 – Programmatic competition in Paraíba (2010-2018)



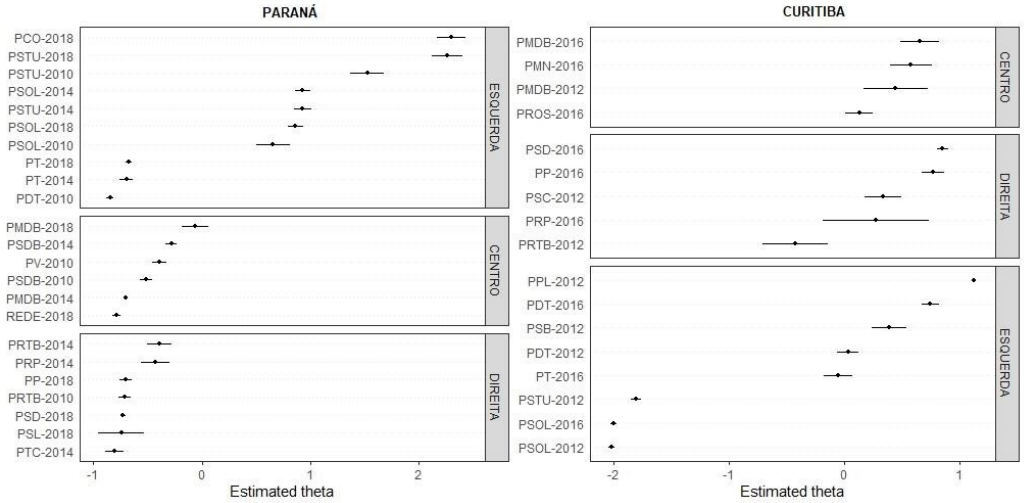
Source: the author.

The programmatic competition for the governor's office in Paraná consistently locates right-wing parties (Graph 19). Center parties appear in more extreme positions than central ones, with the exception of the PSDB (2014) and the PMDB (2018). The biggest issue seems to be the location of the PT and the PDT also

in this space, possibly influenced by the more acute position of the PCO, PSTU, and the PSOL, as seen in other states. When it comes to municipal elections, although there is clustering among the 0 and 1 values, they are more dispersed among themselves when compared to the state level. Nevertheless, that does not ensure ideological alignment, given that center parties are to the right of the PSD, PP, PSC, and the PRP and the extreme positions are occupied by the left: in one of the poles, by the PSOL and, in the other, by the PPL.

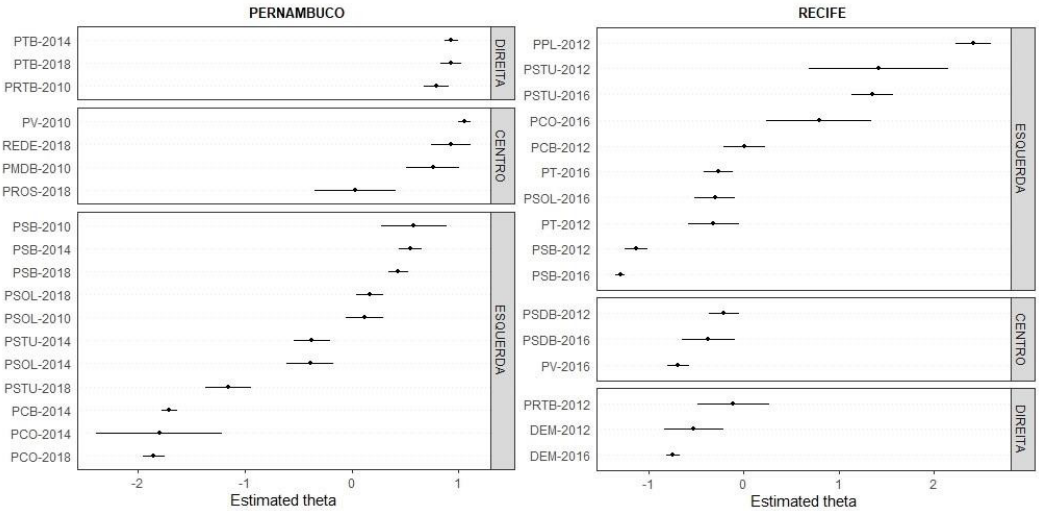
The dispersion of the left is also reproduced in the races at the state and municipal levels in Pernambuco (Graph 20). In both cases, some of the parties are positioned towards the left – PSTU, PSOL (2014), PCB, and PCO, at the state level, and the PCO, PSTU, and the PPL, at the municipal level – while the remaining ones are spread across the center and the right. The case of Pernambuco appears a bit more acute than the others, since even parties such as the PSOL (2010, 2018) at the state level, and the PCB, in Recife, are situated more centrally. In both cases, center and right parties are similarly positioned, between 0 and 1 in the programmatic spectrum. These configurations point more strongly towards the impact of other dimensions other than ideology in the organization of the programmatic space.

Graph 19 – Programmatic competition in Paraná (2010-2018)



Source: the author.

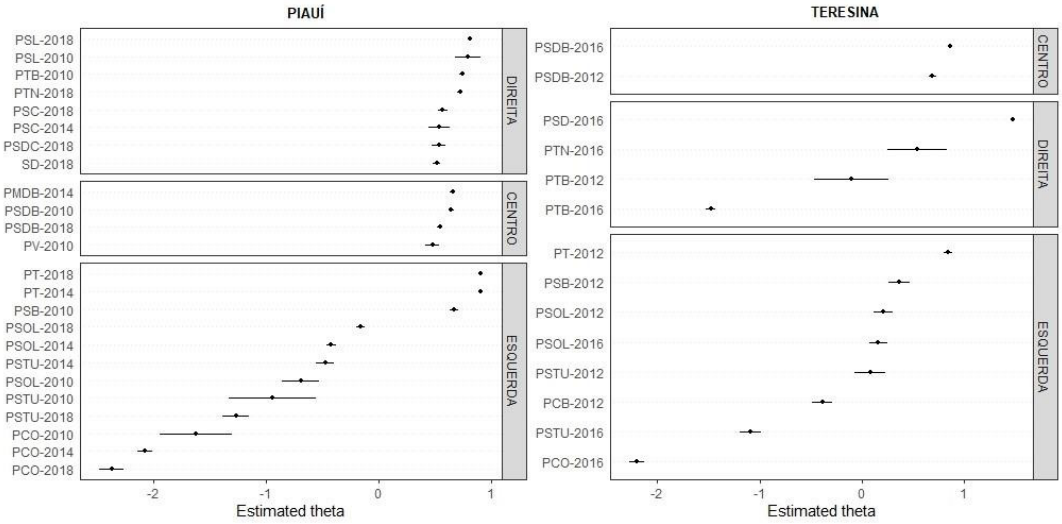
Graph 20 – Programmatic competition in Pernambuco (2010-2018)



Source: the author.

In Piauí, the scenarios laid out from the parties' position in the programmatic spectrum is different (Graph 21). While in state elections there is a concentration of right-wing parties in an opposing pole to where left-wing parties are, the positions of parties in both fields are more dispersed in the Teresina mayoral races. Nonetheless, there is a repetition in both cases: center parties appear to the right without occupying, however, the most extreme position. The right-wing pole is led by the PSD and the left, by the PCO. Despite this, the PSTU, PSOL, and the PTB are situated in central spaces, indicating weak ideological organization of the municipal programmatic competition. With the exception of the PSB's and the PT's positions, ideology seems to have a greater capacity for alignment in gubernatorial elections.

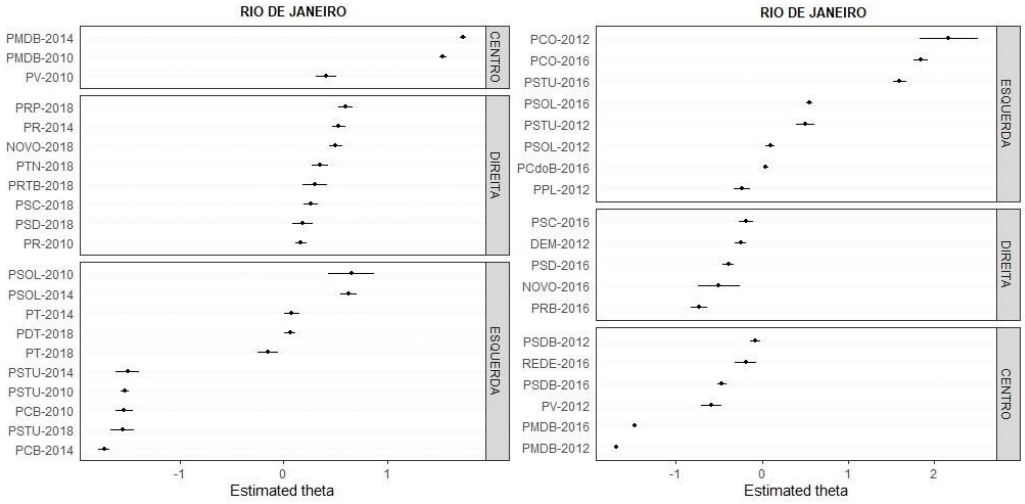
Graph 21 – Programmatic competition in Piauí (2010-2018)



Source: the author.

The configuration of races in Rio de Janeiro is striking due to the extreme position in the right taken by the PMDB, both at the state and local levels (Graph 22). It is notable not only due to the party being classified as center but mainly, because more acute positions are usually attributed to less competitive parties. This is not the only particularity found. In gubernatorial elections, the PSOL appears in the right and the PT and the PDT in the center. In races for the mayor's office of Rio, left-wing parties do not disperse towards the right, but the PPL is centrally located, very close to the PSC, DEM, PSDB, and the REDE. These examples show that the ideological dimension is insufficient to explain the programmatic space of the race, especially at the state level.

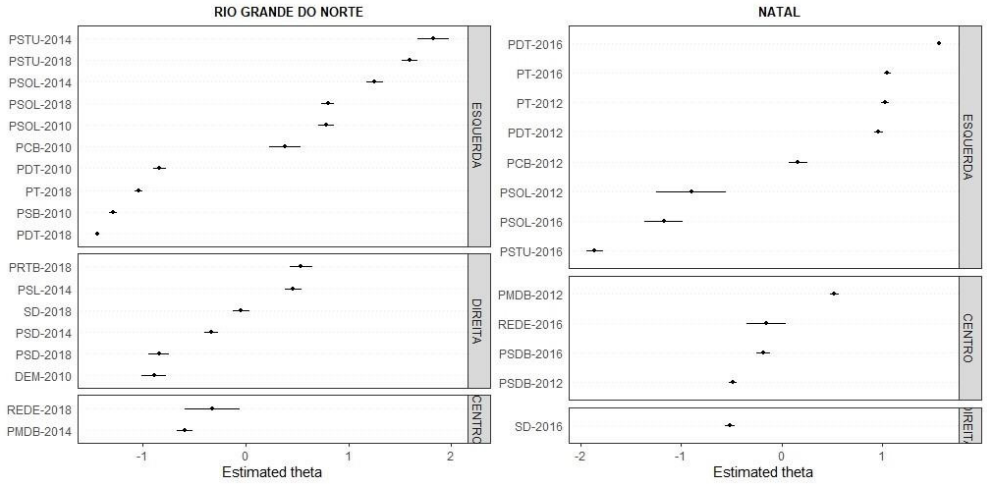
Graph 22 – Programmatic competition in Rio de Janeiro (2010-2018)[16]



Source: the author.

This result seems to extend to Rio Grande do Norte, with the difference that the configuration between the two levels is more aligned than the one observed in Rio de Janeiro (Graph 23). Here, left-wing parties occupy the most extreme positions on both poles. At the state level, the PSTU, the PSOL, and the PCB are placed in polarization against the PDT, the PT, and the PSB. At the municipal level, the opposition occurs between the PSOL and the PSTU on the one hand, and the PDT and the PT on the other, with the PCB occupying the central position. If it is not the first time that this polarized configuration is structured in the left-wing field, it is in Rio Grande do Norte that it occurs simultaneously with opposition among the right-wing parties at the state level, and center parties at the municipal level. These arrangements make it easy to discard the hypothesis the programmatic competition in these places is structured ideologically.

Graph 23 – Programmatic competition in Rio Grande do Norte (2010-2018)

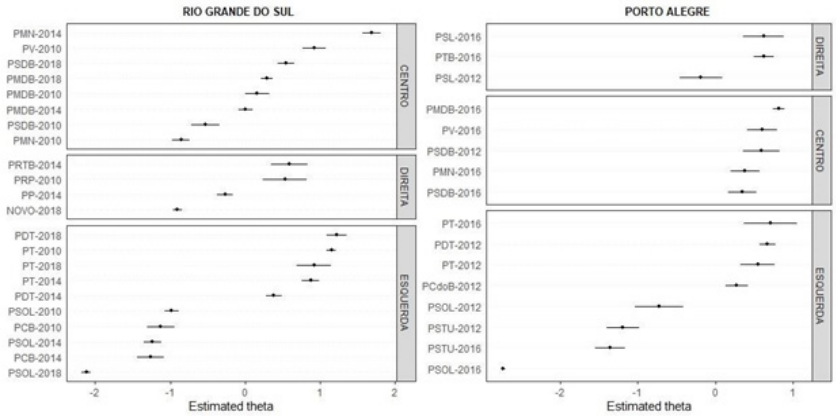


Source: the author.

16 Following the graph's pattern, the first image refers to the state and the second one to the city.

The polarization scenario among parties in the same field is reproduced in the gubernatorial elections in Rio Grande do Sul, in which both center, left, and right-wing parties appear in opposition with each other (Graph 24). This also happens in the elections for the mayor's office of Porto Alegre, but only among left wing parties. In that case, the PSOL and the PSTU you are positioned towards the left, while the PCdoB, the PDT, and the PT appear on the right with positions similar to the PSL and the PTB. Center parties are also located more towards the right than closer to the medium point of the programmatic spectrum. Thus, ideology is not exclusively able to explain neither configuration.

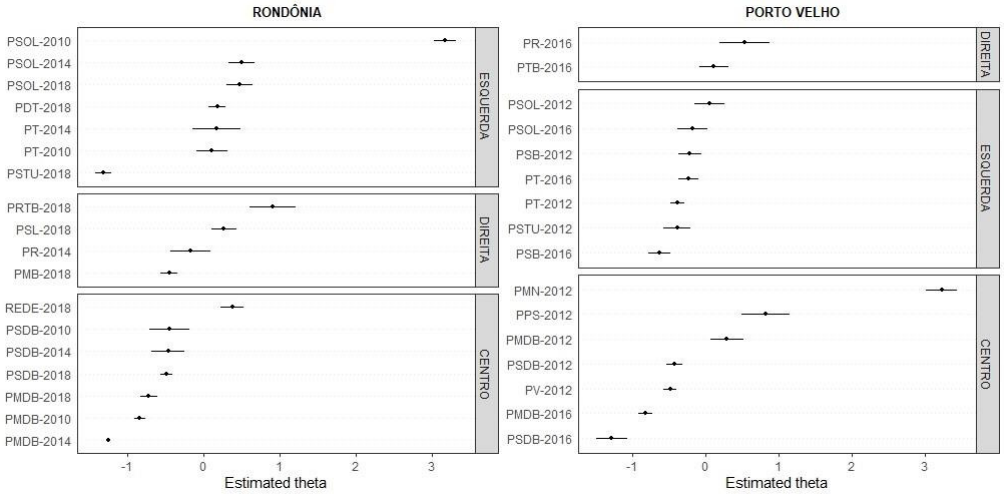
Graph 24 – Programmatic competition in Rio Grande do Sul (2010-2018)



Source: the author.

In Rondônia, the scenario is not so different. At the two levels considered, the extreme positions of both poles are occupied by parties of a same ideological field (Graph 25). At the state level, this occurs among left wing parties, specifically the PSOL and the PSTU. However, at the municipal level, the polarization happens among center parties, with the PMN on one side and the PSDB (2016) on the other. In state elections, right-wing and center parties also disperse among the two poles, as seen in Rio Grande do Sul. In elections for the Porto Velho mayor's office, the left is laid out in a line with a small incline, indicating concentrations of the parties in a space closer to the center. In the right, the PTB is centrally located, while the PR is positioned more towards the extreme. Again, these do not seem to be programmatic competitions organized ideologically.

Graph 25 – Programmatic competition in Rondônia (2010-2018)



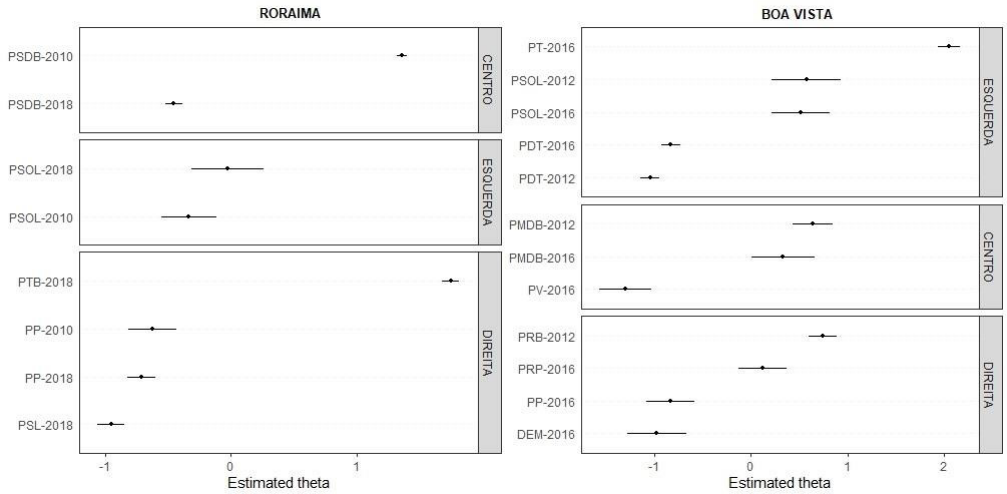
Source: the author.

This result is repeated in the elections in Roraima, where the PSOL occupies the more central positions at the state level and there is dispersion among the parties in the same ideological field at the municipal level (Graph 26). In gubernatorial elections, the PSDB polarizes between manifestos registered in 2010 and 2018, in addition to an opposition in the right between the PTB on the one hand, and the PP and the PSL on the other. Indeed, it is important to note that the PTB and the PSL specifically have the most extreme positions in the whole programmatic spectrum – which cannot be considered ideological – at the state level. At the local level, these positions are occupied by the PT and the PV. But there is also a series of inconsistencies in the configuration presented in the races for the Boa Vista mayor's office, among which I highlight the difficulty in identifying which would be the left and right-wing poles given the dispersion of parties in each field.

In Santa Catarina, the scenario is different – nevertheless, without resulting in the ideological alignment of the programmatic competition (Graph 27). On the left are the PCO, PSTU, PCB, and the PSOL (2018, 2014), at the state level, and only the PSTU at the municipal level. The party closest to the center in both cases is the PSOL in 2010 and 2012. All remaining parties are placed in the right, which includes the PPL, PT, PEN, DEM, PSD, PRP, PSL, PP, PV, REDE, PSDB, and PMDB, at the state level, and the PSOL, PPL, PSB, PCdoB, PSD, PEN, PP, and PMDB, at the municipal level.

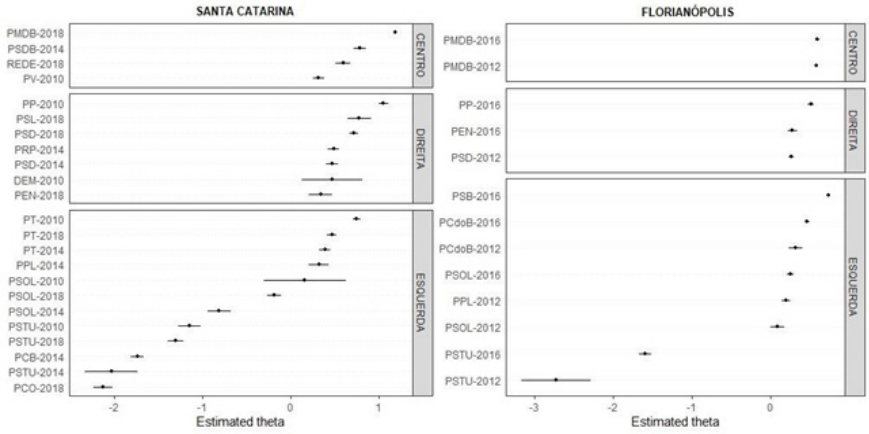
The distribution of the left-wing parties' positions along the two programmatic poles also happens in the state of São Paulo (Graph 28). In addition, center and right-wing parties present a similar location, as seen in other states. In this context, only the PCO, PSTU, PCB, and PSOL are on the left, with the remaining parties dispersing along the center and right. The two most extreme positions of the whole spectrum are also on the left. The election for the São Paulo mayor's office seems strongly aligned to this configuration with the difference that, at this level, center parties are also positioned in the left, specifically the PMDB (2012). At the local level, the extreme positions are not concentrated on the left, being divided among the PCO and the REDE.

Graph 26 – Programmatic competition in Roraima (2010-2018)



Source: the author.

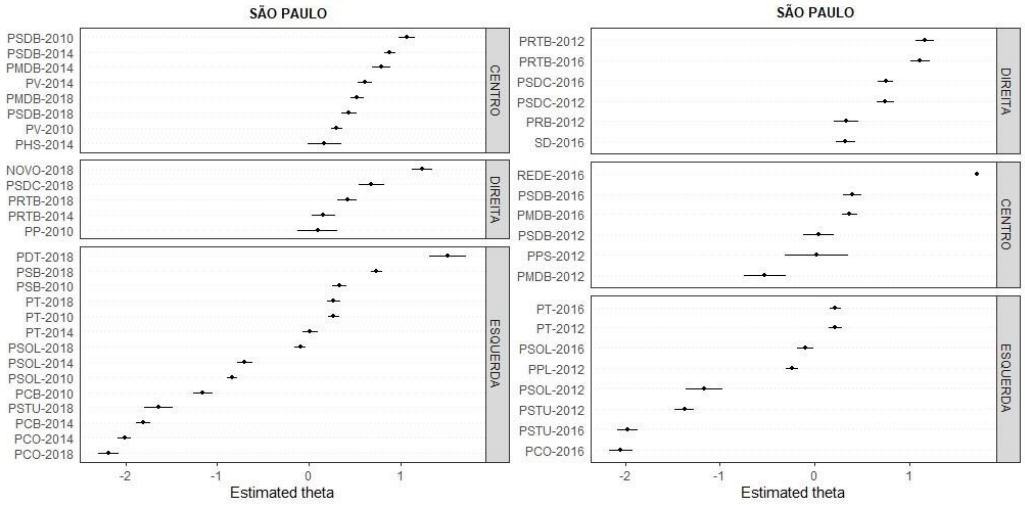
Graph 27 – Programmatic competition in Santa Catarina (2010-2018)



Source: the author.

Sergipe is one of the cases where there is little divergence – and, therefore, programmatic competition – at the state level (Graph 29). This can be identified with the majority of the parties positioned along a vertical line close to the center. In this scenario, the manifestos slightly more distinct are the ones by the PT and the PSB and, in a more acute way, the PSL and the PMDB. The positions obtained in the Aracaju mayoral races, however, present larger divergence. Nevertheless, the ideological dimension seems to be insufficient to explain this configuration, since parties of all ideological classifications are found dispersed in the programmatic spectrum.

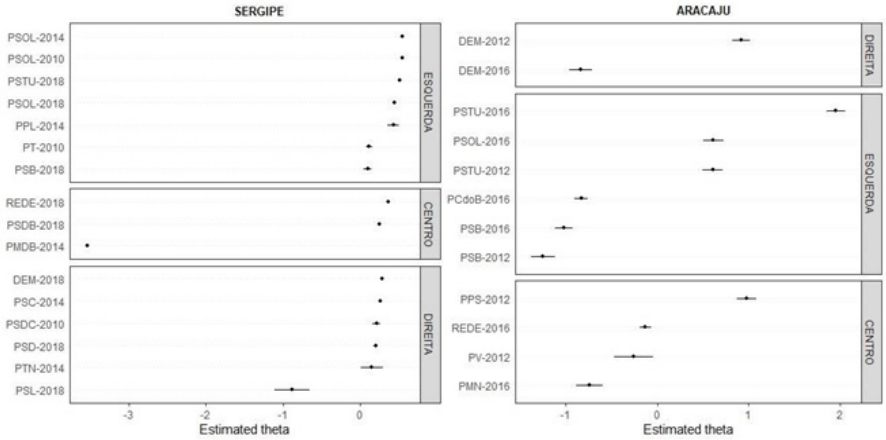
Graph 28 – Programmatic competition in São Paulo (2010-2018)17



Source: the author.

17 Following the graph's pattern, the first image refers to the state and the second one to the city.

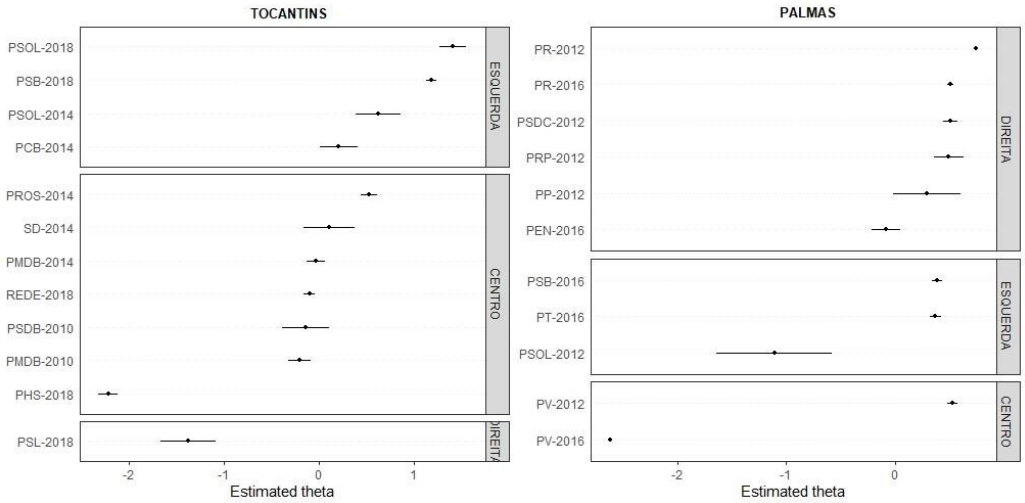
Graph 29 – Programmatic competition in Sergipe (2010-2018)



Source: the author.

The elections for state government in Tocantins were the ones that, along with Goiás, presented stronger ideological alignment (Graph 30). Thus, it is possible to observe that the PSOL, the PSB, and the PCB are located in the left; the PMDB, PSDB, REDE, and the SD in the center; the PSL in the right. Only two positions can be considered inconsistent regarding the ideological classification: PHS's, more towards the right than the PSL, and PROS's in the center-left. It is important to observe that the PCB, although situated on the left, is positioned much closer to the center. At the municipal level, the divergence among the parties is reduced, making them positioned almost in a single pole, between 0 and 1. With the exception of the PV (2016) and the PSOL, which have a more extreme location on the same side, the PR, PSDC, PRP, PP, PEN, PSB, PT, and the PV (2012) share the same programmatic space and, due to this reason, seems to organize themselves from other dimensions than the ideological one.

Graph 30 – Programmatic competition in Tocantins (2010-2018)



Source: the author.

The box below summarizes the results obtained, grouped in three categories. The first refers to the *ideological* organization of the programmatic competition, that is, to the cases in which spatial positions were consistent with the classification given by the literature. The opposing category is the one I call *non-ideological*.

With this I do not mean to say that ideology has no relation with these arrangements but that, when analyzed through the proposed method and as an exclusive dimension, the locations of the parties is significantly distinct from those attributed by the literature. That is, this variable does not have explanatory power. Lastly, an intermediary category, of *partially ideological* programmatic competitions, in which ideology seems to determine the position of parties, but not exclusively. Consequently, it is necessary to include other dimensions in the analysis. In addition, I have also classified the cases in relation to the alignment among levels, that is, if the configurations obtained at state and municipal levels were more or less distinct.

Box 4 – Summary of the ideological configuration of the programmatic competition

State programmatic competition			Municipal programmatic competition			Relationship between levels	
Ideological	Partially ideological	Not ideological	Ideological	Partially ideological	Not ideological	Aligned	Not aligned
GO	AM	AC	BA	AC	AL	AL	AC
TO	AP	AL	PB	MA	AM	AP	AM
	BA	DF		MT	AP	CE	BA
	CE	ES		MS	CE	ES	GO
	MA	MT		MG	ES	MA	MT
	MS	PA		PE	GO	MS	PR
	MG	PE		RJ	PA	MG	PI
	PB	RJ		RS	PR	PA	RS
	PR	RN		SP	PI	PB	RR
	PI	RS			RN	PE	SC
	SP	RO			RO	RJ	SE
		RR			RR	RN	TO
		SC			SC	RO	
		SE			SE	SP	
					TO		

Source: the author.

Overall, the results obtained show the few races, only two cases in each level, are ideologically organized. Nevertheless, at the state level, if these scenarios are added to the ones of partially ideological programmatic competitions, they are nearly half the cases, totaling 13 states. The remainder (14) did not show congruence between the position attributed to parties and their ideological classifications. At the local level, this number is slightly larger, 15 out of 26; in the other nine capitals, the programmatic competition was organized by ideology only partially. Regarding the alignment between the two levels of competition, 14 out of 26 states presented similar configurations of programmatic competition.

These indications confirm that it is not because parties distinguish the content of their manifestos that this divergence will be ideologically organized. In presidential elections, ideological congruence happens much more acutely then at other levels, with more emphasis at the state than the municipal level. This suggests that a competition often cannot be explained unidimensionally, in a left- right scale, but that other dimensions seemed to have an impact. One of them can be competitiveness of the race or the parties themselves, as has been pointed out by *salience theory*, given that in the majority of cases left-wing parties such as the PSTU, PCO, PCB, and the PSOL seemed to exert a centripetal force among others in their ideological field. Therefore, more central

positions in the left usually referred to more competitive parties, such as the PSB and the PT.

Regarding method, *wordfish* seems to have led to more consistent classifications than the ones obtained through the *MARPOR* coding, even in cases where it was adapted to the Brazilian context. However, if in this technique the majority of parties was considered left or center-left, in *wordfish*, in some cases, such as Palmas, for instance, positioning parties more in the center-right. Although the need for new exercises is evident, including new dimensions, and even elaborated through more sophisticated techniques, that advance beyond the frequency of isolated words (bag-of-words), *wordfish* acts as an efficient starting point, indicating paths towards this.

6. Discussion

Manifestos are, by nature, instruments of political positions, since they are a list of problems (and solutions) of the day, that must be faced politically. Despite this, the use of these objects for the ideological classification of parties is not immune from challenges and contradictions, as I have sought to demonstrate throughout this paper.

Overall, there are two main issues related to the concept of ideology itself that have an impact in its measuring through electoral platforms. The first is related to the dimensionality of political space. Given that left and right conform to the same dimension, are they sufficient to explain the parties' positions? In this context, there is a persistent dilemma between reducing the political space to apprehend it without that reduction eliminating essential characteristics that compromise the analysis. The second point refers to the substantive content of left and right categories and how it may vary in different contexts.

In Brazil, these issues have advanced little when it comes to the relationship between ideology and manifestos. As presented, in addition to the studies on this topic being few, the majority of the ones that have done so used the manual coding *MARPOR*. The combination of fixed categories and substantive content of the scale has led to a problematic classification of Brazilian parties, with a nearly exclusive predominance of parties on the left.

This classification also seems to indicate a methodological problem. That is why, in this paper, I proposed an exercise of analyzing manifestos with an automated textual analysis technique that attributes spatial positions from the frequency of words used by candidates. Although that is not the only possible technique, and it has limitations, its use is justified by the fact that ideology is, intrinsically, a spatial concept.

The analysis of manifestos by candidates to executive offices in Brazil since 2010, at the different levels of competition, offered a more complex scenario – albeit imperfect – when compared to the studies that employed the *MARPOR* coding. Supporting it, there is an assumption of a theoretical nature, that a programmatic competition does not necessarily become an ideological dispute, although an ideological dispute is always programmatic. In other words, if the candidates compete programmatically, they elaborate proposals distinct from their opponents. However, they may or may not organize ideologically, as they could or not organized in parties, given a set of contests, and so on. The findings of this paper reinforced this argument. For the presidential elections considered, there was an ideological organization of the race. However, this structuring force partially lost its impact at the state level and then slightly more at the local level. This means that, in these spheres, not always did the position attributed to parties through their electoral platforms converge with the ideological classification found in the literature.

Although this find is not trivial, there are many questions that need to be more deeply observed, as other dimensions and variables that may determine the position of the parties in the programmatic spectrum. Thus, the results of this study point to the need of relating them with electoral data, such as the degree of competitiveness of race and parties. After all, as highlighted, less competitive parties seem to exert a centripetal force over remaining ones in the same ideological field.

From the point of view of method, future efforts should be able to discuss the limitations of *wordfish*, both given its unidimensionality and considering the frequency of words in isolation without pondering other characteristics, such as order and meaning. Given the importance of the theme, especially in contexts of the prominence of the new right, the main and future challenges of an agenda interested in understanding the relationship between programmatic competition and ideology in Brazil are laid out.

BIBLIOGRAPHICAL REFERENCES

ALDRICH, John. (1983), "A downsian spatial model with party activism". *American Political Science Review*, vol. 77, no. 4: 974–990.

ALESINA, Alberto. (1988), "Credibility and policy convergence in a two-party system with rational voters". *The American Economic Review*, vol. 78, no. 4: 796–805.

ALLAN, James; SCRUGGS, Lyle. (2004), "Political partisanship and welfare state reform in advanced industrial societies". *American Journal of Political Science*, vol. 48, no. 3: 496-512.

BABIRESKI, Flávia. (2014), "As diferenças entre a direita do Brasil, Chile e Uruguai: análise dos programas e manifestos partidários". *Paraná Eleitoral*, vol. 3, no. 1: 171–198.

BENOIT, Kenneth; LAVER, Michael. (2006), *Party Policy in Modern Democracies*. New York, Routledge.

BOBBIO, Norberto. (1995). *Direita e esquerda: razões e significados de uma distinção política*. São Paulo, Ed. UNESP.

BUDGE, Ian; FARLIE, Dennis. (1983), *Explaining and predicting elections: issue effects and party strategies in twenty-three democracies*. London/Boston, Unwin Hyman.

CHAPPELL JR., Henry.; KEECH, William. (1986), "Policy motivation and party differences in a dynamic spatial model of party competition". *American Political Science Review*, vol. 80, no. 3: 881–899.

CODATO, Adriano; BERLATTO, Fábia; BOLOGNESI, Bruno. (2018), "Tipologia dos políticos de direita no Brasil: uma classificação empírica". *Análise Social*, no. 229: 870-897.

DE SIO, Lorenzo; WEBER, Till. (2014) "Issue yield: A model of party strategy in multidimensional space". *American Political Science Review*, vol. 108, no. 4: 870–885.

DOWNS, Anthony. (1957), *An economic theory of democracy*. New York, Harper.

GRIMMER, Justin; STEWART, Brandon. (2013), "Text as data: the promise and pitfalls of automatic content analysis methods for political texts". *Political Analysis*, vol. 21, no. 3: 267– 297.

GROFMAN, Bernard. (2004), "Downs and two-party convergence". *Annual Review of Political Science*, vol. 7: 25-46.

HOPKINS, Daniel; KING, Gary. (2010), "A method of automated nonparametric content analysis for social science". *American Journal of Political Science*, vol. 54, no. 1: 229-247.

IZUMI, Maurício; MOREIRA, Davi. (2018), "O texto como dado: desafios e oportunidade para as ciências sociais". *BIB*, vol. 2, no. 86: 138-174.

JORGE, Vladimyr; SILVA, Mayra; FARIA, Alessandra; FERREIRA, Ana Rita. (2018), "Análise dos programas eleitorais dos candidatos a presidente em 2014: o posicionamento ideológico do PT e do PSDB". *Revista de Sociologia e Política*, vol. 26, no. 67: 1–20.

KLINGEMANN, Hans-Dieter; VOLKENS, Andrea; BUDGE, Ian; BARA, Judith; MCDONALD, Michael. (2006), *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union, and OECD 1990-2003*. Oxford, Oxford University Press.

LAVER, Michael; BUDGE, Ian. (1992), *Party policy and coalition policy in Europe*. London, Macmillan.

LAVER, Michael; BENOIT, Kenneth; GARRY, John. (2003), "Extracting policy positions from political texts using word as data". *Political Science Review*, vol. 97, no. 2: 311–331.

LEHMANN, Pola; MATTHIEß, Theres; MERZ, Nicolas; REGEL, Sven; WERNER, Annika. (2019), *Manifesto Corpus*. Version: 2019b. Berlin: WZB Berlin Social Science Center.

LOWE, Will. (2016), "Scaling things we can count". Working paper, *American Political Science Association meeting.*

MADEIRA, Rafael; VIEIRA, Soraia; TAROUCO, Gabriela. (2017), "Agendas, preferências e competição: PT e PSDB em disputas presidenciais". *Caderno CRH*, vol. 30, no. 80: 257–273.

MONROE, Burt; COLARESI, Michael; QUINN, Kevin. (2008), "Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict". *Political Analysis*, vol. 16, no. 4: 372-403.

MONROE, Burt; SCHRODT, Philip. (2008), "Introduction to the special issue: The statistical analysis of political text". *Political Analysis*, vol. 16, no. 4: 351–355.

OLIVEIRA, Augusto. (2011), "Os discursos programáticos dos partidos políticos na América Latina: identidade de classe e política econômica". *Mediações*, vol. 16, no. 1: 201–221.

OLIVEIRA, Augusto. (2014), *Representação Programática em 16 Democracias Presidencialistas: América Latina, 2000-2010*. 226 p. Tese (Doutorado em Ciência Política) – Universidade Federal do Rio Grande do Sul, Porto Alegre.

PETROCIK, John. (1996), "Issue ownership in presidential elections, with a 1980 case study". *American Journal of Political Science*, vol. 40, no. 3: 825–850.

REPASS, David. (1971), "Issue salience and party choice". *American Political Science Review*, vol. 65, no. 2: 389–400.

RIKER, William. (1982), "The two-party system and Duverger's Law: An essay on the history of political science." *American Political Science Review*, vol. 76, no. 4: 753–766.

ROBERTSON, David. (1976), *A Theory of Party Competition*. New York, Wiley.

ROEMER, John. (2001), *Political competition: theory and applications*. Cambridge, Harvard University Press.

SARTORI, Giovanni. (1976), *Parties and Party Systems: a framework for analysis*. Colchester: ECPR Press.

SCHMITTER, Philippe; KARL, Terry. (1991), "What democracy is... and is not". *Journal of Democracy*, vol. 3, no. 2: 75–88.

SALLES, Nara. (2020), "Do paradoxo à competição: o lugar da dimensão programática nas disputas eleitorais". *Revista Brasileira de Ciência Política*, no. 32: 93-134.

SALLES, Nara; GUARNIERI, Fernando. (2019), "Estratégia eleitoral nos municípios brasielrios: componente programático e alinhamento partidário". *Revista de Sociologia e Política*, vol. 27, n. 72: e001.

SLAPIN, Jonathan; PROKSCH, Sven-Oliver. (2008), "A scaling model for estimating time- series party positions from texts". *American Journal of Political Science*, vol. 52, no. 3: 705-722.

STOKES, Donald. (1963), "Spatial models of party competition". American Journal of Political Science, vol. 34, no. 2: 565–598.

TAROUCO, Gabriela. (2011), "Brazilian parties according to their manifestos: Political identity and programmatic emphases". *Brazilian Political Science Review*, vol. 5, no. 1: 54–76.

TAROUCO, Gabriela; MADEIRA, Rafael. (2013), "Partidos, programas e o debate sobre esquerda e direita no Brasil". *Revista de Sociologia e Política*, vol. 21, no. 45: 149–165.

VOLKENS, Andrea; LEHMANN, Pola; MATTHIEß, Theres; MERZ, Nicolas; REGEL, Sven; WERNER, Annika. (2016), *The Manifesto Project Dataset - Codebook.* Manifesto Project (MRG/CMP/MARPOR). Version 2016a. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

WITTMAN, Donald. (1983), "Candidate motivation: a synthesis of alternative theories".

American Political Science Review, vol. 77, no. 1: 142–157.

Mapeando comportamentos com estimação de pontos ideais

Rodrigo Martins - Universidade Federal de Pernambuco

Resumo

Estimação de pontos ideais é uma metodologia de importância crescente na Ciência Política. Desenvolvida desde meados do século passado, seu vínculo com a teoria espacial do voto se mostrou frutífera em diversos campos de estudo, com grandes expectativas sobre seu desenvolvimento futuro. O objeto deste trabalho é apresentar o que é estimação de pontos ideais, e ilustrar didaticamente como aplicar uma das várias técnicas disponíveis, utilizando dados sobre a atuação parlamentar da Câmara dos Deputados no ano de 2019.

Palavras-chave: comportamento; preferências; ideologia; pontos ideais

Abstract

Ideal point estimation is a methodology of growing importance in Political Science. Developed in the mid-20th century, its link with the spatial theory of voting has been fruitful for several fields of study, with great expectations on its future development. The goal of this paper is to present what ideal point estimation is and didactically illustrate how to apply one of the various available techniques, using data on parliamentary behavior from the Chamber of Deputies in 2019.

Keywords: behavior; preferences; ideology; ideal point estimation

1. Introdução

O que estrutura as divisões entre parlamentares, ideologia ou governismo? Mudança de partido afeta o comportamento de legisladores? Qual é o nível de coesão interna dos partidos políticos e como ela oscila ao longo do tempo? A polarização no Congresso Nacional aumentou? Quais políticos possuem comportamento mais imprevisível? Quais tipos de projetos de lei costumam polarizar o comportamento político? Como a capacidade de um presidente estabelecer um governo de coalizão influencia o comportamento parlamentar?

Estas são algumas questões levantadas pela literatura de estudos legislativos para explicar a atividade parlamentar. Em comum a todas elas está a necessidade de acessar preferências, ideologias e comportamentos de representantes eleitos. No entanto, campos diversos da Ciência Política possuem o mesmo interesse com relação a outros tipos de atores políticos e outros tipos de comportamento político. É possível encontrar evidências de comportamento ideológico no poder judiciário? Os discursos de políticos expressam diferenças ideológicas? As preferências ideológicas do eleitorado refletem as preferências de seus representantes? A opção de seguir certas pessoas e não outras no Twitter reflete a preferência política dos usuários da plataforma?

Essas e diversas outras questões são levantadas por estudos que utilizam estimações de pontos ideais para explicar comportamento de atores políticos. A teoria espacial do voto, já teorizada por (Downs, 1957), passou a ser aplicada às mais diversas áreas, não apenas para explicar políticos tradicionais como também para lançar luz sobre as preferências daqueles que estão fora das instituições políticas. A facilidade de se traduzir formalizações teóricas de formação de preferências em fórmulas matemáticas e estatísticas fez com que a teoria espacial do voto obtivesse sucesso para ser implementada de diversas formas.

O objetivo deste trabalho é apresentar, de forma simples e intuitiva, o que é estimação de pontos ideais e ilustrar como utilizá-la didaticamente. Apesar de utilizada para os mais diversos objetivos, o foco aqui será no campo de estudos que mais desenvolveu esta técnica, a área de estudos legislativos. Ilustraremos como estimar pontos ideais utilizando dados sobre da Câmara dos Deputados no ano de 2019, durante o primeiro ano do mandato presidencial de Jair Bolsonaro. Estudos anteriores, que utilizaram estimação de pontos ideais para o caso brasileiro, mostram que a dimensão principal que estrutura o comportamento parlamentar do Senado e Câmara dos Deputados não é a ideológica, mas sim a dimensão governo-oposição (Izumi, 2016; Zucco, 2009), embora em alguns momentos tais dimensões pareçam estar sobrepostas (Leoni, 2002). Diante do diagnóstico de que Bolsonaro alterou a relação entre Executivo e Legislativo ao não compor base de apoio parlamentar por meio de uma coalizão, substituindo "o presidencialismo de coalizão pelo desleixo" (Limongi, 2019), haveria uma mudança no comportamento da Câmara dos Deputados? Haveria uma outra dimensão estruturando as votações nominais em plenário?

2. O que são pontos ideais?

Estimação de pontos ideais deriva do modelo espacial do voto, inicialmente popularizada na Ciência Política por Downs (1957), onde atores politicos podem ser representados como um ponto em um espaço e optam por alternativas políticas, também representadas espacialmente, que estão mais próximas de si. Além de ser um princípio teórico, diversos estudos mostram empiricamente que a maior parte do comportamento político pode ser representado em um espaço unidimensio-

nal, frequentemente tido como uma escala ideológica. A divisão entre esquerda e direita seria uma expressão de como atores políticos, sejam estes eleitores ou representantes, se comportam e manifestam suas preferências de maneira alinhada em assuntos distintos, como política econômica e política social, no mesmo sentido em que Converse (1964) argumenta na direção de um sistema de crença que permeia preferências de assuntos diversos em uma visão unificada de mundo.

A técnica de estimação de pontos ideais nos mostra um mapa de como votantes se distribuem espacialmente, representando a distância entre as preferências de cada um dos indivíduos em análise. De maneira simplificada, o que extraímos é a proximidade entre os votantes quando consideramos o conjunto de todos os indivíduos, e não apenas a semelhança entre eles par a par. Dessa forma é possível, inclusive, estimar a proximidade entre votantes que na prática não participaram juntos de nenhuma votação, pois ao estimar-se a distância destes com relação a todos os outros votantes, saberíamos também a distância entre eles. Um dos pressupostos do modelo é que ao sabermos como um indivíduo decide sobre um tema em específico, seria possível prever seu comportamento em relação a outras decisões, especialmente em temas semelhantes, a partir da derivação da decisão anterior. Dessa forma, ao correlacionar todas as decisões, seria possível estimar coordenadas para os indivíduos, de maneira que aqueles que decidem de forma semelhante fiquem espacialmente próximos, e aqueles que decidem de forma divergente fiquem espacialmente distantes.

Para ilustrar a ideia subjacente à estimação de pontos ideais, podemos recorrer a um exemplo hipotético. Suponhamos que exista um colegiado de três legisladores que precisam tomar uma decisão sobre uma matéria que diz respeito a temática econômica. Considerando que um dos legisladores vota por favorecer mais intervenção do estado na economia e os outros dois votam por menos intervenção, a representação espacial do voto dos legisladores será feita de tal forma que os dois que votaram juntos ocupem o mesmo lado, em oposição ao votante que ficou isolado. Também é possível ilustrar a matéria em votação como uma linha que representa a divisão entre os legisladores em dois lados diferentes. A figura 1 seria a forma de representar espacialmente a divisão neste exemplo. Os círculos vermelho e verde seriam os dois que votaram juntos, o círculo azul o indivíduo que votou isolado e a reta roxa seria a linha da matéria em votação que representa o corte entre os dois lados em questão.

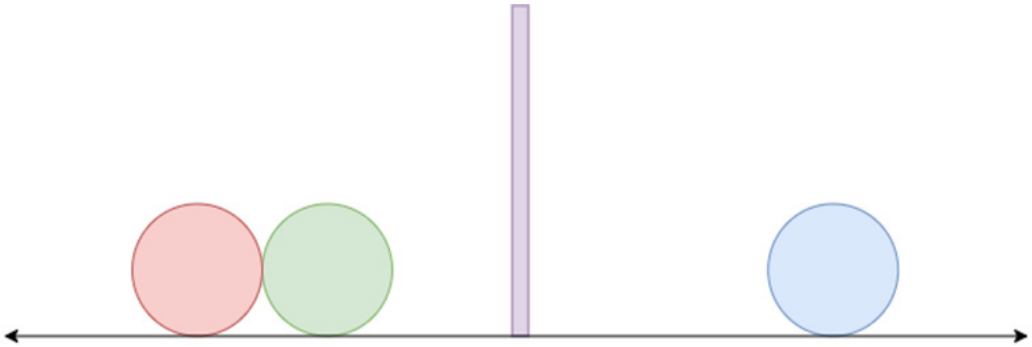


Figura 1: Exemplo de análise espacial com 1 votação e 3 votantes

Suponhamos agora que em uma segunda votação sobre o mesmo tema o legislador representado pela cor verde não vote novamente junto com o votante representado pela cor vermelha, mas sim com o representado pela cor azul. Neste segundo momento, o indivíduo verde deve estar posicionado entre os outros dois de forma equidistante, já que votou uma vez junto com cada um. Ao mesmo tempo, aqueles que não votaram juntos em nenhuma das duas votações devem permanecer em extremos opostos. A linha que representa a primeira votação, em cor roxa, deve manter a divisão entre os votantes da mesma forma que anteriormente. E agora uma segunda linha de votação, que será representada pela cor branca na figura 2, deve deixar isolado o legislador que votou sozinho e manter do mesmo lado aqueles que votaram juntos.

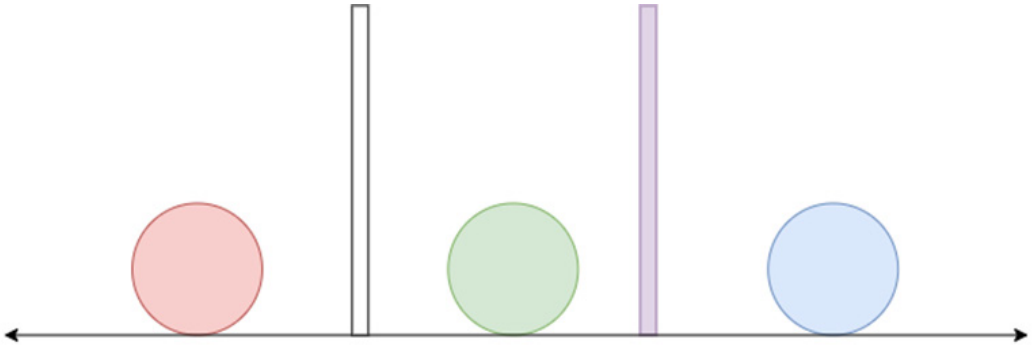


Figura 2: Exemplo de análise espacial com 2 votações e 3 votantes

Como último exemplo, imaginemos que em uma terceira votação sobre outro assunto não relacionado a economia os legisladores que antes se opunham agora votam juntos, deixando o votante que antes estava no centro isolado na minoria. Para que a representação espacial se mantenha mais apropriada a esta situação, o mais adequado é adicionar um eixo vertical onde aquele que foi derrotado se distancie dos outros dois, enquanto os dois vencedores se mantêm próximos um do outro neste segundo eixo. Também podemos representar esta terceira votação como uma linha, representada em amarelo na figura 3, mas desta vez na horizontal.

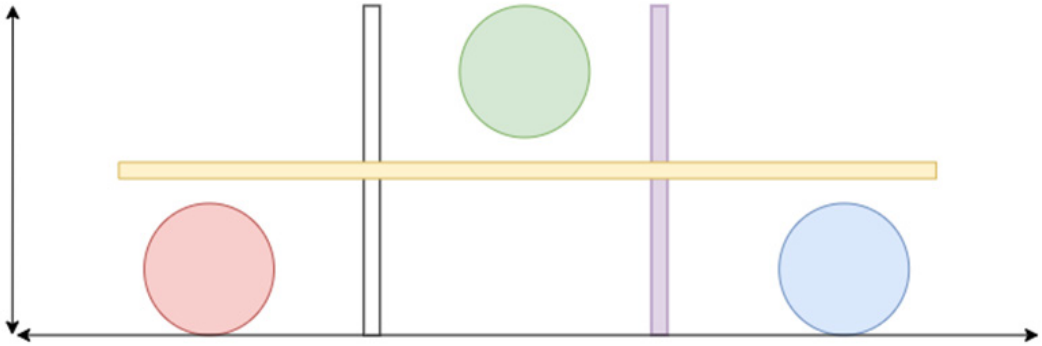


Figura 3: Exemplo de análise espacial com 3 votações e 3 votantes

Esse tipo de modelo projeta as preferências manifestadas através do voto em um espaço (unidimensional, bidimensional e até mesmo além) avaliando as similaridades e divergências entre as decisões de cada um dos votantes. Como não seria factível fazer esse processo manual para muitos votantes em muitas votações, foram desenvolvidas diversas técnicas, com procedimentos estatísticos distintos entre si, com o objetivo de estimar pontos ideais tanto para votantes quanto para votações.

No entanto, é de se imaginar que em algum momento não é mais possível representar perfeitamente a distribuição espacial dos votantes e das votações utilizando-se apenas uma ou duas dimensões. Com o aumento de votos e votações, torna-se inevitável que os pontos de votantes acabem sendo alocados em uma posição que não corresponde ao seu voto em dada matéria. (Poole, 2005) propõe duas medidas para avaliar e comparar o desempenho de modelos de pontos ideais: a porcentagem de classificações corretas e a redução proporcional do erro agregado (RPEA). Para explicar melhor o que são tais medidas e quais as diferenças entre elas, vamos utilizar um exemplo hipotético como ilustração. A figura 4 representa uma votação com 11 votantes identificados por uma letra e uma cor.

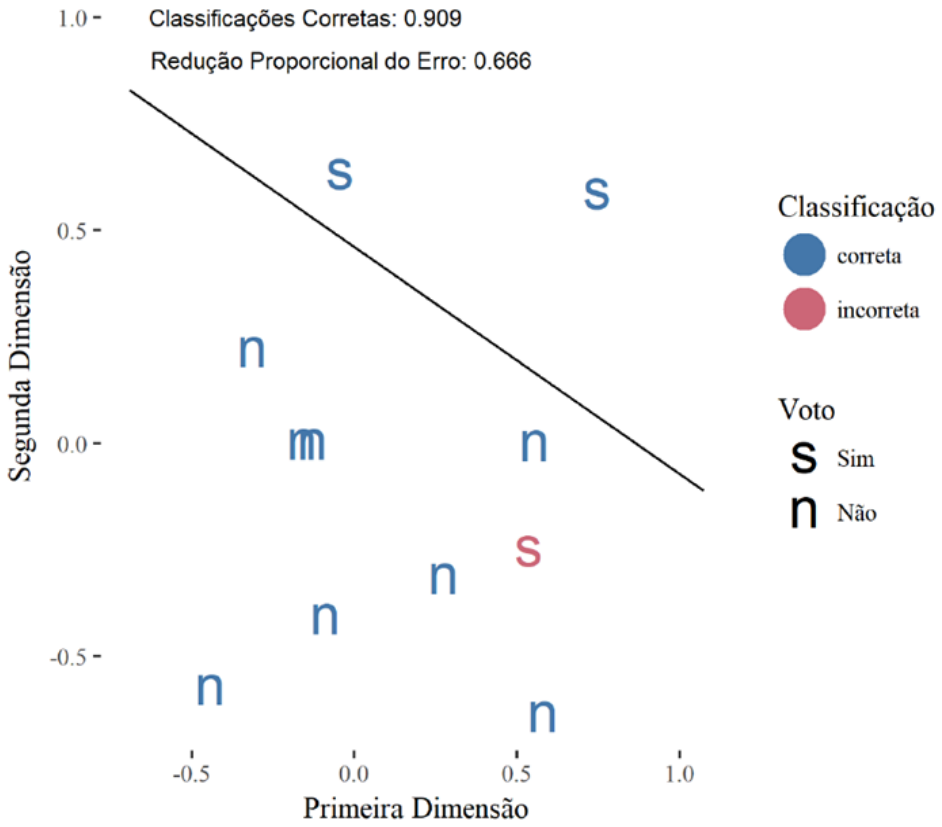


Figura 4: Exemplo para avaliação de medidas de desempenho

Votantes representados pela letra "s" votaram "sim" e votantes representados pela letra "n" votaram "não". O único votante representado pela cor vermelha foi o votante que foi classificado de forma errada pelo modelo, pois votou "sim" enquanto todos os outros representados no mesmo lado da votação votaram "não". Este deveria estar do outro lado da linha, juntamente com os outros que votaram "sim" e foram classificados da forma correta. Dessa forma, a porcentagem de classificação correta desta votação seria de 0.90922[1]. A redução proporcional do erro (RPE) seria de 0.66[2]. Tal medida seria mais adequada e criteriosa para avaliar o desempenho dos modelos. Usando o exemplo de (Armstrong et al., 2014), se em alguma votação houver uma divisão de 65-35 votos a favor de uma proposta, poderíamos classificar corretamente 65% dos votos apenas ao dizer "ingenuamente" que todos os votantes apoiarão a proposta. A redução proporcional do erro mede o quanto o modelo melhora a classificação dos votos em relação a minoria. Voltando ao nosso exemplo hipotético da figura 4, a RPE seria de 0.66 pois classifica corretamente dois dos três votantes da minoria. A RPEA seria uma medida que agrega a RPE de cada uma das votações.

É importante salientar que tal método se utiliza de votações que foram decididas de forma dividida, ou seja, não unânime. Tal critério é importante pois uma vez que se procura identificar como se dá a divergência entre votantes, a matéria julgada deve necessariamente produzir dissenso entre estes. Votações unânimes não têm como nos fornecerem informações sobre agrupamentos e divergências, dessa forma são desconsideradas. Essa característica pode afetar estudos sobre instituições que possuem grande taxa de consenso. No Supremo Tribunal Federal, por exemplo, onde a maior parte das decisões são monocráticas ou unânimes, utilizar esse método nos faria perder capacidade explicativa sobre o tribunal como um todo. No entanto, ao utilizarmos apenas decisões majoritárias, observaríamos casos que tratam de matéria controversa, sujeita então a disputas de interpretações distintas e conflitos políticos. É mais provável que casos importantes e de maior repercussão se enquadrem nesta situação.

É necessário cuidado ao analisar a dispersão dos votantes no espaço, pois o modelo não possui capacidade de inferir o que determina essa dispersão, ou seja, qual é o fator que estrutura o voto nas dimensões espaciais. Este tipo de estimação estatística foi utilizado na Ciência Política brasileira, por exemplo, para analisar a Câmara dos Deputados (Leoni, 2002; Zucco, 2009; Zucco e Lauderdale, 2011), Senado (Izumi, 2016) e Supremo Tribunal Federal (Desposato et al., 2015; Ferreira e Mueller, 2014; Mariano Silva, 2018). No caso da Câmara dos Deputados e Senado, não é a divisão ideológica que acaba sendo evidenciada com as estimações de pontos ideais, mas sim a divisão entre oposição e governo.

O primeiro modelo estatístico de estimação de pontos ideais elaborado na Ciência Política derivado diretamente da teoria espacial do voto foi o NOMINATE (Poole e Rosenthal, 1985, 1991, 1997). Este é um modelo paramétrico aplicado a votações binárias e amplamente utilizado na área de estudos legislativos. Com diversos desenvolvimentos posteriores implementados em R, o mais popular deles é o W-NOMINATE (Poole et al., 2011), enquanto o D-NOMINATE e DW-NOMINATE são utilizados para estimação de pontos ideais dinâmicos, para observar a oscilação do comportamento dos votantes ao longo do tempo.

Com o aumento de performance computacional, novas implementações de estimações de pontos ideais foram elaboradas com a aplicação de estatística bayesiana. Derivados da teoria da resposta ao item, bastante desenvolvida na área de psicometria, e desenvol-

1 número de classificações corretas / número de votantes

2 (votos na minoria - classificações erradas) / votos na minoria

vidos a partir dos trabalhos de Clinton et al. (2004) e Martin e Quinn (2002), a implementação em R do IDEAL no pacote **pscl** é um dos mais populares, junto com o pacote **MCMCpack** que conta com uma implementação do modelo dinâmico bayesiano entre outras. A teoria da resposta ao item é utilizada principalmente em estudos educacionais, onde procuram medir o conhecimento de indivíduos a partir de testes, substituídos na Ciência Política pela mensuração de ideologia a partir de votações.

Com implementações distintas, os modelos da família NOMINATE e os bayesianos acabam por ter pressupostos ligeiramente distintos. Enquanto ambos assumem que os erros são independentes, identicamente distribuídos, com distribuição normal e com menor probabilidade de ocorrer na medida em que o ponto ideal do votante é distante da reta da votação, a forma como cada um dos votantes trata as opções disponíveis ao votarem *sim* ou *não* difere. No NOMINATE, se ambas as alternativas estiverem distantes do ponto ideal do votante, este é mais indiferente as opções oferecidas, pois sua curva de preferência se dá em um formato gaussiano. Nos modelos bayesianos, de maneira distinta, a curva de preferências dos votantes assume um formato quadrático, fazendo com que mantenham a sensibilidade diante de opções distantes de seu ponto ideal. Apesar de, na prática, esse pressuposto não produzir muita diferença de resultado entre os modelos (Carroll et al., 2009), para lidar com este problema foi criado o α-NOMINATE (Carroll et al., 2013). Esta implementação bayesiana do NOMINATE foi feita para substituir as implementações anteriores, e estima se a melhor forma funcional das preferências dos votantes é gaussiana, quadrática ou uma mistura intermediária de ambas.

Neste trabalho, diante de diversas implementações estatísticas do modelo espacial para realizar estimações de pontos ideais, demonstraremos como utilizar o *Optimal Classification* (OC). Esta é uma técnica de estimação de pontos ideais não paramétrica, ao contrário dos modelos anteriores. Seu algorítimo consiste nos seguintes passos (Poole, 2005, p. 82):

1. Gerar valores iniciais para os pontos ideais dos votantes a partir da decomposição em autovalores e autovetores da matriz de concordância entre os pares de votantes

2. Dado os valores dos votantes, encontrar as estimações dos vetores das votações que maximizam o número de classificações corretas.

3. Dados os valores dos vetores de votações, encontrar novos valores dos pontos ideais dos votantes que maximizem o número de classificações corretas.

4. Voltar ao passo 2.

Tal processo se repete até que o número de classificações corretas se estabilize, maximizando o número de acertos. De acordo com Poole, tal técnica é consistente com dados com poucos votantes e poucas votações e uma de suas vantagens seria o fato de possuir menos pressupostos quando comparado aos modelos paramétricos de estimação de ponto ideal, tornando-o menos suscetível a problemas estatísticos de estimação.

Escolhemos esta técnica pois, por ser não-paramétrica, necessita que menos pressupostos estatísticos e sobre o comportamento dos votantes sejam satisfeitos. Rosenthal e Voeten (2004) argumentam que o comportamento estratégico que legislaturas podem apresentar, devido a disciplina partidária, seria menos problemático para o *Optimal Classification* quando comparado com outros modelos paramétricos[3].

3 Este problema não é plenamente resolvido por esta técnica (Spirling e McLean, 2007), pois o pressuposto de monotonicidade ainda persiste. Ver Duck-Mayr e Montgomery ([S.d.]) para um aprofundamento sobre esta questão.

Na seção seguinte, iniciaremos a explicação de como implementar uma estimação de pontos ideais, apresentando o código em R dos comandos e funções necessárias para a utilização deste método.

3. Primeiro passo: criar um objeto *rollcall*

O pacotes que implementam o *Optimal Classification*, *W-NOMINATE*, α*-NOMINATE* e *IDEAL* necessitam que criemos no R um objeto de tipo *rollcall* para que usemos suas respectivas funções. Para isso, é necessário especificar os seguintes elementos na função rollcall():

1. os dados onde cada linha é um votante e cada e coluna é uma votação

2. os valores correspondentes ao voto "sim"

3. os valores correspondentes ao voto "não"

4. os valores correspondentes a *missings*, quando o votante não votou por algum motivo

5. os valores correspondentes a quando o votante não fazia parte daquela legislatura

6. os valores correspondentes aos nomes dos legisladores, na mesma ordem em que aparecem nas linhas do banco de dados

7. os valores correspondentes aos nomes das votações, na mesma ordem em que aparecem nas colunas do banco de dados

8. os valores correspondentes a informações extras sobre os votantes, por exemplo seus partidos, na mesma ordem em que aparecem nas linhas do banco de dados

9. os valores correspondentes a informações extras sobre as votações, por exemplo suas datas, na mesma ordem em que aparecem nas colunas do banco de dados

São necessários obrigatoriamente os 3 primeiros tipos de informações, os outros são opcionais. O código abaixo ilustra o formato de um banco de dados hipotético, e a maneira de se criar um objeto *rollcall* a partir dele:

```
# Observar o banco
banco_hipotetico[1:5,1:13]

##          name      state      party V1 V2 V3 V4 V5 V6 V7 V8 V9 V10
## 1 Legislator1     Alaska   Democrat  1  1  6  6  1  1  6  1  1   1
## 2 Legislator2      Texas   Democrat  1  1  6  1  1  6  6  6  1   1
## 3 Legislator3       Ohio Republican  1  1  6  6  1  6  6  6  6   6
## 4 Legislator4      Texas   Democrat  1  1  1  1  1  1  1  1  1   6
## 5 Legislator5 California   Democrat  6  6  6  1  6  6  6  6  6   1

# Criar um objeto roll call a partir do objeto banco_hipotetico
rollcall_hipotetico <- rollcall(banco_hipotetico[,-c(1:3)],
                                yea = 1,
                                nay = 6,
                                legis.names = banco_hipotetico$name,
                                legis.data = banco_hipotetico[,2:3])

rollcall_hipotetico

## Number of Legislators:    20
## Number of Votes:  100
##
## Using the following codes to represent roll call votes:
## Yea:      1
## Nay:      6
## Abstentions:  NA
## Not In Legislature:   9
##
## Legislator-specific variables:
## [1] “state” “party”
## Detailed information is available via the summary function.
```

Primeiro observamos as primeiras 5 linhas do banco de dados e as 13 primeiras colunas para ter uma ideia de sua estrutura e organização. A primeira coluna representa o nome de cada legislador, a segunda coluna seu estado de origem e a terceira seu partido. Daí em diante as colunas representam votações, onde o valor 1 significa voto favorável e o valor 6 significa voto contrária a matéria. Para criar um objeto de tipo *rollcall*, o primeiro argumento especificado é o banco de dados delimitando apenas as colunas de votação, portanto excluindo-se as colunas de 1 a 3. Em seguida são especificados os argumentos *yea* e *nay* com os valores que correspondem a *sim* e *não*, respectivamente 1 e 6. Depois especificamos os dois argumentos opcionais, *legis.names* indicando a coluna *name* do objeto *banco_hipotetico* como o nome dos legisladores, e as colunas 2 e 3 no argumento *legis.data*, pois estas duas colunas possuem informações extras sobre os legisladores. Caso quiséssemos nomear as votações e incluir informações extras sobre cada uma delas, especificaríamos os argumentos *vote.names* e *vote.data* com tais valores. Ao observamos o objeto criado, *rollcall_hipotetico*, podemos averiguar a estrutura dos dados, indicando o número de legisladores e votações, a codificação dos votos como especificamos, e as variáveis extras sobre os legisladores.

Para analisar o comportamento da Câmara dos Deputados durante o primeiro do governo Bolsonaro, obtemos os dados de votação utilizando o pacote de R *congressbr* (McDonnell et al., 2019), que permite fazer buscas no API[4] da Câmera dos Deputados e Senado para coletar informações diversas sobre o processo legislativo do Congresso Nacional. Quando solicitamos os dados de um determinado ano, o objeto retornado está em formato tidy[5], onde a unidade de análise

4 A sigla API refere-se a *Application Programming Interface*, são interfaces com finalidade de simplificar processos diversos, entre eles o de coletar dados organizados de maneira mais acessível.

5 Para uma explicação sobre os princípios dessa organização de dados, ver (Wickham, 2014)

é o voto de um parlamentar e uma dada votação. O código abaixo ilustra como baixar os dados para deixá-los em formato adequado para a estimação de pontos ideais e convertê-los em um objeto de tipo *rollcall*:

```
# Instalar o pacote tidyverse, necessário para manipulação de dados
install.packages("tidyverse")

# Carregar o pacote tidyverse
library(tidyverse)

# Instalar o pacote congressbr, necessário para obter os dados
# da Câmara dos Deputados
install.packages("congressbr")

# Carregar o pacote congressbr
library(congressbr)

# Instalar o pacote oc
install.packages("oc")

# Carregar o pacote oc
library(oc)

# Baixar os votos de 2019 da Câmara dos Deputados
camara_votos <- cham_votes_year(2019)

# Transformar o camara_votos em formato wide
camara_wide <-
camara_votos %>%
  transmute(legislator_name = str_to_upper(legislator_name),
            legislator_party = str_to_upper(legislator_party),
            rollcall_id, legislator_vote) %>%
  mutate(legislator_vote = recode(legislator_vote,
                                  "Sim" = 1,
                                  "Obstrucao" = 4,
                                  "Art. 17" = 5,
                                  "Nao" = 6,
                                  "Abstencao" = 7)) %>%
  mutate(legislator_id = paste0(legislator_name,
                                  " - ",
                                   legislator_party)) %>%
  pivot_wider(names_from = rollcall_id,
              values_from = legislator_vote)

# Criar um objeto de tipo rollcall
camara_rollcall <- rollcall(camara_wide[,-c(1:3)],
                            yea = 1,
                            nay = 4:7,
                            missing = NA,
                            legis.names = camara_wide$legislator_id,
                            vote.names = colnames(camara_wide)[-c(1:3)],
                            legis.data = camara_wide[,1:3])
```

Primeiro precisamos instalar e carregar os pacotes necessários para manipular e baixar os dados. O *tidyverse* é um conjunto de pacotes de R de utilização bastante difundida para ciência de dados[6]. Em seguida baixamos e instalamos o *congressbr* propriamente dito. A função cham_votes_year() serve para baixarmos todas as votações em plenário de um determinado ano, 2019 no nosso caso.

Precisamos deixar o objeto *camara_votos* em um formato adequado para passar na função *rollcall*. Para isso, precisamos deixar o banco de dados em formato *wide*, onde cada linha é um deputado e cada coluna é uma votação ou informação sobre os deputados. Primeiro selecionamos apenas as colunas que queremos manter, relativas a nome, partido e voto dos parlamentares e a coluna que identifica a votação. Em seguida recodificamos os valores que constam na coluna de votos para valores numéricos. Um passo importante é criar uma nova coluna, que chamamos de *legislador_id*, composta pelo nome do deputado e partido. Assim, caso um parlamentar mude de partido, ele será identificado como outro legislador. Isso é importante pois, como estudos anteriores mostram (Izumi, 2016; McCarty et al., 2001), a mudança partidária e entrada ou saída de parlamentares na coalizão governista faz com que haja uma mudança em seu comportamento. Caso não realizemos esse passo, assumimos que a mudança de partido do legislador não impacta em seu comportamento parlamentar. Ao fazer com que cada ponto seja um deputado em um determinado partido, podemos verificar em que medida seu comportamento se altera com a mudança de sigla. Por fim, convertemos os dados para o formato *wide*, especificando que as novas colunas terão o nome das IDs de cada votação, e seus valores serão os votos dos legisladores.

Finalmente podemos criar o objeto em formato *rollcall*. Primeiro indicamos o objeto que possui os dados, selecionando apenas as colunas que correspondem a votações. Em seguida, estabelecemos quais valores correspondem a votos *sim*, *não*[7] *e missing*. Os últimos argumentos da função se referem a coluna dos nomes dos legisladores (que no caso é a combinação do nome com o partido), o nome das votações, e as informações extras sobre os deputados.

4. Rodando o *Optimal Classification*

Com os dados obtidos e organizados da maneira adequada, podemos utilizá-los para estimar os pontos ideais. O *Optimal Classification* é implementado pelo pacote *oc*. Para rodar a função oc()[8], é necessário especificar o objeto *rollcall* que possuem os dados a serem utilizados, o número de dimensões a serem estimadas, o número mínimo de votos que um legislador precisa ter para ser mantido na análise, a proporção (de 0 a 0.5) mínima que a votação tem que ter como minoria para ser mantida na análise, e a polaridade, que corresponde ao legislador que deve possuir valor positivo em cada uma das dimensões estimadas. Antes de estimar o modelo, cabe algumas explicações sobre esses elementos.

O número de dimensões estimadas, como mencionamos anteriormente, é um problema que todo pesquisador deveria ter em mente ao utilizar qualquer técnica de estimação de pontos ideais. Portanto, cabe ao pesquisador elaborar uma justificativa teórica ou estatística para sua escolha. Uma das formas propostas por Poole (2005) de decidir o número de dimensões, que será a utilizada neste trabalho, é observando o *skree plot*. Amplamente utilizado em estudos psicométricos para avaliar dimensionalidade, a ideia deste tipo de gráfico é obter

6 Para mais detalhes, ver (Wickham et al., 2019)

7 Para simplificar, designamos todos os votos diferentes de "sim" como votos contrários

8 As implementações em R do *W-NOMINATE* e α-NOMINATE também exigem estes mesmos elementos em suas funções

alguma medida de ajuste dos dados variando o número de dimensões estimadas, e observar o momento em que se forma um "cotovelo", fazendo com que a linha fique mais horizontal e menos vertical, indicando que a partir de dada dimensão não existe uma grande melhora ao se adicionar uma dimensão a mais. Este gráfico é gerado pelos pacotes NOMINATE com os valores calculados dos autovalores da matriz de concordância dos votos entre legisladores. Com o exemplo aplicado adiante, tal ideia ficará mais clara. No entanto, tal forma de determinar o número de dimensões não deve ser aplicada sem uma reflexão do sentido substantivo de cada uma delas, pois é frequente que uma melhora na medida de ajuste dos modelos não possua significado algum, sendo resultado apenas de incorporação de ruído aos dados (Armstrong et al., 2014).

O número mínimo de votos que os legisladores devem ter para permanecer na estimação, assim como a proporção mínima que a minoria deve possuir para que dada votação permaneça na estimação, são importantes pois podem influenciar, ainda que ligeiramente por vezes, o resultado final. Votações com minoria muito pequena são pouco úteis para informar a divisão entre seus votantes, e legisladores com poucos votos proporcionam uma confiança baixa sobre a consistência de suas preferências. Tais votações e votantes podem acabar influenciando as estimações por terem características de outliers quando comparados com o conjunto completo de dados. Portanto, é comum excluí-los, mas fica a critério do pesquisador quais critérios serão estabelecidos para tal. O padrão do *Optimum Classification*, caso tais valores não sejam especificados pelos pesquisadores, é que sejam mantidos votantes com no mínimo 20 votos e votações com uma minoria mínima de 2.5%.

A polaridade da votação diz respeito a especificar um votante que possui valor positivo em cada uma das dimensões. Esta parece ser uma escolha crucial para o pesquisador mas, na realidade, ela não afeta o resultado final de maneira significativa. Por exemplo, se acreditamos que a dimensão principal representa uma distribuição ideológica dos parlamentares, e quanto maior o valor mais conservador é o político, deveríamos escolher um votante conservador como referência de polaridade. Caso especifiquemos um progressista, o resultado será apenas um espelhamento do que esperaríamos, progressistas com valores positivos e conservadores com valores negativos. Dessa forma, a posição relativa entre os pontos não se altera, mas apenas o sinal negativo ou positivo de seus valores. Caso não tenhamos certeza ou confiança em quem escolher como polaridade, podemos estimar o *Optimal Classification* com qualquer legislador e verificar o resultado a posteriori.

O código abaixo mostra como rodamos o *Optimal Classification* para a Câmara dos Deputados em 2019, a partir dos dados que obtemos anteriormente:

```
# Roda o algoritmo do Optimal Classification
camara_oc <- oc(camara_rollcall,
                dims = 2,
                polarity = rep("EDUARDO BOLSONARO - PSL", 2))
```

Especificamos o objeto *rollcall* que construímos, pedimos para estimar duas dimensões, e informamos que nas duas dimensões o deputado Eduardo Bolsonaro deve ter valor positivo. Este processo demorou cerca de 20 segundos para rodar[9].

5. Desvendando a dimensionalidade

Podemos ter uma ideia sobre os dados observando o objeto que criamos:

```
# Sumário do objeto gerado pelo Optimal Classification
summary(camara_oc)
##
##
## SUMMARY OF OPTIMAL CLASSIFICATION OBJECT
## ---------------------------
##
## Number of Legislators:    616 (113 legislators deleted)
## Number of Votes:  259 (73 votes deleted)
## Number of Dimensions:     2
## Predicted Yeas:       47153 of 50123 (94.1%) predictions correct
## Predicted Nays:       46840 of 49446 (94.7%) predictions correct
##
## The first 10 legislator estimates are:
##                                coord1D coord2D
## EDIO LOPES - PL                  0.113  -0.360
## HAROLDO CATHEDRAL - PSD          0.143  -0.091
## HIRAN GONCALVES - PP             0.079  -0.380
## JHONATAN DE JESUS - PRB          0.249  -0.075
## JOENIA WAPICHANA - REDE         -0.529   0.703
## NICOLETTI - PSL                  0.283   0.009
## OTACI NASCIMENTO - SOLIDARIED    0.075  -0.126
## SHERIDAN - PSDB                  0.219  -0.025
## ACACIO FAVACHO - PROS            0.136  -0.265
## ALINE GURGEL - PRB              -0.067   0.085
```

Algumas informações úteis são exibidas ao pedirmos um sumário do objeto criado pelo *oc*. O número de legisladores e votação mantidas e excluídas, quantas dimensões foram estimadas, o número e porcentagem de classificações corretas nos votos favoráveis e contra as matérias propostas, e os valores estimados para os 10 primeiros legisladores do banco de dados. É possível obter a porcentagem geral de classificações corretas dos votos e a RPEA com o seguinte comando:

9 Esta é uma vantagem quando comparado com o tempo de estimação de modelos bayesianos, que podem demorar um tempo consideravelmente maior

```
# Obter as medidas de desempenho
camara_oc$fits
## [1] 0.9439986 0.7446653
```

O primeiro valor observado é a porcentagem de classificações correta, seguido da RPEA. Os valores obtidos são relativamente altos quando comparados com os valores apresentados por outros estudos (Leoni, 2002). Cabe observar que quando estimamos apenas uma dimensão, a porcentagem de classificações corretas cai para 92.8% e a RPEA para 0.67. A diferença não é muito grande, nos dando poucos elementos para crer que exista uma segunda dimensão relevante. Uma outra forma de se verificar a relevância de se adicionar mais dimensões é observar o *skree plot* com os autovalores gerados pelo *Optimal Classification*. A lógica desse tipo de gráfico, como mencionado anteriormente, é identificar em qual dimensão se forma um "cotovelo", e considerar apenas o número de dimensões anterior a este ponto. Podemos ver o gráfico com o seguinte código:

```
# Fazer o screeplot
plot.OCskree(camara_oc)
```

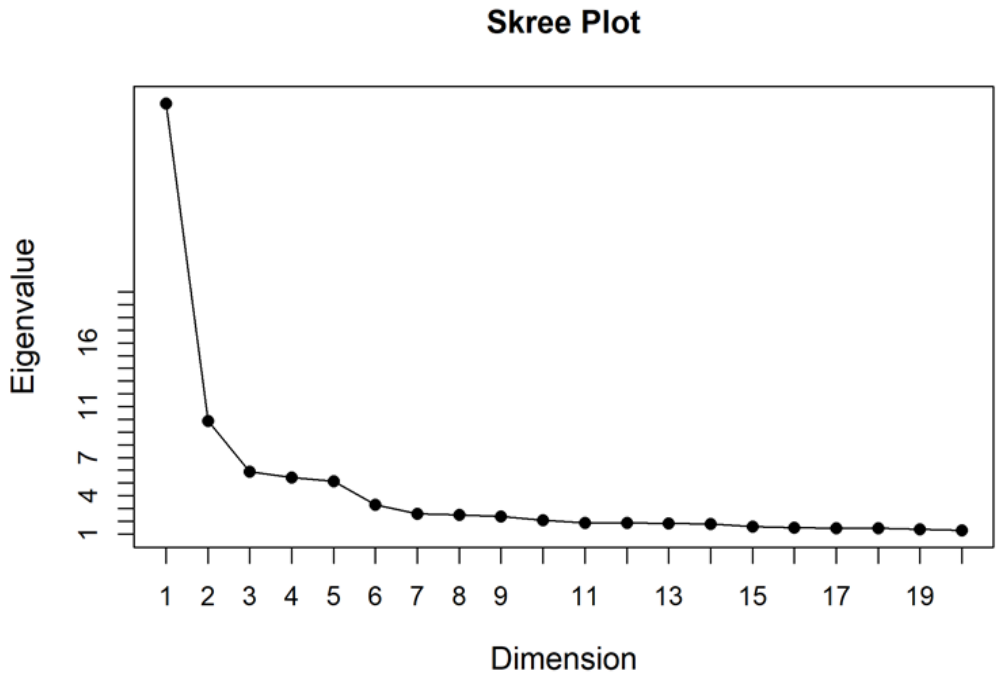


Gráfico 1: Screeplot para avaliação de dimensionalidade

Observamos que na estrutura dos dados predomina uma única dimensão, com uma segunda muito discreta. Esse resultado vai no mesmo sentido de estudos anteriores que identificaram uma única dimensão como relevante na estruturação dos votos da Câmara dos Deputados (Leoni, 2002) e do Senado (Izumi, 2016). Tais estudos também afirmam que essa dimensão diz respeito a divisão entre governo e oposição. Como mencionado anteriormente, a interpretação do significado das dimensões estimadas deve ser feita pelo pesquisador, estas não são dadas ou informadas pelas técnicas de estimação de pontos ideais.

Podemos verificar se em 2019, no primeiro ano de um governo que não estabeleceu um governo de coalizão nos moldes anteriores, essa dimensão permanece caracterizada pelo conflito entre governo e oposição. O gráfico abaixo mostra a correlação entre os pontos estimados para a primeira dimensão e a correlação do voto dos deputados com a posição que o líder de governo orientou[10].

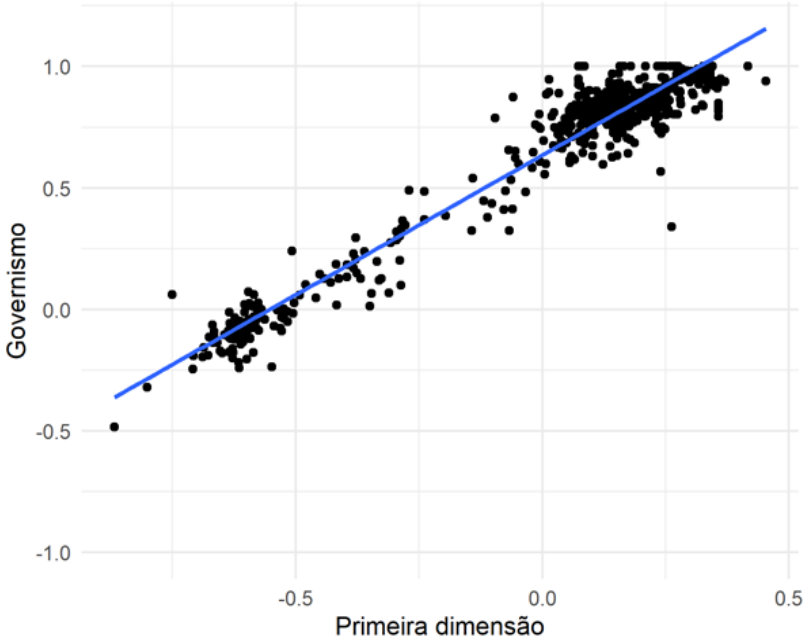


Gráfico 2: Correlação entre pontos ideais e governismo

O gráfico mostra uma alta correlação entre os pontos estimados da primeira dimensão e o voto dos deputados na mesma direção em que o líder do governo indicou. Essa associação é de 0.97. Ou seja, os deputados que possuem costumam ter o voto na mesma direção que o lider do governo estão localizados mais a direita no gráfico, com maiores valores em seus pontos ideais. Portanto, mesmo não havendo um presidencialismo de coalizão nos mesmos moldes que os governos anteriores, a dimensão do conflito político da Câmara dos Deputados se dá de maneira semelhante.

6. Distribuição dos pontos ideais

Os valores dos pontos ideais de cada legislador podem ser verificados no objeto que foi criado pelo *Optimal Classification*, em uma lista chamada *legislators*. Nela consta para cada votante a quantidade de votos corretamente classificados, as coordenadas estimadas e as variáveis sobre os legisladores que especificamos no argumento *legis.data* na função `rollcall()`.

10 Excluímos as votações nas quais o líder do governo liberou a bancada

```
glimpse(camara_oc$legislators)
## Rows: 729
## Columns: 11
## $ legislator_name  <chr> "EDIO LOPES", "HAROLDO CATHEDRAL", "HIRAN GONCALVE...
## $ legislator_party <chr> "PL", "PSD", "PP", "PRB", "REDE", "PSL", "SOLIDARI...
## $ legislator_id    <chr> "EDIO LOPES - PL", "HAROLDO CATHEDRAL - PSD", "HIR...
## $ rank             <dbl> 282.0, 333.0, 220.0, 511.0, 95.0, 548.0, 212.0, 46...
## $ correctYea       <dbl> 69, 106, 79, 44, 67, 104, 100, 105, 95, 33, 97, 80...
## $ wrongYea         <dbl> 1, 8, 4, 4, 6, 0, 5, 2, 2, 5, 0, 1, 4, 0, 5, 0, 0,...
## $ wrongNay         <dbl> 1, 3, 0, 1, 20, 1, 4, 5, 4, 10, 1, 4, 4, 4, 6, 0, ...
## $ correctNay       <dbl> 47, 87, 71, 34, 91, 90, 82, 77, 83, 44, 83, 125, 6...
## $ volume           <dbl> 0.311, 0.019, 0.068, 0.006, 0.003, 0.070, 0.020, 0...
## $ coord1D          <dbl> 0.11288447, 0.14326323, 0.07852262, 0.24926999, -0...
## $ coord2D          <dbl> -3.600102e-01, -9.135407e-02, -3.802511e-01, -7.45...
```

Com essas informações, podemos elaborar um gráfico para mostrar os pontos ideais dos parlamentares. O código a seguir mostra como fazer um gráfico aproveitando a variável partidária para analisar separadamente o comportamento dos partidos na primeira dimensão, utilizando a variável *coord1D*:

```
camara_oc$legislators %>%
  filter(legislator_party %in% c("PSL", "NOVO", "PT", "PSOL",
                                 "PCDOB", "PSB", "PDT", "PSD",
                                 "PSDB", "MDB", "PP", "DEM")) %>%
ggplot(aes(coord1D)) +
  geom_density() +
  facet_wrap(~legislator_party, drop = T, ncol = 3, scales = "free_y") +
  geom_vline(aes(xintercept = mean(camara_oc$legislators$coord1D,
                                   na.rm = T))) +
  theme_minimal() +
  labs(x = "Primeira dimensão", y = "Densidade") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

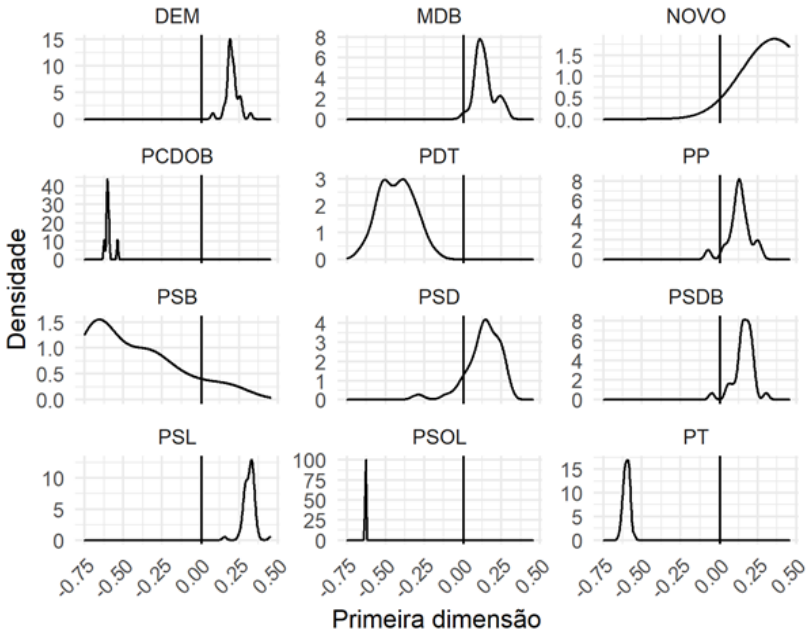


Gráfico 3: Distribuição partidária na primeira dimensão

Optamos por manter no gráfico apenas 12 partidos pois, caso representássemos as 33 diferentes agremiações e aqueles sem partido, o gráfico seria pouco útil para distinguir o comportamento partidário. Representamos também com uma linha vertical o valor médio dos pontos obtidos por todos os parlamentares. Dessa forma é possível verificar partidos com diferentes comportamentos de acordo com a dispersão de seus pontos ideais ao longo da primeira dimensão. PT, PSOL e PC do B concentrados na esquerda com pouca dispersão. PSB e PDT concentrados a esquerda da média, mas bastante dispersos. Os outros partidos, PSL, NOVO, DEM, MDB, PP, PSD e PSDB se concentram na direita.

Além da dispersão dos parlamentares em uma única dimensão, podemos visualizar a dispersão de seus pontos nas duas dimensões estimadas:

```
ggplot(camara_oc$legislators, aes(coord1D, coord2D)) +
  geom_point(alpha = 0.5) +
  theme_minimal() +
  labs(x = "Primeira dimensão", y = "Segunda dimensão")
```

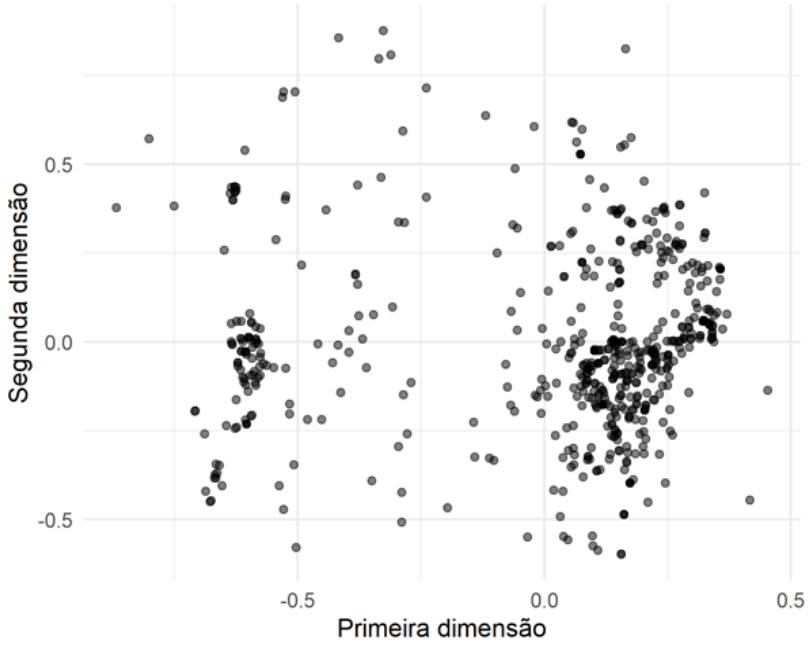


Gráfico 4: Distribuição dos pontos ideais nas duas dimensões

7. Distribuição das retas das votações

Mencionamos anteriormente que as estimações de pontos ideais estimam tanto os pontos dos legisladores quanto as retas das votações que os dividem. Para isso, usamos as informações que constam na lista *rollcalls* que compõe o objeto criado pelo *Optimal Classification*. Podemos verificar, em cada votação, quantos votos foram classificados de maneira correta e incorreta, a redução proporcional do erro, e as informações necessárias para determinarmos as retas que dividem os parlamentares.

```
head(camara_oc$rollcalls)
##      correctYea wrongYea wrongNay correctNay       PRE normVector1D
## [1,]         NA       NA       NA         NA        NA           NA
## [2,]         NA       NA       NA         NA        NA           NA
## [3,]        297       12        1         72 0.8452381    0.9985689
## [4,]         94        4       36        216 0.6923077    0.9254343
## [5,]        121       31       13        244 0.6716418    0.2525467
## [6,]        212       26       20        112 0.6666667    0.2703514
##      normVector2D   midpoints
## [1,]           NA          NA
## [2,]           NA          NA
## [3,]  -0.05347965  0.23146515
## [4,]  -0.37890820 -0.06372496
## [5,]  -0.96758471  0.08072122
## [6,]  -0.96276172  0.08239360
```

A análise de votações específicas, de forma qualitativa, também é uma forma importante para termos evidências do significado substantivo das dimensões estimadas. Não só podemos identificar as coalizões mais frequentes por meio dessa análise, como também podemos saber o ângulo dessas retas e determinar quais delas são predominantemente pertencentes a primeira dimensão, quais possuem influência de uma segunda dimensão e quais não são bem explicadas por nenhuma das dimensões. Para acrescentar as linhas das votações ao gráfico anterior, é necessário o seguinte código:

```
ggplot() +
  geom_point(data = camara_oc$legislators,
             aes(coord1D, coord2D)) +
  geom_segment(data = data.frame(camara_oc$rollcall),
               aes(x = (midpoints*normVector1D)+normVector2D,
                   y = (midpoints*normVector2D)-normVector1D,
                   xend = (midpoints*normVector1D)-normVector2D,
                   yend = (midpoints*normVector2D)+normVector1D),
               color = "grey", alpha = 0.4) +
  coord_cartesian(xlim = c(-1,1), ylim = c(-1,1)) +
  theme_minimal() +
  labs(x = "Primeira dimensão", y = "Segunda dimensão")
```

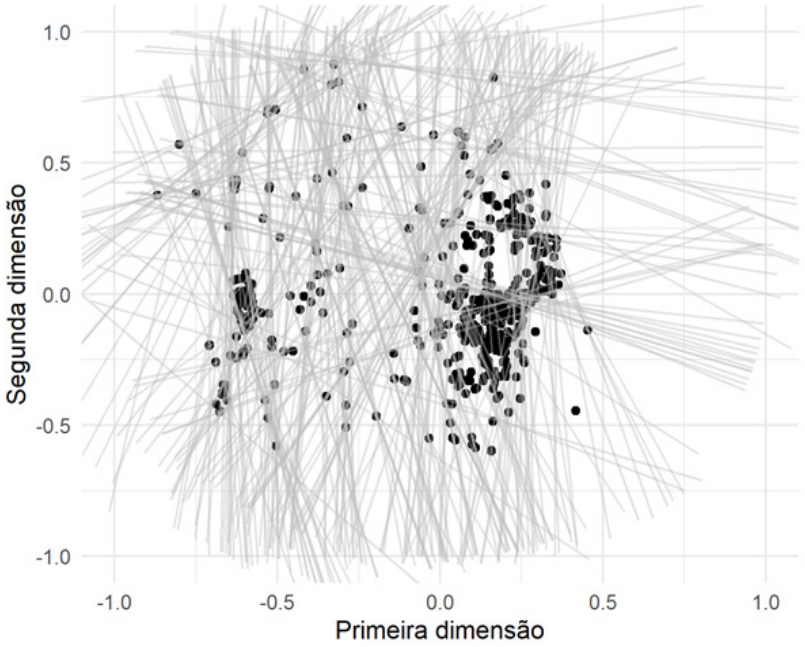


Gráfico 5: Distribuição dos pontos ideais nas duas dimensões e as linhas das votações

Observe que na função geom_point() especificamos os valores dos pontos ideais, que está na lista *legislators* do objeto criado pelo *Optimal Classification*. Na função geom_segment() especificamos os valores necessários para o cálculo dos pontos iniciais e finais das retas nas coordenadas *x* e *y*, utilizando as informações das votações que está na lista *rollcall* do mesmo objeto criado pelo *Optimal Classification*.

Com o gráfico, temos a impressão que a maioria das linhas possuem uma inclinação mais vertical, cruzando o eixo *x* do gráfico, indicando o predomínio da primeira dimensão na estruturação do comportamento dos votantes. No entanto, podemos analisar com um pouco mais de profundidade votações específicas para avaliar quais delas são bem explicadas pela estimação de pontos ideais. A primeira forma de fazer uma seleção das votações, é observar quais tiveram melhor redução proporcional do erro. Essa informação está disponível na lista *legislators* do objeto criado pelo *Optimal Classification*. Com o gráfico abaixo, podemos observar a distribuição desse valor em todas as votações:

```
ggplot(data.frame(camara_oc$rollcalls), aes(PRE)) +
  geom_histogram() +
  geom_vline(aes(xintercept = median(PRE, na.rm = T))) +
  theme_minimal() +
  labs(x = "RPE", y = "Número de votações")
```

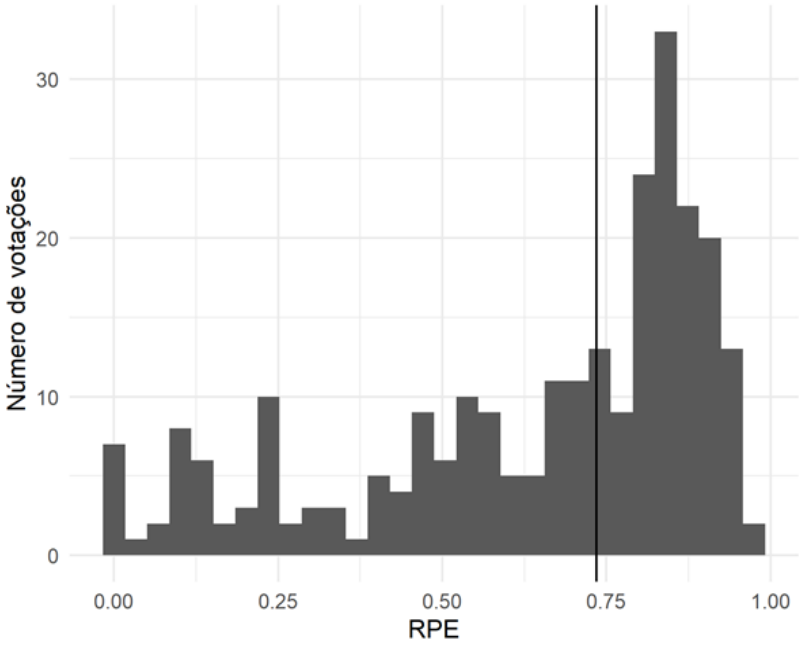


Gráfico 6: Distribuição da redução proporcional do erro

O valor mediano da redução proporcional do erro é próximo de 0.75. Portanto um primeiro critério para averiguar o sentido substantivo da primeira dimensão seria observar as votações com RPE maior do que 0.75. Essa informação sozinha não nos ajuda a saber quais votações dizem respeito a primeira ou segunda dimensão, ou ainda a ambas, mas sim quais votações são bem explicadas pela estimação como um todo. Uma forma de acessarmos quais delas correspondem a primeira dimensão é utilizamos as mesmas variáveis utilizadas para estimar as retas que dividem os parlamentares nas votações. Fazemos isso da seguinte forma:

```
x1 <- (camara_oc$rollcalls[,"midpoints"]*camara_oc$rollcalls[,"normVector1D"])+
  camara_oc$rollcalls[,"normVector2D"]

x2 <- (camara_oc$rollcalls[,"midpoints"]*camara_oc$rollcalls[,"normVector1D"])-
  camara_oc$rollcalls[,"normVector2D"]

x_diff <- abs(x2 - x1)

ggplot(NULL, aes(x_diff)) +
  geom_histogram() +
  geom_vline(aes(xintercept = median(x_diff, na.rm = T))) +
  theme_minimal() +
  labs(y = "Número de votações", x = "Medida de inclinação das retas")
```

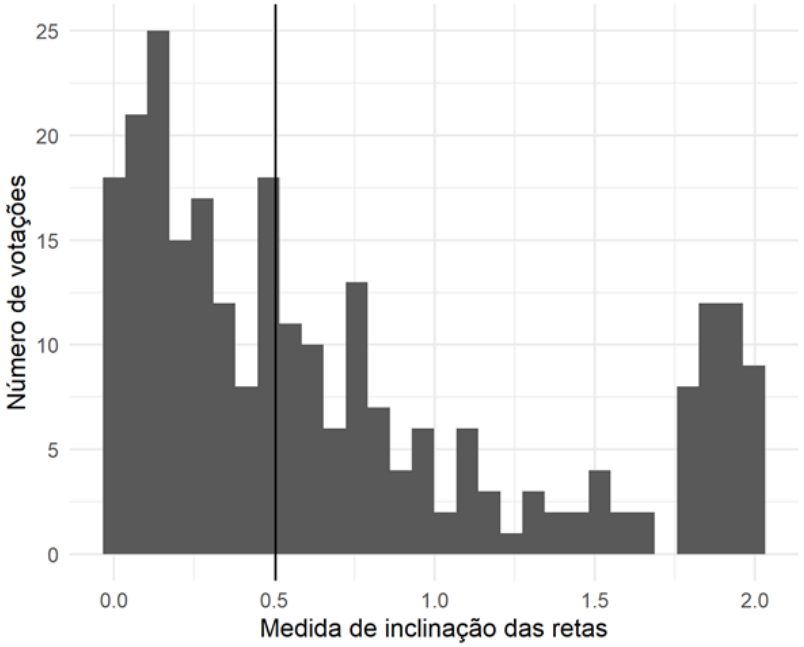


Gráfico 7: Medida de inclinação das retas

A ideia por trás do que fizemos foi: 1) calcular qual é o valor no eixo x, que corresponde a primeira dimensão, do ponto inicial da reta que divide os parlamentares; 2) fazer o mesmo para o ponto final; 3) saber o valor absoluto da diferença entre os dois valores. Se a diferença for baixa, quer dizer que a linha tende a ter uma posição mais vertical, sendo melhor explicada pela primeira dimensão. Com o gráfico, verificamos que a mediana dessa diferença é muito próxima de 0.5. Portanto valores inferiores a este seria um segundo indicador de predominância da primeira dimensão. Por fim, juntamos os dois critérios para selecionar as votações:

```
votacoes_selecionadas <- which(camara_oc$rollcalls[,"PRE"] > 0.75 &
                                 x_diff < 0.5)

colnames(camara_rollcall$votes)[votacoes_selecionadas]
##  [1] "PL-1292-1995-3"  "PL-1292-1995-7"  "PL-4742-2001-1"  "PLP-441-2017-7"
##  [5] "PLP-441-2017-9"  "PLP-441-2017-11" "MPV-852-2018-1"  "MPV-866-2018-1"
##  [9] "MPV-867-2018-3"  "MPV-867-2018-19" "MPV-871-2019-3"  "MPV-871-2019-4"
## [13] "MPV-871-2019-5"  "PEC-6-2019-2"    "PEC-6-2019-3"    "PEC-6-2019-4"
## [17] "PEC-6-2019-5"    "PEC-6-2019-6"    "PEC-6-2019-7"    "PEC-6-2019-8"
## [21] "PEC-6-2019-9"    "PEC-6-2019-13"   "PEC-6-2019-14"   "PEC-6-2019-18"
## [25] "PEC-6-2019-22"   "PEC-6-2019-24"   "PEC-6-2019-31"   "PEC-6-2019-32"
## [29] "PEC-6-2019-33"   "PEC-6-2019-35"   "PEC-6-2019-37"   "PEC-6-2019-38"
## [33] "PEC-6-2019-39"   "PEC-6-2019-40"   "PEC-6-2019-41"   "PEC-6-2019-42"
## [37] "PEC-34-2019-4"   "MPV-881-2019-1"  "MPV-881-2019-5"  "MPV-881-2019-6"
## [41] "MPV-881-2019-7"  "MPV-881-2019-8"  "MPV-881-2019-9"  "MPV-881-2019-11"
## [45] "PL-2787-2019-1"  "PL-2788-2019-2"  "REQ-1464-2019-1" "PL-2999-2019-6"
## [49] "PL-3261-2019-5"  "PL-3261-2019-6"  "PL-3261-2019-7"  "PL-3261-2019-9"
## [53] "PL-3261-2019-10" "PL-3261-2019-12" "PL-3261-2019-13" "MPV-886-2019-1"
## [57] "MPV-886-2019-2"  "PL-3715-2019-2"  "PL-3715-2019-9"  "MPV-890-2019-3"
## [61] "MPV-890-2019-5"  "MPV-890-2019-6"  "MPV-890-2019-7"  "MPV-890-2019-9"
## [65] "PL-4162-2019-1"  "PL-4162-2019-2"  "REQ-2110-2019-1" "REQ-2157-2019-1"
## [69] "REQ-2234-2019-1" "PDL-523-2019-1"  "PDL-523-2019-4"  "PL-5815-2019-1"
## [73] "PL-5815-2019-2"  "PL-5815-2019-3"
```

Nos códigos acima, verificamos quais votações atendem os dois critérios que estabelecemos e visualizamos os nomes destas votações. Podemos observar que diversas votações que dizem respeito a PEC 6/2019, sobre a reforma da previdência, correspondem a primeira dimensão. Podemos checar se essas linhas realmente se comportam da maneira esperada, ou seja, se são verticais com pouca inclinação:

```
ggplot() +
  geom_point(data = camara_oc$legislators,
             aes(coord1D, coord2D)) +
  geom_segment(data = data.frame(camara_oc$rollcall[votacoes_selecionadas,]),
               aes(x = (midpoints*normVector1D)+normVector2D,
                   y = (midpoints*normVector2D)-normVector1D,
                   xend = (midpoints*normVector1D)-normVector2D,
                   yend = (midpoints*normVector2D)+normVector1D),
               color = "grey", alpha = 0.5) +
  theme_minimal() +
  labs(x = "Primeira dimensão", y = "Segunda dimensão")
```

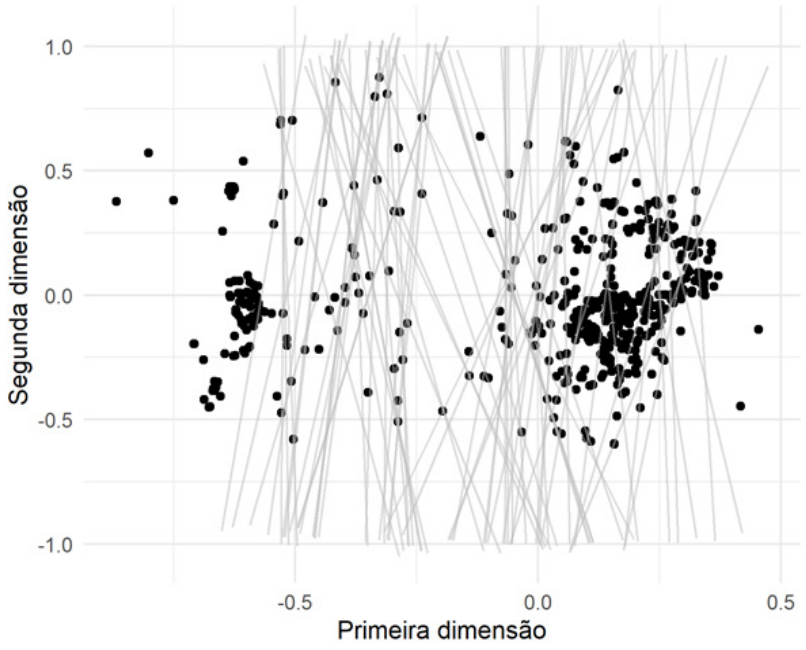


Gráfico 8: Pontos ideais e retas das votações correspondentes a primeira dimensão

Note que a diferença está na especificação da função *geom_segment()*, onde indicamos no objeto *camara_oc$rollcall* que queremos apenas as linhas que correspondem as votações selecionadas. Nosso objetivo foi bem sucedido em selecionar votações bem explicadas pela técnica e que diz respeito a primeira dimensão. Podemos notar que existem votações que acabam dividindo tanto o bloco de parlamentares localizados a esquerda quanto o bloco de deputados localizados a direita.

Com as informações obtidas sobre os pontos ideais de cada votante e as retas de cada votação, é possível analisar mais qualitativamente decisões específicas. Dessa forma é possível se aprofundar em como se dá a estrutura das preferências e do conflito político entre os deputados, contribuindo ainda mais para interpretação do significado substantivo das dimensões estimadas e dos resultados de forma geral. Abaixo mostramos um exemplo com a representação de uma das votações da PEC da reforma da previdência.

```
camara_oc$legislators %>%
  mutate(Voto = camara_rollcall$votes[,147]) %>%
  mutate(Voto = case_when(Voto == 1 ~ "Sim",
                          Voto != 1 ~ "Não")) %>%
  filter(!is.na(Voto)) %>%
ggplot() +
  geom_point(aes(coord1D, coord2D)) +
  geom_segment(data = data.frame(camara_oc$rollcall) %>%
                 slice(147),
               aes(x = (midpoints*normVector1D)+normVector2D,
                   y = (midpoints*normVector2D)-normVector1D,
                   xend = (midpoints*normVector1D)-normVector2D,
                   yend = (midpoints*normVector2D)+normVector1D),
               color = "grey", alpha = 0.5) +
  theme_minimal() +
  facet_wrap(vars(Voto), nrow = 1) +
  labs(x = "Primeira dimensão", y = "Segunda dimensão")
```

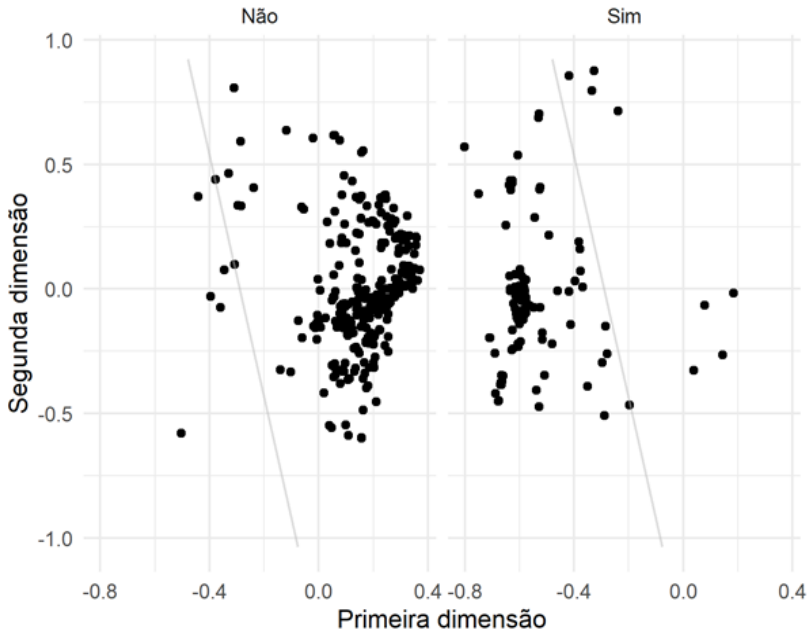


Gráfico 9: Pontos ideais e reta da votação de uma das votações da PEC 6/2019

No código acima, utilizamos o objeto que criamos com o *Optimal Classification* e adicionamos a coluna referente aos votos da coluna 147, que se refere a votação selecionada sobre a reforma da previdência. Em seguida recodificamos os valores para indicar se o voto foi favorável ou contrário a matéria, removendo aqueles que não votaram, e geramos o gráfico. Especificamos que sejam feitos dois painéis separados, de acordo com o voto, e mantemos apenas a reta da votação selecionada. Com o gráfico, podemos visualizar que os governistas, que neste caso votaram contra a proposta, se concentram ao lado direito, com maiores valores de ponto ideal na primeira dimensão, enquanto a oposição votou a favor da matéria. Podemos observar que a estimação erra 5 deputados que votaram *não* e 8 deputados que votaram *sim*, por estarem localizados do lado oposto da reta ao que seria esperado. Essa votação tem 97.1% dos votos corretamente classificados e uma RPE de 0.88. Com este gráfico é possível analisar qualitativamente esta votação específica e votantes específicos. Um exemplo é o com-

portamento da deputada Tabata Amaral. No contexto da reforma da previdência, houve muita crítica, tanto nas redes sociais quanto internamente no PDT, sobre alguns do votos que a deputada fez a favor do texto que foi aprovado. O ponto ideal dela é o que votou *não* localizado mais acima do gráfico. Como observamos, o voto é corretamente classificado pela estimação, colocando ao lado direito da reta que divide a votação. Dessa forma, o voto dela é coerente com seu comportamento em outras votações, fazendo-a ocupar este ponto no espaço. Neste caso, mesmo localizada mais próxima a esquerda, seu voto acompanhando aqueles localizados a direita não seria atípico dada a forma como o conflito político se estrutura na Câmara dos Deputados ao analisarmos espacialmente todos os deputados e votações.

Outra possibilidade de análise do comportamento dos votantes é observar quais deles são mais "imprevisíveis", com maior número de classificações incorretas pois seu comportamento não seguiria a lógica das dimensões encontradas, mas por outras não estimadas. Existem modelos específicos para lidar com tais tipos de votantes de forma mais adequada (Lauderdale, 2010). No objeto criado pelo *Optimal Classification*, são apresentadas colunas com o resultado de quantos *sim* e *não* foram classificados de maneira correta e incorreta. Dessa forma, é possível saber a porcentagem de votos de cada um dos votantes que a estimação de pontos ideais conseguiu alocar de maneira adequada no espaço. Podemos obter os resultados e elaborar um gráfico da seguinte forma:

```
camara_oc$legislators %>%
  mutate(classificacoes_corretas = 100 * (correctYea + correctNay) /
           (correctYea + wrongYea + wrongNay + correctNay)) %>%
  ggplot(aes(classificacoes_corretas)) +
  geom_histogram() +
  geom_vline(aes(xintercept = mean(classificacoes_corretas, na.rm = T))) +
  theme_minimal() +
  labs(x = "Porcentagem de classificações corretas", y = "Quantidade")
```

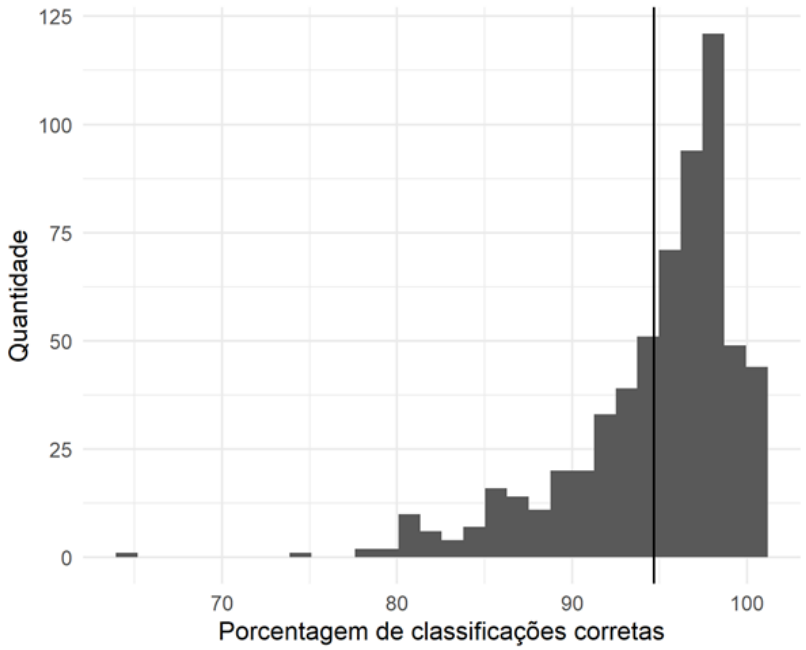


Gráfico 10: Distribuição da porcentagem de classificações corretas dos votantes

Com o histograma, observamos que a média de classificações corretas dos votantes é de 94.6% e a grande maioria dos deputados está acima dos 90%. Para observar quais são os votantes cujos comportamentos são menos explicados pelas duas dimensões estimadas, utilizamos o seguinte código:

```
camara_oc$legislators %>%
  mutate(classificacoes_corretas = 100 * (correctYea + correctNay) /
           (correctYea + wrongYea + wrongNay + correctNay)) %>%
  arrange(classificacoes_corretas) %>%
  slice(1:20) %>%
  mutate(legislator_id = fct_reorder(legislator_id, classificacoes_corretas)) %>%
  ggplot(aes(classificacoes_corretas, legislator_id)) +
  geom_point() +
  theme_minimal() +
  labs(x = "Classificações corretas (%)", y = "Votante")
```

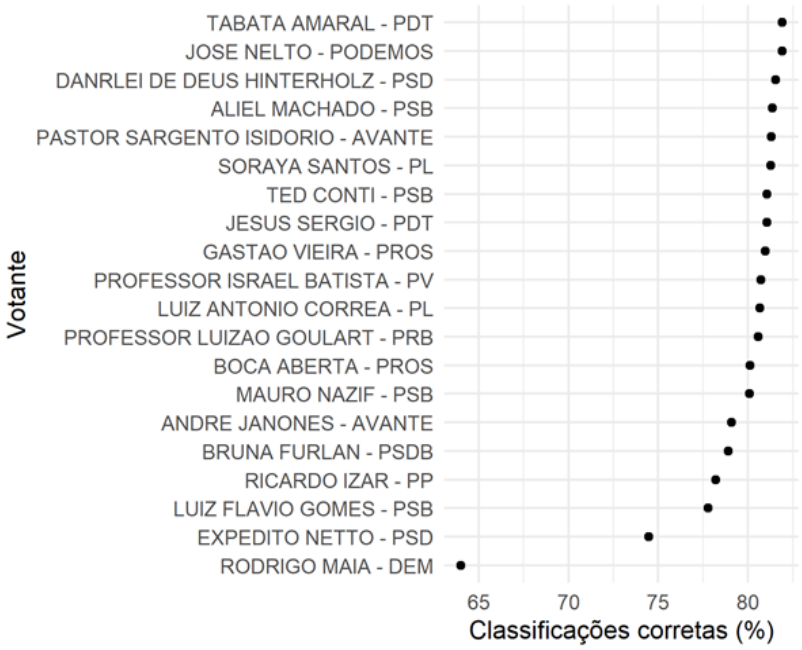


Gráfico 11: Votantes com menor porcentagem de classificações corretas

Mantemos no gráfico apenas os 20 votantes com menor porcentagem de votos corretamente classificados. Observamos que Rodrigo Maia é o que possui menor porcentagem, com menos de 65%, seguido por Expedito Netto com pouco menos que 75%. Curioso notar que a deputada Tabata Amaral, mencionada anteriormente pelas críticas ao seu comportamento, está entre aqueles cujo comportamento não é tão bem explicado pelas duas dimensões estimadas.

8. Considerações finais

O presente trabalho apresentou como estimar pontos ideais e algumas formas de se explorar os resultados de uma das diversas aplicações que a técnica pode ter. Muitos outros pacotes de R foram implementados para explicar o comportamento de atores políticos diversos com tipos de dados variados.

Em estudos que trabalham com textos, ao invés de votos em propostas legislativas, a escolha dos atores estão no uso de palavras. Ao tratarem sobre um determinado tópico, usar certas palavras e omitir outras revelaria uma preferência ideológica sobre como abordar uma questão. Ainda mais, a própria prevalência de certos tópicos em detrimento de outros em um discurso político revelaria um conteúdo ideológico. Implementações como *Wordfish* (Slapin e Proksch, 2008), *Wordscores* (Laver et al., 2003) e *Wordshoal* (Lauderdale e Clark, 2014) procuram lidar com esse tipo de questão, analisando programas partidários, discursos parlamentares, entre outros.

A relativa facilidade de se obter dados do Twitter também tem possibilitado a utilização desse tipo de dado para estimação de pontos ideais. Não apenas o texto dos tweets feitos pelos usuários podem ser analisados com as implementações que acabamos de mencionar, como também a maneira como os usuários interagem e se conectam uns com os outros. O ato de escolher seguir determinadas contas em detrimento de outras relaria uma preferência ideológica dos usuários da rede social em obter certos tipos de conteúdos (Barberá, 2015).

Diversos pacotes de R que surgem a cada dia, além de expandir quais tipos de dados podem ser utilizados para a estimação de pontos ideais, também procuram expandir as possibilidades de uma melhor inferência estatística a partir de dados já utilizados a tempos. A comparação intertemporal de uma mesma instituição, por exemplo, não é trivial. Não podemos estimar pontos ideais, por exemplo, da Câmara dos Deputados em 2019 e comparar diretamente os resultados com uma estimação relativa ao ano anterior. Para isso são necessários pressupostos teóricos e estatísticos que são implementados por outros algorítimos (Bailey et al., 2017; Martin e Quinn, 2002; Poole e Rosenthal, 2001). Além da comparação intertemporal, é possível fazer uma comparação de pontos ideais entre instituições e atores distintos, com os dados e métodos adequados (Bailey, 2007; Bonica, 2014; Tausanovitch e Warshaw, 2013).

REFERÊNCIAS BIBLIOGRÁFICAS

ARMSTRONG, David A et al. (2014), *Analyzing spatial models of choice and judgment with R*. [S.l.] CRC Press.

BAILEY, Michael A. (2007), "Comparable preference estimates across time and institutions for the court, congress, and presidency". *American Journal of Political Science*, vol. 51,no. 3: 433–448.

BAILEY, Michael A & STREZHNEV, Anton & VOETEN, Erik. (2017), "Estimating dynamic state preferences from United Nations voting data". *Journal of Conflict Resolution*, vol. 61,no. 2: 430–456.

BARBERÁ, Pablo. (2015), "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data". *Political analysis*, vol. 23,no. 1: 76–91.

BONICA, Adam. (2014), "Mapping the ideological marketplace". *American Journal of Political Science*, vol. 58,no. 2: 367–386.

CARROLL, Royce et al. (2009), "Comparing NOMINATE and IDEAL: Points of difference and Monte Carlo tests". *Legislative Studies Quarterly*, vol. 34,no. 4: 555–591.

CARROLL, Royce et al. (2013), "The structure of utility in spatial models of voting". *American Journal of Political Science*, vol. 57,no. 4: 1008–1028.

CLINTON, Joshua & JACKMAN, Simon & RIVERS, Douglas. (2004), "The statistical analysis of roll call data". *American Political Science Review*, vol. 98,no. 2: 355–370.

CONVERSE, Philip E. (1964), "The nature of belief systems in mass publics. Ideology and discontent". *Ideology and Discontent*, 206–261.

DESPOSATO, Scott W. & INGRAM, Matthew C. & LANNES JR., Osmar P. (2015), "Power, Composition, and Decision Making: The Behavioral Consequences of Institutional Reform on Brazil's Supremo Tribunal Federal". *Journal of Law, Economics, and Organization*, vol. 31,no. 3: 534–567.

DOWNS, Anthony. (1957), *An economic theory of democracy*. New York, Harper; Row.

DUCK-MAYR, JBrandon & MONTGOMERY, Jacob M. *Ends Against the Middle: Scaling Votes When Ideological Opposites*. [S.d.].

FERREIRA, Pedro Fernando Almeida Nery & MUELLER, Bernardo. (2014), "How judges think in the Brazilian Supreme Court: Estimating ideal points and identifying dimensions". *Economia*, vol. 15,no. 3: 275–293.

IZUMI, Mauricio. (2016), "Governo e Oposição no Senado Brasileiro (1989-2010)". *Dados*, vol. 59,no. 1: 91–138.

LAUDERDALE, Benjamin. (2010), "Unpredictable voters in ideal point estimation". *Political Analysis*, no. 18: 151–171.

LAUDERDALE, Benjamin E. & CLARK, Tom S. (2014), "Scaling politically meaningful dimensions using texts and votes". *American Journal of Political Science*, vol. 58,no. 3: 754–771.

LAVER, Michael & BENOIT, Kenneth & GARRY, John. (2003), "Extracting policy positions from political texts using words as data". *American Political Science Review*, vol. 97,no. 02: 311–331.

LEONI, Eduardo. (2002), "Ideologia, Democracia e Comportamento Parlamentar: A Camara dos Deputados (1991-1998)". *Dados*, vol. 45,no. 3: 361–386.

LIMONGI, Fernando. (2019), "Presidencialismo do Desleixo". *Piauí*, no. 158: Disponível em: <https://piaui.folha.uol.com.br/materia/presidencialismo-do-desleixo/>.

MARIANO SILVA, Jeferson. (2018), "Mapeando o Supremo. As posições dos ministros do STF na jurisdição constitucional (2012-2017)". *Novos Estudos - CEBRAP*, vol. 37,no. 1: 35–54.

MARTIN, Andrew D. & QUINN, Kevin M. (2002), "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis*, vol. 10,no. 2: 134–153.

MCCARTY, Nolan & POOLE, Keith T. & ROSENTHAL, Howard. (2001), "The Hunt for PartyDiscipline in Congress". *American Political Science Review*, vol. 95,no. 3: 673–87.

MCDONNELL, Robert Myles & DUARTE, Guilherme Jardim & FREIRE, Danilo. (2019), "congressbr: An R Package for Analyzing Data from Brazil's Chamber of Deputies and Federal Senate". *Latin American Research Review*, vol. 54,no. 4:

POOLE, Keith et al. (2011), "Scaling Roll Call Votes with wnominate in R". *Journal of Statistical Software*, vol. 42,no. 14: 1–21.

POOLE, Keith T. (2005), *Spatial models of parliamentary voting*. [S.l.] Cambridge University Press.

POOLE, Keith T & ROSENTHAL, Howard. (1985), "A spatial model for legislative roll call analysis". *American Journal of Political Science*, 357–384.

POOLE, Keith T & ROSENTHAL, Howard. (1997), *Congress: A political-economic history of roll call voting*. [S.l.] Oxford University Press on Demand.

POOLE, Keith T & ROSENTHAL, Howard. (2001), "D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting". *Legislative Studies Quarterly*, 5–29.

POOLE, Keith T & ROSENTHAL, Howard. (1991), "Patterns of congressional voting". *American Journal of Political Science*, 228–278.

ROSENTHAL, Howard & VOETEN, Erik. (2004), "Analyzing roll calls with perfect spatial voting: France 1946–1958". *American Journal of Political Science*, vol. 48,no. 3: 620–632.

SLAPIN, Jonathan B & PROKSCH, Sven-Oliver. (2008), "A scaling model for estimating time-series party positions from texts". *American Journal of Political Science*, vol. 52,no. 3: 705–722.

SPIRLING, Arthur & MCLEAN, Iain. (2007), "UK OC OK? Interpreting optimal classification scores for the UK House of Commons". *Political Analysis*, 85–96.

TAUSANOVITCH, Chris & WARSHAW, Christopher. (2013), "Measuring constituent policy preferences in congress, state legislatures, and cities". *The Journal of Politics*, vol. 75,no. 2: 330–342.

WICKHAM, Hadley. (2014), "Tidy data". *Journal of Statistical Software*, vol. 59,no. 10: 1–23.

WICKHAM, Hadley et al. (2019), "Welcome to the Tidyverse". *Journal of Open Source Software*, vol. 4,no. 43: 1686.

ZUCCO, Cesar. (2009), "Ideology or What? Legislative Behavior in Multiparty Presidential Settings". *The Journal of Politics*, vol. 71,no. 3: 1076–1092.

ZUCCO, Cesar & LAUDERDALE, Benjamin E. (2011), "Distinguishing Between Influences on Brazilian Legislative Behavior". *Legislative Studies Quarterly*, vol. 3,no. 36: 363–396.

Mapping behavior with ideal point estimation

Rodrigo Martins - Universidade Federal de Pernambuco

Ideal point estimation is a methodology of growing importance in Political Science. Developed in the mid-20th century, its link with the spatial theory of voting has been fruitful for several fields of study, with great expectations on its future development. The goal of this paper is to present what ideal point estimation is and didactically illustrate how to apply one of the various available techniques, using data on parliamentary behavior from the Chamber of Deputies in 2019.

1. Introduction

What structures divisions among parliamentarians, ideology or government support? Does change in party affect the behavior of legislators? What is the level of internal cohesion of political parties and how does it oscillate over time? Has polarization in the National Congress increased? Which political parties have the most unpredictable behavior? What kinds of bills tend to polarize political behavior? How does a president's ability to establish a coalition government influence parliamentary behavior?

These are some of the questions raised by the legislative studies literature to explain parliamentary behavior. In common among them is the need to access preferences, ideologies, and behaviors of elected representatives. However, several fields in Political Science have the same interest regarding other types of political actors and political behavior. Is it possible to find evidence of ideological behavior in the judiciary branch? Do politicians' speeches express ideological differences? Do the electorate's ideological preferences reflect those of their representatives? Does the option to follow certain people and not others on Twitter reflect the political preference of the platform's users?

These and several other questions are raised by studies that use ideal point estimation to explain the behavior of political actors. The spatial theory of voting, theorized by Downs (1957), began being applied to the most diverse areas, not only to explain traditional politicians, but also to shed light on the preferences of those outside political institutions. The ease in translating theoretical formalizations of preference formation into mathematical and statistical formulas made the spatial theory of voting find success in various manners of implementation.

The goal of this paper is to paper, in a simple and intuitive way, what is ideal point estimation and to illustrate how to use it didactically. Despite being used for varied purposes, the focus here will be on the field of studies that most developed the technique, the area of legislative studies. I will illustrate how to estimate ideal points using data from the Chamber Deputies in 2019, the first year of Jair Bolsonaro's presidential term. Previous studies that used ideal point estimation for the Brazilian case show that the main dimension that structures parliamentary behavior in the Senate and Chamber of Deputies is not ideological, but the government-opposition dimension (Izumi, 2016; Zucco, 2009), although at some moments these dimensions seem to be overlayed (Leoni, 2002). Faced with the analysis that Bolsonaro altered the relationship between the executive and the legislative by not forming a parliamentary support base via a coalition, replacing "coalition presidentialism with negligence" (Limongi, 2019), would there be a change in the Chamber of Deputies' behavior? Would there be another dimension structuring nominal votes in the plenary?

2. What are ideal points?

Ideal point estimation is derived from the spatial model of voting, initially made popular in Political Science by Downs (1975), in which political actors may be represented as a point in a space and choose between political alternatives. These also represented spatially, which are closer together. In addition to being a theoretical principle, several studies empirically demonstrate that the majority of political behavior can be represented by a unidimensional space, frequently assumed to be an ideological scale. The division between left and right then is an expression of how political actors, be they voters or representatives, behave and manifest their preferences in ways

that are aligned on distinct subjects, such as economic and social policies. This is similar to Converse's (1964) argument regarding a belief system that permeates preferences on varied issues in a unified worldview.

The ideal point estimation technique shows us a map of how voters are distributed spatially, representing the distance among the preferences of each individual under analysis. Simply put, what we extract is the proximity among voters when we consider all individuals, and not just their similarities two-by-two. In this way, it is even possible to estimate the proximity among voters that in practice did not participate together in any vote since, by estimating their distance in relation to all other voters, we would also know the distance among them. One of the model's assumptions is that when knowing how an individual decides on a specific theme, it would be possible to predict their behavior in relation to other decisions, especially on similar themes, by deriving from the previous decision. Thus, when correlating all decisions, it would be possible to estimate individuals' coordinates, so that those who decide in similar ways are spatially closer and those who diverge are spatially distant.

To illustrate the idea underlying ideal point estimation, we can turn to a hypothetical example. Let's suppose there is a board of three legislators that needs to make a decision on an economic issue. Considering that one of the legislators votes for more state intervention in the economy and the other two who vote for less intervention, the legislators' spatial representation will be done in such a way that the two who voted together occupy the same side, in opposition to the voter who remained isolated. It is also possible to illustrate the issue being voted as a line that represents the division among legislators on two different sides. Figure 1 is the way to spatially represent the division in this example. The red and green circle are the two who voted together, the blue circle the individual who voted alone and the purple line is the issue being voted on which represents the split between the two sides in question.

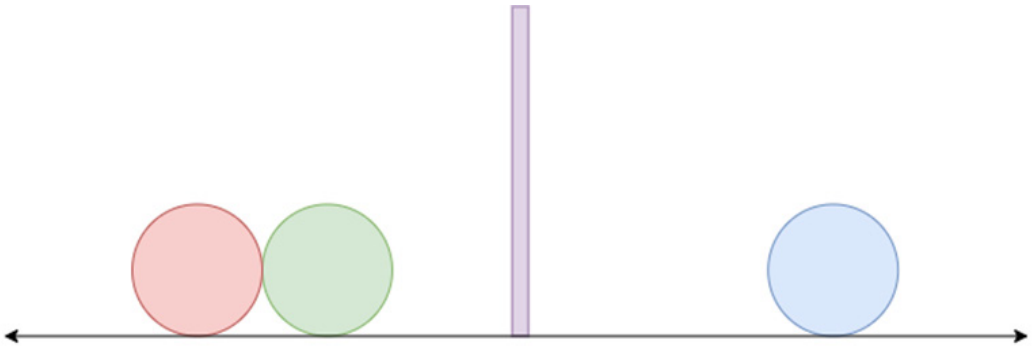


Figure 1: Example of spatial analysis with 1 vote and 3 voters

Let's suppose now that on a second vote on the same issue, the legislator represented by green does not vote again with the voter represented by red, but with the one represented by blue. In this second moment, the green individual must be placed between the others equidistantly, since they voted together with each of them. At the same time, those who did not vote together on either vote must remain at extreme opposition. The line that represents the first vote, in purple, must keep the division among voters as it did previously. And now a second vote line, which will be represented in white in figure 2, must isolate the legislator who voted alone and keep together those who voted together.

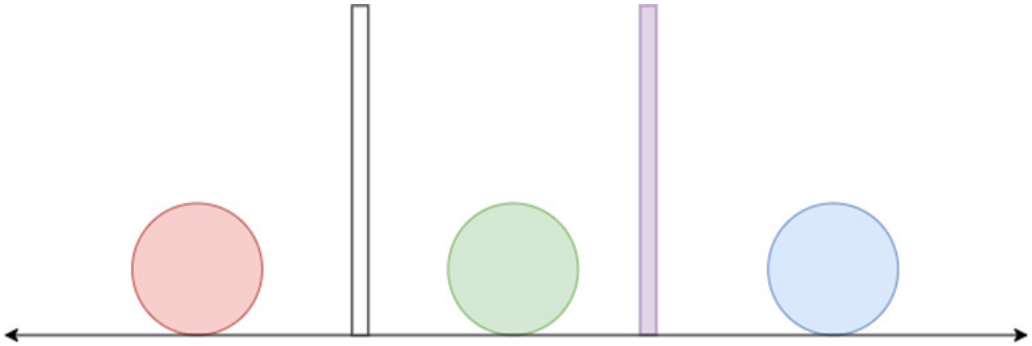


Figure 2: Example of spatial analysis with 2 votes and 3 voters

As a last example, let's imagine that on a third vote on matter unrelated to the economy, the legislators who before were opposed now vote together, leaving the voter who was once in the center, isolated in the minority. So that the spatial representation stays appropriate in this situation, it is best to add a vertical axis where the one who was defeated is distanced from the other two, while the victors are kept close to each other in this second axis. We can also represent this third vote as a line, represented in yellow in figure 3, this time horizontally.

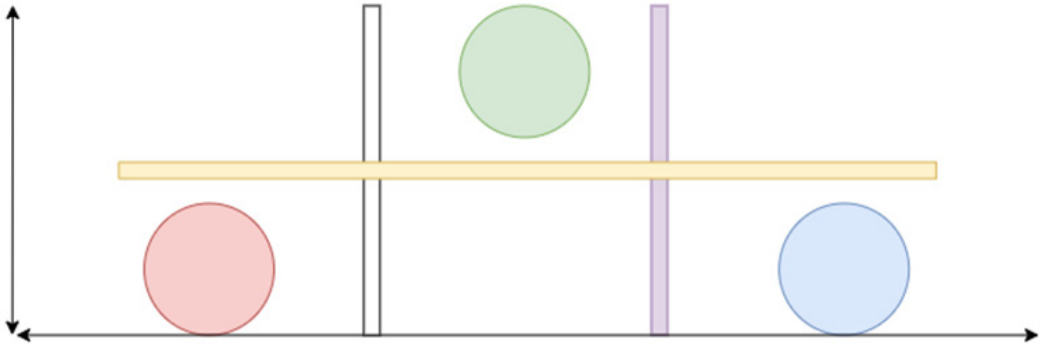


Figure 3: Example of spatial analysis with 3 votes and 3 voters

This type of model projects the preferences manifested through voting in a space (unidimensional, bidimensional, and even beyond), evaluating similarities and divergences in the decisions of each voter. As it would not be practical to conduct this manual process for many voters over many votes, several techniques were developed, with differing statistical procedures, with the goal of estimating ideal points both for voters and votes.

However, one must wonder that at any point it will no longer possible to perfectly represent the spatial distribution of voters and votes using only two dimensions. With an increase in voters and votes, it is inevitable that voters' points end up allocated in a position that does not correspond with their vote on a give issue. Poole (2005) proposes two measures to evaluate and compare the performance of ideal point models: the percentage of correct classifications and the aggregate proportional reduction in error (APRE). To better explain what these measures are and the difference between them, I will use a hypothetical example as illustration. Figure 4 represents a vote with 11 voters identified by a letter and a color.

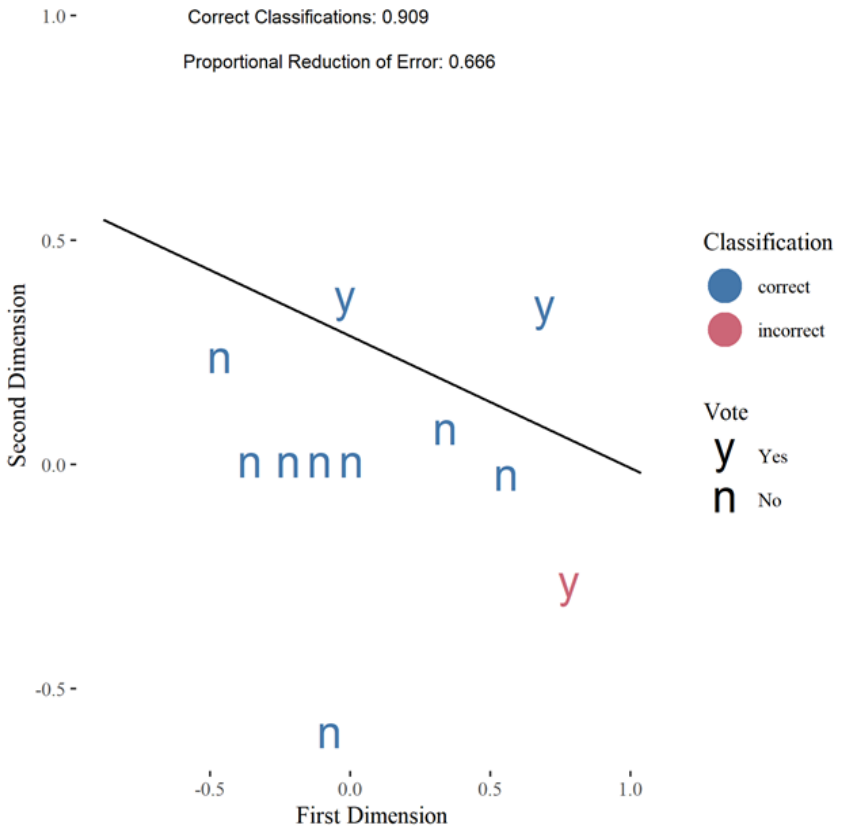


Figure 4: Example of evaluation of performance measures

Voters represented by the letter "y" voted "yes" and voters represented by the letter "n" voted "no". The only voter represented in red was the voter who was wrongly classified by the model, given that they voted "yes" when all the others represented in the same side of the vote voted "no". They should have been on the other side of the line, along with the others who voted "yes" and were correctly classified. Thus, the percentage of correct classification of this vote is 0.90922[1]. The proportional reduction in error (PRE) is 0.66[2]. This measure is more adequate and stricter in evaluating the performance of the models. Using the example of Armstrong et al. (2014), if at any vote there is a 65-35 split in favor of a proposal, we could correctly classify 65% of the votes, just by "naively" saying that all voters will support it. The proportional reduction in error measures how much the model improves the classification of votes in relation to the minority. Going back to our hypothetical example in figure 4, the PRE is 0.66 because it correctly classifies two out of three minority voters. The APRE is a measure that aggregates the PRE of each vote.

It is important to highlight that this method uses split votes, that is, votes that were not unanimous. This criterion is important, given that once one is looking to identify how divergence happens among voters, the issue at hand must necessarily produce dissent among them. Unanimous votes cannot give information on groupings and divergences and are thus disregarded. This characteristic may affect studies on institutions with a high rate of consensus. For the Federal Supreme Court, for example, where most decisions are monocratic or unanimous, the use of this method would lead to loss of explanatory capacity of the court as a whole. However, by using it only on majoritarian decisions, we would observe cases that deal with controversial issues, subject to dispute of distinct interpretations and polit-

1 number of correct classifications / number of voters

2 (votes in the minority – incorrect classifications) / votes in the minority

ical conflict. It is more likely that important cases, that have larger repercussion, fit into this scenario.

Care is needed when analyzing the dispersion of voters in space, because the model does not have the capacity to infer what determines that dispersion, that is, what is the factor that structures voting in spatial dimensions. This type of statistical estimation has been used in Brazilian Political Science, for example, when analyzing the Chamber of Deputies (Leoni, 2002; Zucco, 2009; Zucco and Lauderdale, 2011), Senate (Izumi, 2016), and Federal Supreme Court (Desposato et al., 2015; Ferreira and Mueller, 2014; Mariano Silva, 2018). For the cases of the Chamber of Deputies and the Senate, it is not ideological division that ends up being highlighted with ideal point estimation, but the division between opposition and government.

The first statistical ideal point estimation model elaborated in Political Science directly derived from the spatial theory of voting was NOMINATE (Poole and Rosenthal, 1985, 1991, 1997). This is a parametric model applied to binary votes and widely used in the area of legislative studies. There are several subsequent developments implemented in R, the most popular of which being W-NOMINATE (Poole et al., 2011), while D-NOMINATE and DW-NOMINATE are used to estimate dynamic ideal points, to observe the oscillation of voters' behaviors over time.

With the increase in computational performance, new implementations of ideal point estimations were elaborated with the application of Bayesian statistics. Derived from item-response theory, well developed in the field of psychometrics, and developed from the works of Clinton et al. (2004) and Martin and Quinn (2002), the implementation in R of IDEAL in the **pscl** package is one of the most popular, along with the MCMCpack package, which relies on an implementation of the Bayesian dynamic model, among others. Item-response theory is used mainly in educational studies, where they seek to measure the knowledge of individuals through tests, replaced in Political Science by measuring ideology from votes.

With distinct implementations, the NOMINATE family and Bayesian models have slightly different assumptions. While both assume that errors are independent, identically distributed, with normal distribution, and lower probability of occurring as the voter's ideal point is distant the vote line, the way each voter treats the available options when voting *yes* or *no* differs. In NOMINATE, if both alternatives are distant from the voter's ideal point, they are indifferent to the options offered, given that the preference curve takes on a Gaussian shape. In Bayesian models, the voter's preference curve takes on a quadratic shape, allowing them to maintain sensibility when facing options distant from their ideal point. In practice, despite this assumption not having much effect on the models' results (Carroll et al., 2009), to deal with this issue α-NOMINATE was created (Carroll et al., 2013). This Bayesian implementation of NOMINATE was made to replace previous implementations and it estimates if the best functional form of the voters' preferences is Gaussian, quadratic, or an intermediary mix of both.

In this paper, given the various statistical implementations of the spatial model to conduct ideal point estimation, I will demonstrate how to use *Optimal Classification* (OC). This is a non-parametric ideal point estimation technique, unlike the previous models. Its algorithm consists of the following steps (Poole, 2005, p. 82):

1. To generate initial values for the voters' ideal points from the decomposition into eigenvalues and eigenvectors from the concordance matrix between voter pairs.

2. Given the voters' values, to find the estimations of the votes' vectors that maximize the number of correct classifications.

3. Given the values of the votes' vectors, to find new values of the voters' ideal points that maximize the number of correct classifications.

4. Return to step 2.

This process is repeated until the number of correct classifications is stabilized, maximizing their number. According to Poole, this technique is consistent with data with few voters and few votes and one of its advantages is the fact that it has less assumptions when compared to parametric models of ideal point estimation, making it less susceptible to statistical estimation problems.

I chose this technique because, being non-parametric, it requires that less statistical assumptions on the voters' behaviors be satisfied. Rosenthal and Voeten (2004) argue that the strategic behavior that legislatures may present due to party discipline would be less problematic for *Optimal Classification* when compares to other parametric models[3].

In the following section, I will begin the explanation of how to implement an ideal point estimation, presenting the code in R for the required commands and functions to use this method.

3. First step: create a rollcall object

The packages that implement *Optimal Classification*, *W-NOMINATE*, α*-NOMINATE,* and *IDEAL* require that a *rollcall* type object be created in R, so that we can use its respective functions. To do so, it is necessary to specify each of the following elements in the rollcall() function:

1. the data where each line is a voter, and each column is a roll call vote

2. the values corresponding to a "yes" vote

3. the values corresponding to a "no" vote

4. the values corresponding to *missing*, when a voter did not vote for some reason

5. the values corresponding to when a voter was not part of the legislature

6. the values corresponding to the legislators' names, in the same order that they appear in the lines of the database

7. the values corresponding to the roll call votes' names, in the same order that they appear in the columns of the database

8. the values corresponding to additional information on the voters, such as their parties, in the same order that they appear in the lines of the database

9. the values corresponding to additional information on the roll call votes, such as their dates, in the same order that they appear in the columns of the database

The first three types of information are mandatory, while the others are optional. The code below illustrates the shape of a hypothetical database and how to create a *rollcall* object from it:

3 This problem is not fully solved by this technique (Spirling and McLean, 2007) since the monotonicity assumption still applies. See Duck-Mayr and Montgomery ([S.d.]) for more on this issue.

```
# Observe the database
hypotethical_data[1:5,1:13]
##          name      state      party V1 V2 V3 V4 V5 V6 V7 V8 V9 V10
## 1 Legislator1     Alaska   Democrat  1  1  6  6  1  1  6  1  1   1
## 2 Legislator2      Texas   Democrat  1  1  6  1  1  6  6  6  1   1
## 3 Legislator3       Ohio Republican  1  1  6  6  1  6  6  6  6   6
## 4 Legislator4      Texas   Democrat  1  1  1  1  1  1  1  1  1   6
## 5 Legislator5 California   Democrat  6  6  6  1  6  6  6  6  6   1
# Create a rollcall object from the hypothetical_database object
hypotethical_rollcall<- rollcall(hypotethical_data[,-c(1:3)],
                                 yea = 1,
                                 nay = 6,
                                 legis.names = hypotethical_data$name,
                                 legis.data = hypotethical_data[,2:3])

hypotethical_rollcall
## Number of Legislators:    20
## Number of Votes:  100
##
## Using the following codes to represent roll call votes:
## Yea:      1
## Nay:      6
## Abstentions:  NA
## Not In Legislature:   9
##
## Legislator-specific variables:
## [1] "state" "party"
## Detailed information is available via the summary function.
```

First, we observe the database's first five lines and 13 columns to get an idea of its structure and organization. The first column represents the name of each legislator, the second their state of origin, and the third, their party. Following these, the columns represent votes, where 1 means a favorable vote and 6 means a contrary vote on the issue. To create a *rollcall* object, the first argument specified is the database, delimiting the voting columns only, and therefore excluding columns 1 through 3. Next the *yea* and *nay* arguments are specified, respectively 1 and 6. Then we specify the two optional arguments, *legis.names*, indicating the name column in the *hypotethical_data* object as the legislators' names, and columns 2 and 3 in the *legis.data* argument, given that these columns have additional information on the legislators. If we wanted to name the votes and include additional information on each one of them, we would specify the arguments *vote.names* and *vote.data* with their values. When observing the object created, *hypotethical_rollcall*, we can verify the structure of the data, indicating the number of legislators and votes, the votes' codifications as specified, and the additional variables on the legislators.

To analyze the Chamber of Deputies' behavior during the first year of the Bolsonaro government, we obtained voting data using the R package *congressbr* (McDonnell et al., 2019), which allows searches in the Chamber of Deputies and Senate's API[4] to collect various information on the National Congress's legislative process. When we request data from a given year, the object returned is in tidy format[5] where the unit of analysis is a parliamentarian's vote and a given roll call vote. The code below illustrates how to download the data to put them in the proper format for ideal point estimation and convert them to a *rollcall* object:

4 The API acronym means *Application Programming Interface*, interfaces that simplify several processes; among them, collection of data in a more accessible manner.

5 For an explanation on the principles of this data organization, see (Wickham, 2014)

```
# Install the tidyverse package, required for data manipulation
install.packages("tidyverse")

# Load the tidyverse package
library(tidyverse)

# Install the congressbr package, required to access the data from the #Chamber of
Deputies
install.packages("congressbr")

# Load the congressbr package
library(congressbr)

# Install the oc package
install.packages("oc")

# Load the oc package
library(oc)

# Download the Chamber of Deputies' votes from 2019
chamber_votes <- cham_votes_year(2019)

# Transform chamber_votes to wide format
chamber_wide <-
chamber_votes %>%
  transmute(legislator_name = str_to_upper(legislator_name),
            legislator_party = str_to_upper(legislator_party),
            rollcall_id, legislator_vote) %>%
  mutate(legislator_vote = recode(legislator_vote,
                                  "Sim" = 1,
                                  "Obstrucao" = 4,
                                  "Art. 17" = 5,
                                  "Nao" = 6,
                                  "Abstencao" = 7)) %>%
  mutate(legislator_id = paste0(legislator_name,
                                " - ",
                                 legislator_party)) %>%
  pivot_wider(names_from = rollcall_id,
              values_from = legislator_vote)

# Create a rollcall object
chamber_rollcall <- rollcall(chamber_wide[,-c(1:3)],
                             yea = 1,
                             nay = 4:7,
                             missing = NA,
                             legis.names = chamber_wide$legislator_id,
                             vote.names = colnames(chamber_wide)[-c(1:3)],
                             legis.data = chamber_wide[,1:3])
```

First, we need to install and load the packages required to manipulate and download the data. *Tidyverse* is a set of R packages with wide usage in data science[6]. Next, we download and install *congressbr* itself. The cham_votes_year() function downloads all the plenary votes in a given year, 2019 in this case.

We need to put the *chamber_votes* object in the adequate format to use the *rollcall* function. To do so, we need the database in *wide* format, where every line is a deputy, and every column is a roll call vote or information on the deputies. First, we select only the columns we want to keep, regarding parliamentarians' name, party, and vote and the column that identifies the roll call vote. Then, we recode the values in the votes column to numerical values. An important step is to create a new column, which I called *legislador_id*, comprised of the deputy's name and party. Then, if a parliamentarian changes party, they will be identified as another legislator. This is important as previous studies show (Izumi, 2016; McCarty et al., 2001) that party change and the entry or exit of parliamentarians from the governing coalition causes change in their behavior. If we do not take this step, we assume that a legislator's change in party has no impact in their parliamentary behavior. By making each point a legislator in a certain party, we can verify how much their behavior changes with the move. Lastly, we convert the data to the *wide* format, specifying that new columns will have the name of each roll call votes' ID, and its values will be the legislators' votes.

Finally, we can create the object in *rollcall* format. First, we indicate the object that has the data, selecting only the columns that correspond to roll call votes. Then we establish which values correspond to *yes*, *no*[7], and *missing*. The last arguments in the function refer to the column of legislators' names (which, in this case, is the combination of name and party), the name of the votes, and the additional information on the deputies.

4. Running Optimal Classification

With the data obtained and organized in the adequate manner, we case use them to estimate the ideal points. *Optimal Classification* is implemented by the *oc* package. To run the oc()[8] function it is necessary to specify the *rollcall* object that has the data to be used, the number of dimensions to be estimated, the minimum number of votes a legislator needs to be kept in the analysis, the minimum proportion (0 to 0.5) that the vote must have as minority to be kept in the analysis, and the polarity, which corresponds to the legislator that must have positive value in each of the estimated dimensions. Before estimating the model, some explanations about these elements are required.

The number of estimated dimensions, as mentioned previously, is a problem every researcher must keep in mind when using any ideal point estimation technique. Therefore, it is up to the researcher to elaborate a theoretical or statistical justification for their choice. One of the forms proposed by Poole (2005) of deciding the number of dimensions, which will be used in this paper, is by observing the *scree plot*. Widely used in psychometric studies to evaluate dimensionality, the idea behind this graph is to obtain some adjustment measure of the data by varying the number of dimensions estimated and observing the moment in which an "elbow" is formed, making the line more horizontal and less vertical, indicating that from a given dimension there is no great improvement in adding another dimension. This graph

6 For more details, see (Wickham et al., 2019)

7 To make it simple, we assigned all votes different from "yes" as contrary votes.

8 *W-NOMINATE*'s and α-NOMINATE's R implementations also require these same elements in their functions.

is generated by the NOMINATE packages with values calculated from the eigenvalues of the concordance matrix of votes among legislators. With the example used ahead, this will become clearer. However, this way of eliminating the number of dimensions should not be applied without reflecting on the substantive meaning of each of them, because it is frequent that an improvement in the measure of the models' fit has no meaning at all, being the result of a mere incorporation of noise to the data (Armstrong et al., 2014).

The minimum number of votes that legislators must have to remain in the estimation, as the minimum proportion that the minority must have so that a given vote remains in the estimation, are important because they can influence, albeit slightly at times, the final result. Votes with a very small minority are of little use to inform the division among its voters, and legislators with few votes bring low confidence to the consistency of their preferences. These votes and voters may end up influencing the estimations due to their outlier characteristics when compared to the data set. Therefore, it is common to exclude them, but it is up to the researcher which criteria will be established for that. The standard for *Optimum Classification*, if these values are not specified by researchers, is that voters with a minimum of 20 votes and votes with minimum minority of 2.5% be kept.

The polarity of the vote concerns specifying a voter who has a positive value in each dimension. This may seem like a crucial choice for the researcher but, in reality, it does not affect the final result in a significant way. For example, if we believed that the main dimension represents an ideological distribution of parliamentarians, and the higher the value the more conservative is the politician, then we should choose a conservative voter as polarity reference. If we specified a progressive, the result would only mirror our expectation, progressives with positive values and conservatives with negative values. In this way, the relative position among the points is not altered, but only the negative or positive sign of their values. If we are not certain or confident in who to choose as polarity, we can estimate the *Optimal Classification* with any legislator and verify the result a posteriori.

The code below shows how I ran the *Optimal Classification* for the Chamber of Deputies in 2019, from the data obtained previously:

```
# Running Optimal Classification's algorithm
chamber_oc <- oc(chamber_rollcall,
                 dims = 2,
                 polarity = rep("EDUARDO BOLSONARO - PSL", 2))
```

I specified the *rollcall* object built, asked to estimate two dimensions, and informed that in both dimensions deputy Eduardo Bolsonaro must have positive value. This process took roughly 20 seconds to run[9].

Unveiling dimensionality

We can have an idea on about the data by observing the object created:

9 This is an advantage when compared to the time for estimation of Bayesian models, which can take considerably longer.

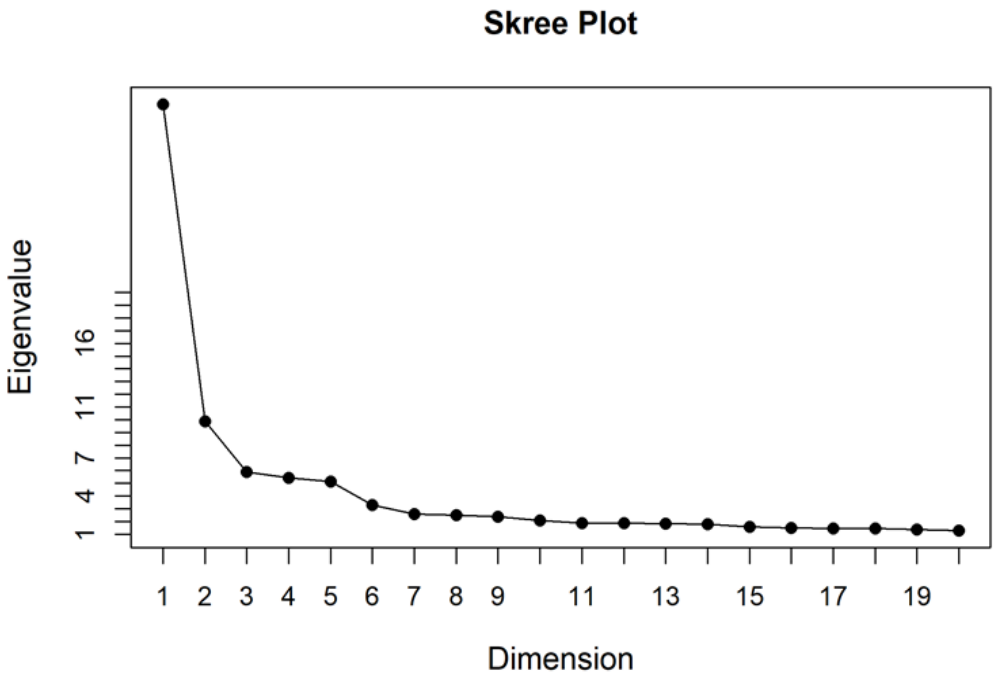
```
# Summary of the object generated by Optimal Classification
summary(chamber_oc)
##
##
## SUMMARY OF OPTIMAL CLASSIFICATION OBJECT
## --------------------------
##
## Number of Legislators:    616 (113 legislators deleted)
## Number of Votes:  259 (73 votes deleted)
## Number of Dimensions:     2
## Predicted Yeas:       47153 of 50123 (94.1%) predictions correct
## Predicted Nays:       46840 of 49446 (94.7%) predictions correct
##
## The first 10 legislator estimates are:
##                                coord1D coord2D
## EDIO LOPES - PL                  0.113  -0.360
## HAROLDO CATHEDRAL - PSD          0.143  -0.091
## HIRAN GONCALVES - PP             0.079  -0.380
## JHONATAN DE JESUS - PRB          0.249  -0.075
## JOENIA WAPICHANA - REDE         -0.529   0.703
## NICOLETTI - PSL                  0.283   0.009
## OTACI NASCIMENTO - SOLIDARIED    0.075  -0.126
## SHERIDAN - PSDB                  0.219  -0.025
## ACACIO FAVACHO - PROS            0.136  -0.265
## ALINE GURGEL - PRB              -0.067   0.085
```

Some useful information is displayed when we request a summary of the object created by *oc*. The number of legislators and roll call votes kept and excluded, how many dimensions were estimated, the number and percentage of correct classifications of favorable and contrary votes on the issues proposed, and the values estimated for the first 10 legislators in the database. It is possible to obtain the overall percentage of correct classifications of the votes and the APRE with the following command:

```
# Obtain performance measures
chamber_oc$fits
## [1] 0.9439986 0.7446653
```

The first value observed is the percentage of correct classifications, followed by the APRE. The values obtained are relatively high when compared to the values presented by other studies (Leoni, 2002). It is worth noting that when we estimate only one dimension, the percentage of correct classifications falls to 92.8% and the APRE to 0.67. The difference is small, giving us few elements to believe that there is a relevant second dimension. Another way to check the relevance of adding more dimensions is to observe the *scree plot*, with the eigenvalues generated by *Optimal Classification*. The logic behind this type of graph, as mentioned earlier, is to identify at which dimension an "elbow" is formed and to consider only the number of dimensions prior to that point. We can see the graph with the following code:

```
# Make scree plot
plot.OCskree(chamber_oc)
```

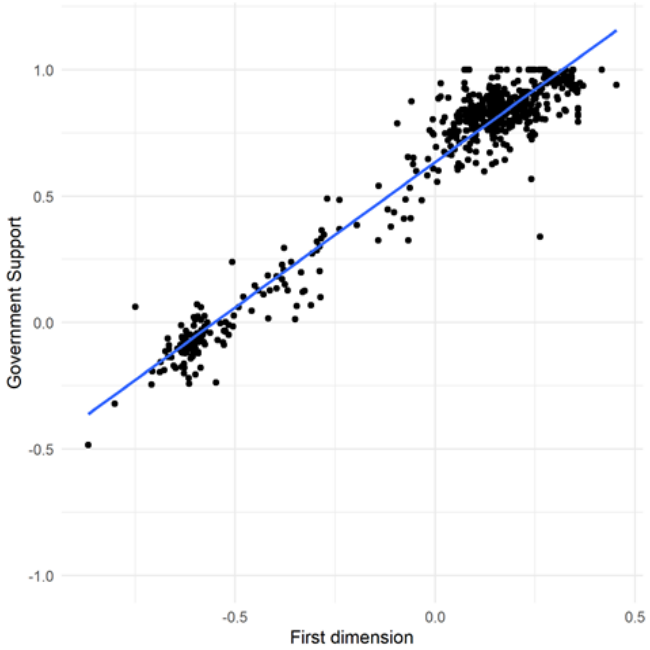


Graph 1: Screeplot for dimensionality evaluation

We can see that a single dimension is predominant in the data structure, with a second one being very discreet. This result follows previous studies that identified a single dimension as relevant in the structuring of votes in the Chamber of Deputies (Leoni, 2002) and Senate (Izumi, 2016). These studies also state that this dimension concerns the division between government and opposition. As mentioned previously, the interpreting the meaning of the estimated dimensions must be done by the researcher, as these are not given or informed by the ideal point estimation techniques.

We can check if, in 2019, in the first year of a government that did not establish a coalition government similar to ones prior, this dimension remains characterized by the government and opposition conflict. The graph below shows the correlation between estimated points for the first dimension and correlation of deputies' votes with the position the leader of the government laid out[10].

10 I excluded the votes for which the leader of the government allowed the deputies to vote freely.



Graph 2: Correlation between ideal points and government support

The graph shows a high correlation between the points estimated in the first dimension and the deputies' vote in the same direction stated by the leader of the government. This association is of 0.97. That is, deputies who usually vote in the same way as the leader of the government are further to the right of the graph, with higher values at their ideal points. Therefore, even though there is no coalition presidentialism in the same manner of prior governments, the dimension of political conflict in the Chamber of Deputies occurs in a similar way.

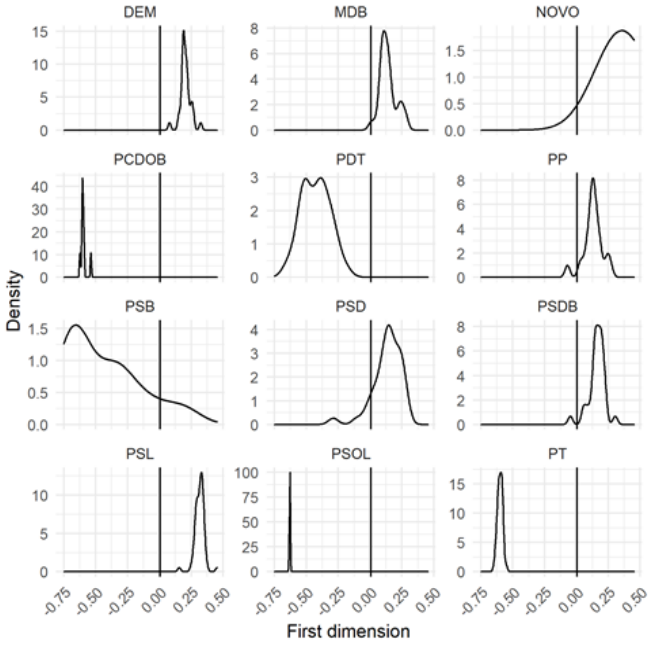
5. Distribution of ideal points

The values of each legislators' ideal points can be verified in the object created by *Optimal Classification*, in a list named *legislators*. In it can be found, for each voter, the amount of votes correctly classified, the estimated coordinates, and the variables on the legislators that were specified in the *legis.data* argument in the rollcall() function

```
glimpse(chamber_oc$legislators)
## Rows: 729
## Columns: 11
## $ legislator_name  <chr> "EDIO LOPES", "HAROLDO CATHEDRAL", "HIRAN GONCALVE...
## $ legislator_party <chr> "PL", "PSD", "PP", "PRB", "REDE", "PSL", "SOLIDARI...
## $ legislator_id    <chr> "EDIO LOPES - PL", "HAROLDO CATHEDRAL - PSD", "HIR...
## $ rank             <dbl> 282.0, 333.0, 220.0, 511.0, 95.0, 548.0, 212.0, 46...
## $ correctYea       <dbl> 69, 106, 79, 44, 67, 104, 100, 105, 95, 33, 97, 80...
## $ wrongYea         <dbl> 1, 8, 4, 4, 6, 0, 5, 2, 2, 5, 0, 1, 4, 0, 5, 0, 0,...
## $ wrongNay         <dbl> 1, 3, 0, 1, 20, 1, 4, 5, 4, 10, 1, 4, 4, 4, 6, 0, ...
## $ correctNay       <dbl> 47, 87, 71, 34, 91, 90, 82, 77, 83, 44, 83, 125, 6...
## $ volume           <dbl> 0.311, 0.019, 0.068, 0.006, 0.003, 0.070, 0.020, 0...
## $ coord1D          <dbl> 0.11288447, 0.14326323, 0.07852262, 0.24926999, -0...
## $ coord2D          <dbl> -3.600102e-01, -9.135407e-02, -3.802511e-01, -7.45...
```

With this information, we can elaborate a graph to show the parliamentarians' ideal points. The following code shows how to make a graph using the party variable to separately analyze the behavior of parties in the first dimension, using the *coord1D* variable:

```
chamber_oc$legislators %>%
  filter(legislator_party %in% c("PSL", "NOVO", "PT", "PSOL",
                                 "PCDOB", "PSB", "PDT", "PSD",
                                 "PSDB", "MDB", "PP", "DEM")) %>%
ggplot(aes(coord1D)) +
  geom_density() +
  facet_wrap(~legislator_party, drop = T, ncol = 3, scales = "free_y") +
  geom_vline(aes(xintercept = mean(chamber_oc$legislators$coord1D,
                                   na.rm = T))) +
  theme_minimal() +
  labs(x = "First dimension", y = "Density") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

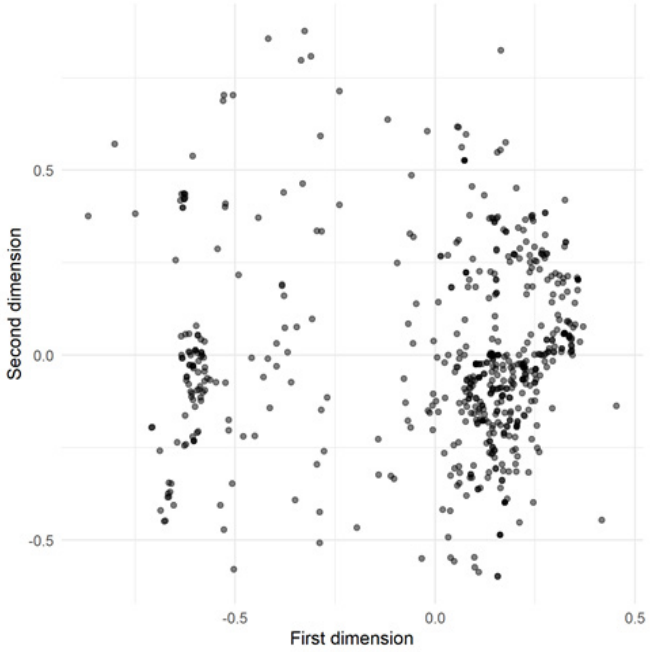


Graph 3: Party distribution in the first dimension

I opted to keep in the graph only 12 parties, given that if all 33 different parties and those without parties were added, the graph would be of little use to distinguish party behavior. I have also represented with a vertical line the mean value of points obtained of all parliamentarians. That way, it is possible to verify parties with different behaviors according to the dispersion of their ideal points over the first dimension. PT, PSOL, and PC do B appear concentrated in the left with little dispersion. PSB and PDT appear concentrated to the left of the mean, but very dispersed. Other parties, PSL, NOVO, DEM, MDB, PP, PSD, and PSDB are concentrated on the right.

In addition to the parliamentarians' dispersion in a single dimension, we can visualize the dispersion of their points on the two estimated dimensions:

```
ggplot(chamber_oc$legislators, aes(coord1D, coord2D)) +
  geom_point(alpha = 0.5) +
  theme_minimal() +
  labs(x = "First dimension", y = "Second dimension")
```



Graph 4: Distribution of ideal points on both dimensions

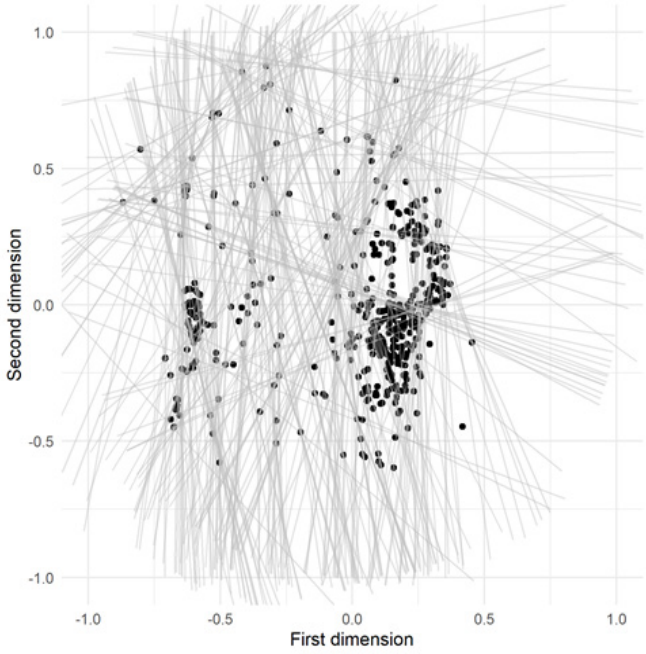
6. Distribution of the vote lines

I mentioned previously that the estimation of ideal points estimates both legislators' points and the vote lines that divide them. To do so, I used the information present in the *rollcalls* list that make up the object created by *Optimal Classification*. We can check, for every roll call vote, how many votes were correctly and incorrectly classified, the proportional reduction in error, and the necessary information to determine the lines that divide the parliamentarians.

```
head(chamber_oc$rollcalls)
##      correctYea wrongYea wrongNay correctNay       PRE normVector1D
## [1,]         NA       NA       NA         NA        NA           NA
## [2,]         NA       NA       NA         NA        NA           NA
## [3,]        297       12        1         72 0.8452381    0.9985689
## [4,]         94        4       36        216 0.6923077    0.9254343
## [5,]        121       31       13        244 0.6716418    0.2525467
## [6,]        212       26       20        112 0.6666667    0.2703514
##      normVector2D   midpoints
## [1,]           NA          NA
## [2,]           NA          NA
## [3,]  -0.05347965  0.23146515
## [4,]  -0.37890820 -0.06372496
## [5,]  -0.96758471  0.08072122
## [6,]  -0.96276172  0.08239360
```

The qualitative analysis of specific votes is also an important way of getting evidence on the substantive meaning of the dimensions estimated. We can not only identify the most frequent coalitions through this analysis, but we can also know the angle of these lines and determine which ones predominantly belong to the first dimension, which ones are influenced by a second dimension, and which are not explained well by either dimension. To add vote lines to the previous graph, the following code is necessary:

```
ggplot() +
  geom_point(data = chamber_oc$legislators,
             aes(coord1D, coord2D)) +
  geom_segment(data = data.frame(chamber_oc$rollcall),
               aes(x = (midpoints*normVector1D)+normVector2D,
                   y = (midpoints*normVector2D)-normVector1D,
                   xend = (midpoints*normVector1D)-normVector2D,
                   yend = (midpoints*normVector2D)+normVector1D),
               color = "grey", alpha = 0.4) +
  coord_cartesian(xlim = c(-1,1), ylim = c(-1,1)) +
  theme_minimal() +
  labs(x = "First dimension", y = "Second dimension")
```

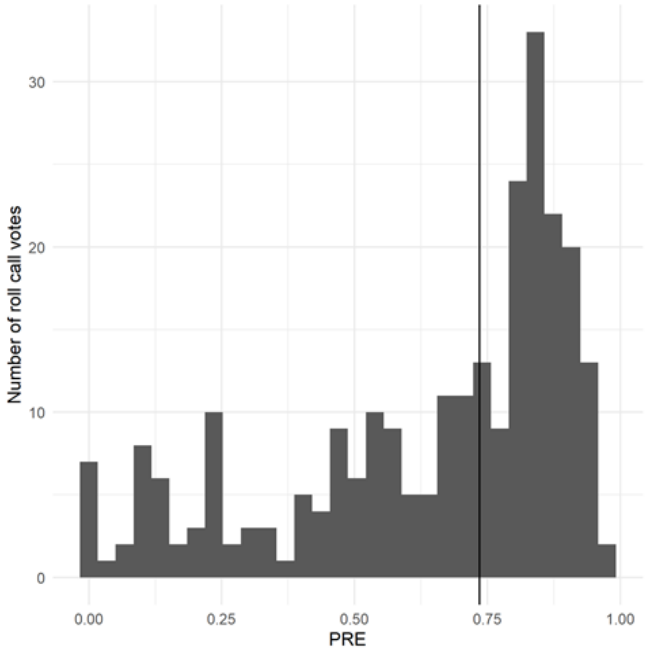


Grpah 5: Distribution of ideal points on both dimensions and vote lines

Note that in the geom_point() function, I specified the values of the ideal points, which are in the *legislators* list in the object created by *Optimal Classification*. In the geom_segment() function, I specified the values needed to calculate the initial and final points of the lines on the *x* and *y* coordinates, using the information from the roll call votes in the *rollcall* list of the same object created by *Optimal Classification*.

The graph gives us the impression that most lines have a more vertical inclination, crossing its *x* axis, indicating a dominance by the first dimension in the structuring of the voters' behaviors. However, we can analyze in more depth specific votes, to evaluate which are explained well by the estimation of ideal points. The first way of selecting votes is to observe which ones had a better proportional reduction in error. That information is available in the *legislators* list in object created by *Optimal Classification*. With the graph below we can observe the distribution of that value on all votes:

```
ggplot(data.frame(chamber_oc$rollcalls), aes(PRE)) +
  geom_histogram() +
  geom_vline(aes(xintercept = median(PRE, na.rm = T))) +
  theme_minimal() +
  labs(x = "PRE", y = "Number of roll call votes")
```



Graph 6: Distribution of the proportional reduction in error

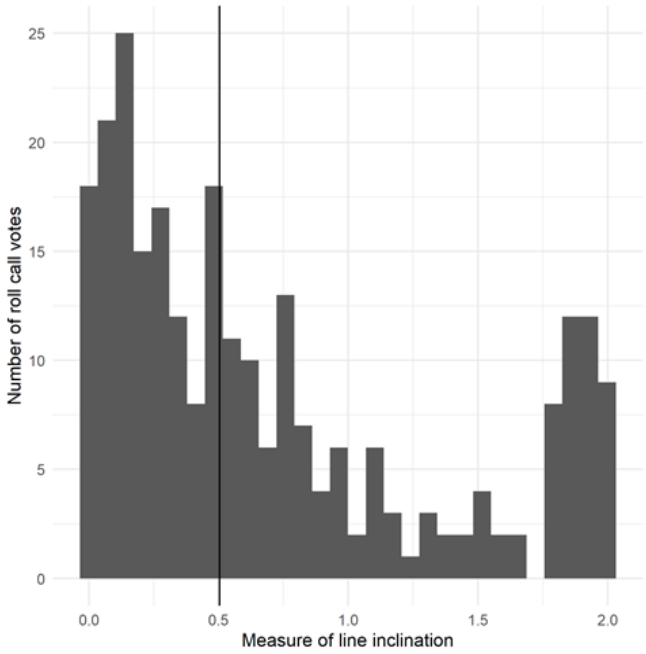
The median value of the proportional reduction in error is close to 0.75. Therefore, a first criterion to verify the substantive meaning of the first dimension should be to note the roll call votes with PRE higher than 0.75. That information alone does not help us in knowing which votes concern the first or second dimension, or both, but which votes are explained well by the estimation overall. One way of accessing which ones correspond to the first dimension is to use the same variables that were used to estimate the lines that divide the parliamentarians in roll call votes. We can do that in the following way:

```
x1 <- (chamber_oc$rollcalls[,"midpoints"]*chamber_oc$rollcalls[,"normVector1D"])+
  chamber_oc$rollcalls[,"normVector2D"]

x2 <- (chamber_oc$rollcalls[,"midpoints"]*chamber_oc$rollcalls[,"normVector1D"])-
  chamber_oc$rollcalls[,"normVector2D"]

x_diff <- abs(x2 - x1)

ggplot(NULL, aes(x_diff)) +
  geom_histogram() +
  geom_vline(aes(xintercept = median(x_diff, na.rm = T))) +
  theme_minimal() +
  labs(y = "Number of roll call votes", x = "Measure of line inclination")
```

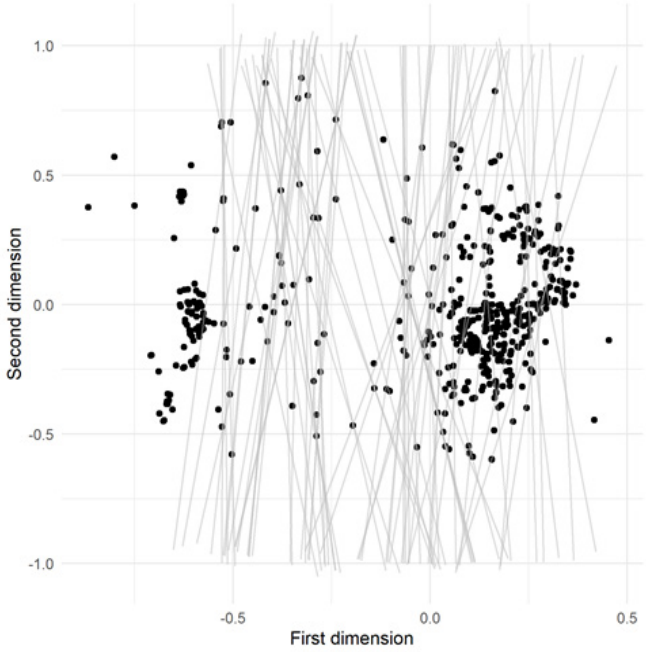


Graph 7: Measure of line inclination

The idea behind what I have done is: 1) calculate the value on the x axis, which corresponds to the first dimension, from the initial point in the line that divides the parliamentarians; 2) do the same for the final point; 3) know the absolute value of the difference between the two values. If the difference is low, that means that the line tends to have a more vertical position, being better explained by the first dimension. With the graph, we can verify that the median of that difference if very close to 0.5. therefore, values inferior to this are a second indicator of the first dimension's predominance. Lastly, I joined both criteria to select the votes:

```
selected_rollcalls <- which(chamber_oc$rollcalls[,"PRE"] > 0.75 &
                              x_diff < 0.5)

colnames(chamber_rollcall$votes)[selected_rollcalls]
##  [1] "PL-1292-1995-3"  "PL-1292-1995-7"  "PL-4742-2001-1"  "PLP-441-2017-7"
##  [5] "PLP-441-2017-9"  "PLP-441-2017-11" "MPV-852-2018-1"  "MPV-866-2018-1"
##  [9] "MPV-867-2018-3"  "MPV-867-2018-19" "MPV-871-2019-3"  "MPV-871-2019-4"
## [13] "MPV-871-2019-5"  "PEC-6-2019-2"    "PEC-6-2019-3"    "PEC-6-2019-4"
## [17] "PEC-6-2019-5"    "PEC-6-2019-6"    "PEC-6-2019-7"    "PEC-6-2019-8"
## [21] "PEC-6-2019-9"    "PEC-6-2019-13"   "PEC-6-2019-14"   "PEC-6-2019-18"
## [25] "PEC-6-2019-22"   "PEC-6-2019-24"   "PEC-6-2019-31"   "PEC-6-2019-32"
## [29] "PEC-6-2019-33"   "PEC-6-2019-35"   "PEC-6-2019-37"   "PEC-6-2019-38"
## [33] "PEC-6-2019-39"   "PEC-6-2019-40"   "PEC-6-2019-41"   "PEC-6-2019-42"
## [37] "PEC-34-2019-4"   "MPV-881-2019-1"  "MPV-881-2019-5"  "MPV-881-2019-6"
## [41] "MPV-881-2019-7"  "MPV-881-2019-8"  "MPV-881-2019-9"  "MPV-881-2019-11"
## [45] "PL-2787-2019-1"  "PL-2788-2019-2"  "REQ-1464-2019-1" "PL-2999-2019-6"
## [49] "PL-3261-2019-5"  "PL-3261-2019-6"  "PL-3261-2019-7"  "PL-3261-2019-9"
## [53] "PL-3261-2019-10" "PL-3261-2019-12" "PL-3261-2019-13" "MPV-886-2019-1"
## [57] "MPV-886-2019-2"  "PL-3715-2019-2"  "PL-3715-2019-9"  "MPV-890-2019-3"
## [61] "MPV-890-2019-5"  "MPV-890-2019-6"  "MPV-890-2019-7"  "MPV-890-2019-9"
## [65] "PL-4162-2019-1"  "PL-4162-2019-2"  "REQ-2110-2019-1" "REQ-2157-2019-1"
## [69] "REQ-2234-2019-1" "PDL-523-2019-1"  "PDL-523-2019-4"  "PL-5815-2019-1"
## [73] "PL-5815-2019-2"  "PL-5815-2019-3"
```

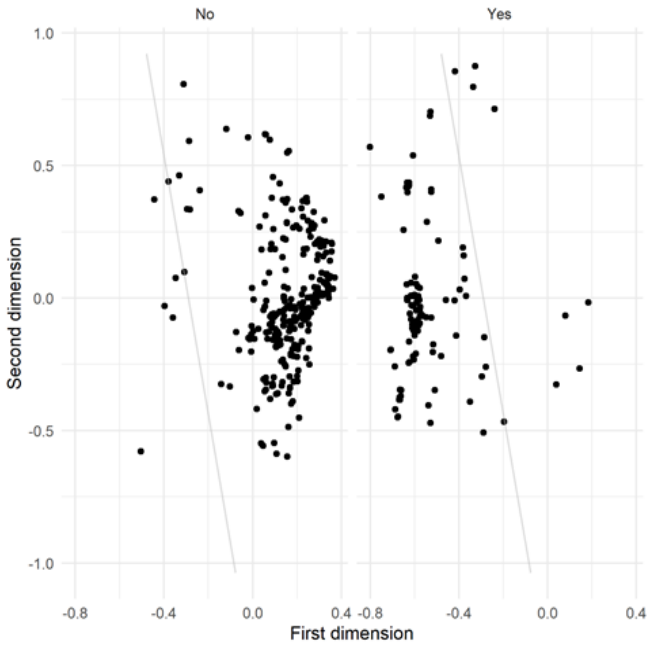


Graph 8: Ideal points and vote lines corresponding to the first dimension

Note that the difference is in the specification of the geom_segment() function, where I indicated in the *chamber_oc$rollcall* object that I only wanted the lines that correspond to the votes selected. We succeeded in our goal of selecting only the roll call votes that were explained well by the technique and concern the first dimension. We can observe that there are votes that end up splitting both the left-wing and the right-wing party blocs.

With the information gathered on each voter's ideal points and the lines for each roll call vote, it is possible to analyze specific decisions qualitatively. Thus, it is possible to better understand how the structure of preferences and political conflict is formed among deputies, contributing even more towards an interpretation of the substantive meaning of the dimensions estimated and the results overall. Below I show an example, with of one of the votes for the social security reform PEC.

```
chamber_oc$legislators %>%
  mutate(Vote = chamber_rollcall$votes[,147]) %>%
  mutate(Vote = case_when(Vote == 1 ~ "Yes",
                          Vote != 1 ~ "No")) %>%
  filter(!is.na(Vote)) %>%
ggplot() +
  geom_point(aes(coord1D, coord2D)) +
  geom_segment(data = data.frame(chamber_oc$rollcall) %>%
                 slice(147),
               aes(x = (midpoints*normVector1D)+normVector2D,
                   y = (midpoints*normVector2D)-normVector1D,
                   xend = (midpoints*normVector1D)-normVector2D,
                   yend = (midpoints*normVector2D)+normVector1D),
               color = "grey", alpha = 0.5) +
  theme_minimal() +
  facet_wrap(vars(Vote), nrow = 1) +
  labs(x = "First dimension", y = "Second dimension")
```

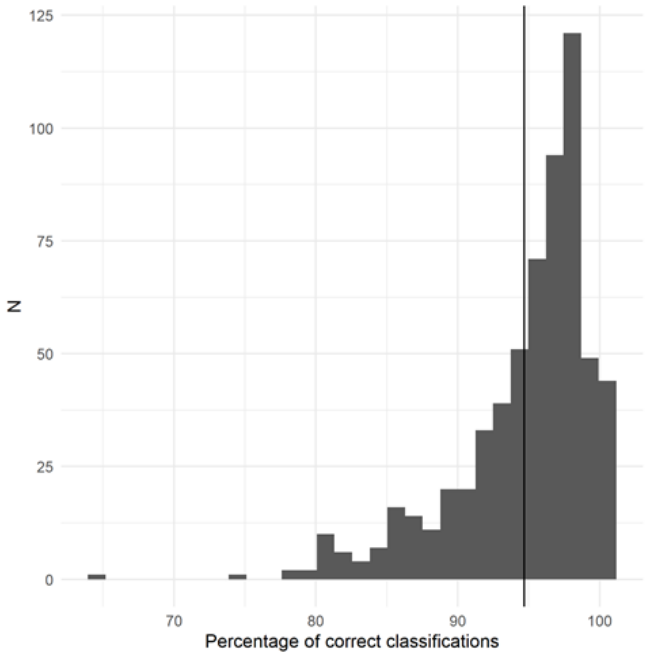


Graph 9: Ideal points and vote line for one of the votes concerning PEC 6/2019

In the code above, I used the object created by *Optimal Classification* and added a column for the votes from column 147, which refers to the selected roll call vote on social security reform. Next, I recoded the values to indicate if the vote was in favor or against the issue, removing those who did not vote, and generated the graph. I specified that two separate panels be made, according to the vote, and kept only the line for the selected roll call vote. We can see on the graph that government supporters who, in this case, voted against the proposal, are concentrated on the right side, with higher ideal points values in the first dimension, while the opposition voted for it. We can observe that the estimation gets wrong five deputies who voted *no* and eight deputies who voted *yes*, given that they are located at the opposite side the line than it would be expected. That roll call vote has 97.1% of the votes correctly classified and a PRE of 0.88. This graph enables a qualitative analysis of this specific vote and specific voters. One example is the behavior of deputy Tabata Amaral. Within the context of the social security reform, there was a lot of criticism, both on social media and internally in the PDT, about some of the votes that the deputy had in favor of the approved text. Her ideal point is the one that voted *no*, located higher on the graph. As we have noted, her vote is correctly classified by the estimation, placing her on the right side of the line that splits the vote. Thus, her vote is coherent with her behavior on other votes, placing her point in that space. In this case, even if located more to the left, her vote with those on the right would not be atypical, given how political conflict is structured in the Chamber of Deputies when we spatially analyze all the deputies and roll call votes.

Another possibility for analysis of voter behavior is to observe which ones are more "unpredictable", with a higher number of incorrect classifications because their behavior does not follow the logic of the dimensions found, but others not estimated. There are specific models to handle these types of voters more appropriately (Lauderdale, 2010). In the object created by *Optimal Classification*, columns are presented with the result of how many *yes* and *no* were correctly and incorrectly classified. Thus, it is possible to know the percentage of votes for each voter that the estimation of ideal points was able to adequately allocate in the space. We can get the results and elaborate a graph in the following way:

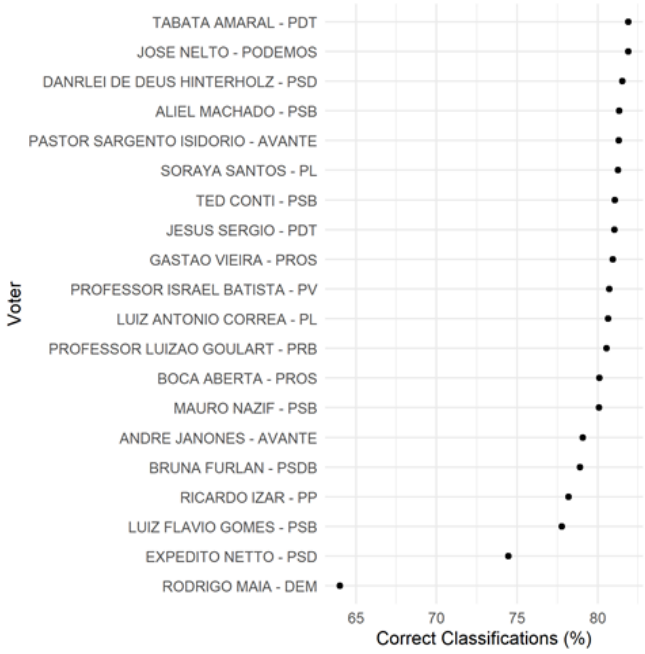
```
chamber_oc$legislators %>%
  mutate(correct_classifications = 100 * (correctYea + correctNay) /
           (correctYea + wrongYea + wrongNay + correctNay)) %>%
  ggplot(aes(correct_classifications)) +
  geom_histogram() +
  geom_vline(aes(xintercept = mean(correct_classifications, na.rm = T))) +
  theme_minimal() +
  labs(x = "Percentage of correct classifications", y = "N")
```



Graph 10: Distribution of the percentage of the voters' correct classifications

With the histogram, we can observe that the average of correct classifications of voters is 94.6% and the overwhelming majority of deputies is above 90%. To observe who are the voters whose behaviors are less explained by the two dimensions estimated, I have used the following code:

```
chamber_oc$legislators %>%
  mutate(correct_classifications = 100 * (correctYea + correctNay) /
           (correctYea + wrongYea + wrongNay + correctNay)) %>%
  arrange(correct_classifications) %>%
  slice(1:20) %>%
  mutate(legislator_id = fct_reorder(legislator_id, correct_classifications)) %>%
  ggplot(aes(correct_classifications, legislator_id)) +
  geom_point() +
  theme_minimal() +
  labs(x = "Correct Classifications (%)", y = "Voter")
```



Graph 11: Voters with smaller percentage of correct classifications

I kept in the graph only the 20 voters with the smallest percentage of correctly classified votes. We can observe that Rodrigo Maia is the one who the smallest percentage, with less than 65%, followed by Expedito Netto, with a little under 75%. It is curious that deputy Tabata Amaral, mentioned previously due to the critiques on her behavior, is among those whose behavior is not well explained by the two dimensions estimated.

7. Discussion

This paper showed how to estimate ideal points and some ways to explore the results from one of the technique's several applications. Many other R packages have been implemented to explain the behavior of various political actors with varied types of data.

For studies that work with text, instead of votes on bills, the choice of actors is in the use of words. When dealing with a certain topic, using certain words and omitting others would reveal an ideological preference on how to approach an issue. Moreover, the relevance of a topic itself, to the detriment of others, in a political speech, would reveal ideological content. Implementations such as *Wordfish* (Slapin and Proksch, 2008), *Wordscores* (Laver et al., 2003), and *Wordshoal* (Lauderdale and Clark, 2014) seek to deal with this type of issue, analyzing party manifestos, parliamentary speeches, among others.

The relative ease in obtaining Twitter data has also enable the use of this type of data for ideal point estimation. Not only the text of the tweets made by users can be analyzed with the implementations just mentioned, but also the way users interact and connect to one another. The act of choosing to follow certain accounts to the detriment of others would reveal an ideological preference of users of that social network in getting certain types of content (Barberá, 2015).

Several R packages that surface daily, more than expanding what type of data can be used for ideal point estimation, also look to expand the possibilities for better statistical inference from data already in use for a long time. We cannot estimate ideal points, for example, of the 2019 Chamber of Deputies and directly compare the results with an estimation from the previous year.

For that, theoretical and statistical assumptions are required that are implemented by other algorithms (Bailey et al., 2017; Martin e Quinn, 2002; Poole and Rosenthal, 2001). In addition to intertemporal comparison, it is possible compare ideal points between different institutions and actors, with the proper data and methods (Bailey, 2007; Bonica, 2014; Tausanovitch and Warshaw, 2013).

Regardless of future developments, it is necessary to understand the possibilities that these techniques offer in the present, with the limitations and virtues. I hope to have contributed so that new studies can adequately implement these techniques.

BIBLIOGRAPHICAL REFERENCES

ARMSTRONG, David A et al. (2014), *Analyzing spatial models of choice and judgment with R*. [S.l.] CRC Press.

BAILEY, Michael A. (2007), "Comparable preference estimates across time and institutions for the court, congress, and presidency". *American Journal of Political Science*, vol. 51,no. 3: 433–448.

BAILEY, Michael A & STREZHNEV, Anton & VOETEN, Erik. (2017), "Estimating dynamic state preferences from United Nations voting data". *Journal of Conflict Resolution*, vol. 61,no. 2: 430–456.

BARBERÁ, Pablo. (2015), "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data". *Political analysis*, vol. 23,no. 1: 76–91.

BONICA, Adam. (2014), "Mapping the ideological marketplace". *American Journal of Political Science*, vol. 58,no. 2: 367–386.

CARROLL, Royce et al. (2009), "Comparing NOMINATE and IDEAL: Points of difference and Monte Carlo tests". *Legislative Studies Quarterly*, vol. 34,no. 4: 555–591.

CARROLL, Royce et al. (2013), "The structure of utility in spatial models of voting". *American Journal of Political Science*, vol. 57,no. 4: 1008–1028.

CLINTON, Joshua & JACKMAN, Simon & RIVERS, Douglas. (2004), "The statistical analysis of roll call data". *American Political Science Review*, vol. 98,no. 2: 355–370.

CONVERSE, Philip E. (1964), "The nature of belief systems in mass publics. Ideology and discontent". *Ideology and Discontent*, 206–261.

DESPOSATO, Scott W. & INGRAM, Matthew C. & LANNES JR., Osmar P. (2015), "Power, Composition, and Decision Making: The Behavioral Consequences of Institutional Reform on Brazil's Supremo Tribunal Federal". *Journal of Law, Economics, and Organization*, vol. 31,no. 3: 534–567.

DOWNS, Anthony. (1957), *An economic theory of democracy*. New York, Harper; Row.

DUCK-MAYR, JBrandon & MONTGOMERY, Jacob M. *Ends Against the Middle: Scaling Votes When Ideological Opposites*. [S.d.].

FERREIRA, Pedro Fernando Almeida Nery & MUELLER, Bernardo. (2014), "How judges think in the Brazilian Supreme Court: Estimating ideal points and identifying dimensions". *Economia*, vol. 15,no. 3: 275–293.

IZUMI, Mauricio. (2016), "Governo e Oposição no Senado Brasileiro (1989-2010)". *Dados*, vol. 59,no. 1: 91–138.

LAUDERDALE, Benjamin. (2010), "Unpredictable voters in ideal point estimation". *Political Analysis*, no. 18: 151–171.

LAUDERDALE, Benjamin E. & CLARK, Tom S. (2014), "Scaling politically meaningful dimensions using texts and votes". *American Journal of Political Science*, vol. 58,no. 3: 754–771.

LAVER, Michael & BENOIT, Kenneth & GARRY, John. (2003), "Extracting policy positions from political texts using words as data". *American Political Science Review*, vol. 97,no. 02: 311–331.

LEONI, Eduardo. (2002), "Ideologia, Democracia e Comportamento Parlamentar: A Camara dos Deputados (1991-1998)". *Dados*, vol. 45,no. 3: 361–386.

LIMONGI, Fernando. (2019), "Presidencialismo do Desleixo". *Piauí*, no. 158: Disponível em: <https://piaui.folha.uol.com.br/materia/presidencialismo-do-desleixo/>.

MARIANO SILVA, Jeferson. (2018), "Mapeando o Supremo. As posições dos ministros do STF na jurisdição constitucional (2012-2017)". *Novos Estudos - CEBRAP*, vol. 37,no. 1: 35–54.

MARTIN, Andrew D. & QUINN, Kevin M. (2002), "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis*, vol. 10,no. 2: 134–153.

MCCARTY, Nolan & POOLE, Keith T. & ROSENTHAL, Howard. (2001), "The Hunt for Party Discipline in Congress". *American Political Science Review*, vol. 95,no. 3: 673–87.

MCDONNELL, Robert Myles & DUARTE, Guilherme Jardim & FREIRE, Danilo. (2019), "congressbr: An R Package for Analyzing Data from Brazil's Chamber of Deputies and Federal Senate". *Latin American Research Review*, vol. 54,no. 4:

POOLE, Keith et al. (2011), "Scaling Roll Call Votes with wnominate in R". *Journal of Statistical Software*, vol. 42,no. 14: 1–21.

POOLE, Keith T. (2005), *Spatial models of parliamentary voting*. [S.l.] Cambridge University Press.

POOLE, Keith T & ROSENTHAL, Howard. (1985), "A spatial model for legislative roll call analysis". *American Journal of Political Science*, 357–384.

POOLE, Keith T & ROSENTHAL, Howard. (1997), *Congress: A political-economic history of roll call voting*. [S.l.] Oxford University Press on Demand.

POOLE, Keith T & ROSENTHAL, Howard. (2001), "D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting". *Legislative Studies Quarterly*, 5–29.

POOLE, Keith T & ROSENTHAL, Howard. (1991), "Patterns of congressional voting". *American Journal of Political Science*, 228–278.

ROSENTHAL, Howard & VOETEN, Erik. (2004), "Analyzing roll calls with perfect spatial voting: France 1946–1958". *American Journal of Political Science*, vol. 48,no. 3: 620–632.

SLAPIN, Jonathan B & PROKSCH, Sven-Oliver. (2008), "A scaling model for estimating time-series party

positions from texts". *American Journal of Political Science*, vol. 52,no. 3: 705–722.

SPIRLING, Arthur & MCLEAN, Iain. (2007), "UK OC OK? Interpreting optimal classification scores for the UK House of Commons". *Political Analysis*, 85–96.

TAUSANOVITCH, Chris & WARSHAW, Christopher. (2013), "Measuring constituent policy preferences in congress, state legislatures, and cities". *The Journal of Politics*, vol. 75,no. 2: 330–342.

WICKHAM, Hadley. (2014), "Tidy data". *Journal of Statistical Software*, vol. 59,no. 10: 1–23.

WICKHAM, Hadley et al. (2019), "Welcome to the Tidyverse". *Journal of Open Source Software*, vol. 4,no. 43: 1686.

ZUCCO, Cesar. (2009), "Ideology or What? Legislative Behavior in Multiparty Presidential Settings". *The Journal of Politics*, vol. 71,no. 3: 1076–1092.

ZUCCO, Cesar & LAUDERDALE, Benjamin E. (2011), "Distinguishing Between Influences on Brazilian Legislative Behavior". *Legislative Studies Quarterly*, vol. 3,no. 36: 363–396.

Tempo, tempo, tempo: regressão de Cox na ciência política

Rodrigo Lins - Universidade Federal de Sergipe

Resumo

Este artigo tem como objetivo apresentar o modelo de riscos proporcional de Cox. Para tanto, será feito um debate teórico geral dos modelos de sobrevivência, assim como a apresentação de termos comuns para a técnica. Além disso, será feita uma apresentação passo a passo de como elaborar um modelo de Cox no software R. Espera-se que este trabalho ajude na difusão da técnica dentro das ciências sociais.

Palavras-chave: Análise de Sobrevivência; Modelo de Cox; R.

Abstract

This article aims to present Cox Proportional-Hazards Model. To do so, a general theoretical debate on survival models will be made, as well as the presentation of common terms for the technique. In addition, there will be a step-by-step presentation of how to develop a Cox model in the R software. It is hoped that this work will help to disseminate the technique within the social sciences.

Keywords: Survival Analysis; Cox Model; R.

1. Introdução[1]

A análise de sobrevivência é uma técnica de vários nomes: análise de sobrevida; análise de evento histórico; modelo de duração; e tempo até evento (*time-to-event*) são apenas alguns dos nomes dados a essa ferramenta estatística. No entanto, apesar da variação de nome, todos eles apresentam uma característica chave: o tempo. Assim, não é surpresa que a preocupação dos modelos de sobrevivência não seja só o "se", mas também "quando" um dado evento[2] acontece (GUO, 2010). Isto é: o interesse do pesquisador é a observação de uma mudança qualitativa da variável dependente em um ponto específico no tempo (ALLISON, 1984).

Dentro das possíveis abordagens de uma análise de sobrevivência, o modelo de Cox de riscos proporcionais tem lugar de destaque (BOX-STEFFENSMEIER & JONES, 2004). Conceituado como um modelo semiparamétrico[3], ele é bastante maleável. Ele não assume uma distribuição estatística. Assume apenas que a ação das variáveis independentes sobre o risco é multiplicativa (CARVALHO ET AL., 2011), isto é, para cada incremento de unidade da variável explicativa, o risco é multiplicado pelo seu coeficiente exponenciado (ALLISON, 1984).

E como anda a utilização da análise de sobrevivência – e mais especificamente, do modelo de Cox – na ciência política brasileira? É possível notar que o uso ainda não decolou. Analisando os resumos de artigos disponíveis nos arquivos de cinco das principais revistas brasileiras de ciência política[4], só houve um retorno que indicasse o uso de uma análise de sobrevivência[5]. Assim, este trabalho pretende apresentar a técnica ao grande público das ciências sociais para que seja mais considerada.

O restante deste artigo está dividido da seguinte forma: primeiro focarei no modelo de Cox, com abordagem mais teórica; em seguida, apresentarei um breve glossário de termos comuns nas análises de sobrevivência, mas que não costumam ser tão utilizados em outras técnicas estatísticas. A seção seguinte apresentará as possíveis aplicações em diferentes áreas do conhecimento. Finalmente, na última etapa, apresentarei um passo a passo de como realizar uma análise de sobrevivência no software R e como interpretar os seus resultados.

2. O modelo de Cox: o que é, para que serve, quando usar?

O que é o modelo de riscos proporcionais de Cox? Antes, é preciso entender que as modelagens em análise de sobrevivência podem ser divididas em dois grupos: paramétricos e o semiparamétrico[6]. Os paramétricos são aqueles que pressupõem uma determinada distribuição, como o modelo de Weibull ou seu formato exponencial (BOX-STEFFENSMEIER & JONES, 2004). Enquanto a última espera que o risco seja constante no tempo, a de Weibull permite que o risco aumente ou diminua linearmente no tempo (ALLISON, 1984).

1 Agradeço aos pareceristas anônimos que me ajudaram a melhorar substancialmente este artigo. Erros remanescentes são de responsabilidade exclusiva do autor.

2 Na literatura de análise de sobrevivência, a o termo "evento" também pode ser chamado de "falha". No entanto, como a mudança qualitativa nem sempre indica uma falha de fato, podendo ser apenas uma transição de governo, por exemplo, optei por sempre usar "evento".

3 Modelos paramétricos e semiparamétricos são explicados na próxima seção.

4 A DADOS foi a única que apresentou um retorno, em artigo publicado em 2019. As outras quatro revistas incluídas na busca foram: Brazilian Political Science Review; Opinião Pública; Revista Brasileira de Ciências Sociais; e Revista de Sociologia e Política. Os termos de busca foram dois: Cox e sobrevivência. A pesquisa foi feita em 2 de junho de 2020.

5 Isso também pode indicar que os resumos dos artigos não apresentam claramente o desenho de pesquisa e, assim, não antecipam as ferramentas analíticas utilizadas.

6 Ainda é possível encontrar estimações não-paramétricas. No entanto, não é possível estimar o efeito das variáveis explicativas nesses modelos. O estimador produto de Kaplan-Meier, por exemplo, compara a curva de sobrevivência de dois ou mais grupos, não considerando variáveis independentes. Ver Lee e Wang (2003), especialmente os capítulos 4 e 5.

Já o modelo semiparamétrico – o modelo Cox – não assume qualquer distribuição para a função de risco basal. Essa seria uma característica que poderia o colocar em um grupo não-paramétrico. No entanto, é esperado que suas variáveis ajam sempre com efeito multiplicativo sobre o risco. E esta é a característica que o credenciaria para o grupo paramétrico. É por isso, portanto, que o modelo de Cox é considerado um modelo semiparamétrico.

Respondendo, então, de forma mais direta o questionamento do início desta seção, podemos afirmar que o modelo de Cox: (1) assim como modelos paramétricos, permite a inclusão de variáveis explanatórias; (2) o efeito destas variáveis é multiplicativo; e (3) ele não carrega nenhum pressuposto no que diz respeito à distribuição estatística para a função de risco basal. O modelo de Cox é, então, uma técnica estatística para explorar a relação de um evento específico com uma série de variáveis exploratórias (WALTERS, 2009).

Ele foi desenvolvido por Cox (1972), de quem herdou o nome, e surgiu como forma de flexibilizar os pressupostos já mencionados nos modelos paramétricos. É chamado de modelo de riscos proporcionais porque para quaisquer duas observações em qualquer ponto no tempo, a razão de seus riscos deve ser constante. Em outras palavras, a razão dos riscos de uma observação A para uma observação B deve ser o mesmo tanto em t_0 quanto em t_{100}, por exemplo (ALLISON, 1984). Pode ser simbolizado pela equação a seguir:

$$\log h(t) = a(t) + b_1 x_1 + b_n x_n \quad (1)$$

É possível notar que o lado esquerdo da equação é o logaritmo natural da função de risco. Do outro lado, a função de risco basal a(t) pode ser qualquer função de tempo. Como já vimos, é justamente por não precisar ser especificado que o modelo de Cox é tido como parcialmente paramétrico ou semiparamétrico (ALLISON, 1984). Em b_1x_1, b representa o coeficiente e x o valor da variável independente 1. A fórmula (1) é aplicada ao tipo de dados que será visto na tabela 1, em que cada observação ocupa uma linha, pois as variáveis X_1 e X_n são constantes no tempo.

Modelo de Cox estendido[7]

É comum que quando se trabalhe com dados longitudinais, algumas das variáveis explicativas variem de acordo com o tempo. Digamos que o PIB per capita seja uma das variáveis independentes. É difícil imaginar que ela se mantenha constante com o passar dos anos. Dessa forma, como incluí-la em um modelo? É justamente nestes casos, quando existem variáveis dependentes do tempo (ou que variam no tempo), que o modelo de Cox estendido deve ser usado. Pode ser representado da seguinte forma:

$$\log h(t) = a(t) + b_1 x_1 + b_n x_n(t) \quad (2)$$

Aqui, a(t) segue podendo ser qualquer função de tempo e a variável X_1 segue constante no tempo. No entanto, a variável X_n varia, sim, em função do tempo. Para esse tipo de desenho, o pesquisador deve construir um banco de dados pelo processo de contagem, similar ao que será visto na tabela 2, onde uma nova linha é adicionada sempre que houver mudança nos valores de X_n.

Os resíduos de Schoenfeld

Existem diferentes formas de testar o ajuste do modelo de uma regressão de Cox. No entanto, considerando que a proporcionalidade dos riscos é uma das características definidoras do modelo, é importante avaliar se tal pressuposto está sendo cumprido. Os resíduos de Schoenfeld servem justamente para essa finalidade[8].

7 Vale ressaltar que além de sua forma estendida, o modelo de Cox também se permite ser estratificado. Para uma visão mais completa do modelo de Cox estratificado, ver capítulo 6 de Carvalho et al. (2011).

8 Existem outras formas de analisar a proporcionalidade do risco. Um caminho comum é via resíduos martingale (ver Lin et al.,1993). Ainda existem as possibilidades de resíduos deviance e resíduos escore (ver Carvalho et al., 2011).

Ele apresenta a proporcionalidade de cada covariável definido para cada indivíduo (CARVALHO ET AL., 2011). A equação que ilustra o cálculo dos resíduos é a seguinte:

$$r_{ik} = \delta_i(x_{ik} - a_{ik}) \quad (3)$$

onde o a ocorrência da falha para a observação *i* é simbolizada por δ_i. O valor da variável específica se dá por x_{ik} e, finalmente, a média ponderada das covariáveis em risco em um determinado tempo *t* é a_{ik}.

Embora existam diferentes formas de analisar os resíduos, a maneira mais intuitiva de compreender a análise, como argumentado por Schoenfeld (1982), é a de que os resíduos devem ser independentes do tempo. Assim, se uma correlação entre os resíduos e o tempo não for estatisticamente significativa, isso quer dizer que os coeficientes não são dependentes do tempo e, assim, o pressuposto foi observado. Caso seja estatisticamente significativa, é indicação de que a proporcionalidade dos riscos não foi atendida.

Como será visto na seção prática, a análise dos resíduos de Schoenfeld também pode ser vista graficamente, facilitando sua interpretação. Como é esperado que a correlação seja aleatória, a linha do gráfico de dispersão deve seguir uma reta horizontal. Caso exista alguma tendência, positiva ou negativa, a proporcionalidade dos riscos foi ferida.

Em caso de não observância da proporcionalidade de risco, o primeiro passo é verificar se isso acontece de forma natural ou artificial. Assim, é importante verificar a presença de *outliers*. Caso a quebra de proporcionalidade seja real, duas abordagens são possíveis: (1) buscar um novo modelo, como o modelo de Cox estratificado (ATA & SÖZER, 2007; BORUCKA, 2013); ou (2) adicionar um efeito de interação entre a variável que fere o pressuposto de proporcionalidade e alguma função de tempo[9] (BOX-STEFFENSMEIER & ZORN, 2001).

3. Pequeno glossário de sobrevivência

Antes de seguir adiante, é importante que seja feito um breve elucidário. Como os modelos de sobrevivência lidam com vocábulos que são pouco usuais a outras técnicas estatísticas, julgo que seja justificado apresentar ao menos os principais termos. Os conceitos apresentados nessa seção são válidos não só para os modelos de Cox, mas para análise de sobrevivência em geral. Esse pequeno glossário não tem intenção de ser exaustivo. Isso implica dizer que só serão apresentados aqui os termos mais comuns e relevantes para a compreensão dos modelos de sobrevivência[10]. Para começar, tais modelos podem lidar com dois tipos de dados incompletos: censurados e trucados.

Censurados/censured: Uma observação é considerada censurada quando seu início ou fim não são observados, fazendo com que seja difícil determinar um ponto específico no tempo. Existem três diferentes tipos de censura: (1) censura à direita; (2) censura à esquerda; e (3) censura intervalar. A primeira acontece quando a observação chega ao fim do estudo sem experimentar o evento ($T>t^-$)[11]. Assim, o pesquisador sabe apenas que a observação sobreviveu até determinado momento. Mas não sabe quando,

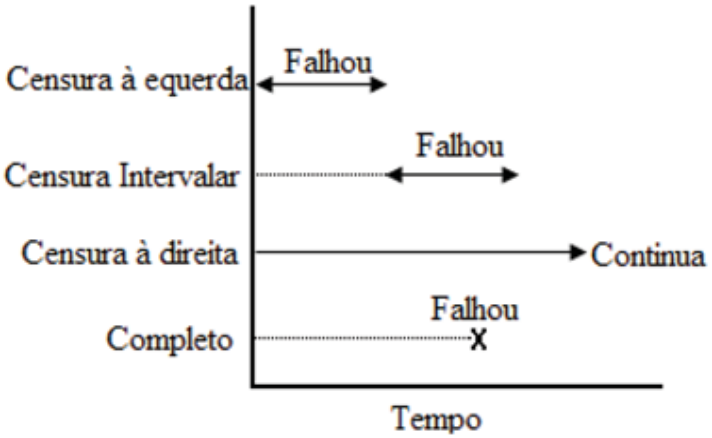
9 Em geral, é utilizado o logaritmo natural. Ver Box-Steffensmeier e Zorn (2001).

10 Os termos serão apresentados em português, com seu correspondente em inglês. O único objetivo é facilitar o estudo de quem quer trabalhar com análise de sobrevivência, evitando que o leitor se perca na tradução de termos técnicos.

11 T representa o tempo de sobrevivência ou tempo até o desfecho (CARVALHO et al., 2011). Ele é o intervalo entre a inserção da observação no estudo até o momento que ele experimenta a falha ou é censurado. t^- indica que o tempo de estudo é menor do que a sobrevivência da observação e t+ simboliza que o tempo de estudo supera o momento em que a observação foi censurada.

ou se chegou a experimentar o evento (GUJARATI, 2011). Dizemos que os dados têm censura a esquerda quando não sabemos quando o evento aconteceu, mas sabemos que ocorreu antes de um dado momento (T<t+). Finalmente, a censura intervalar ($t^- < T < t+$) indica que o evento aconteceu dentro de um período de tempo, mas o pesquisador não sabe exatamente quando (CARVALHO ET AL., 2011; KLEINBAUM & KLEIN, 2005). A figura 1 ilustra esses três tipos de censura.Testo

Figura 1 – Censura



Fonte: Lins et al. (2017).

Existem ainda outros três tipos de censura: de tipo I, tipo II e aleatória, comuns em estudos clínicos (COLOSIMO & GIOLO, 2006). A censura tipo 1, de maneira similar à censura à direita, acontece quando o estudo é encerrado após um tempo pré-estabelecido e, neste momento, algumas observações ainda não tinham sofrido o evento de interesse. A censura de tipo II ocorre quando o estudo é encerrado depois que um número pré-estabelecido de eventos acontece. Finalmente, a censura de tipo aleatório é aquela em que uma observação é retirada do estudo sem que a falha tenha acontecido (COLOSIMO & GIOLO, 2006).

Truncados/truncated: Os dados também podem ser truncados à direita e à esquerda. Truncamento à direita acontece quando são selecionados casos que sofreram o evento. Isso implica dizer que todas as observações sofreram o evento de interesse antes do limite final do estudo. Já o truncamento à esquerda conta com os casos que sofreram o evento após o limite inferior da janela de estudo (CARVALHO ET AL., 2011). Vale destacar também as observações com entradas atrasadas. São os casos que entraram no banco de dados uma vez que a pesquisa já havia iniciado. Esses casos costumam ser erroneamente tratados também como truncamento à esquerda (HARRELL, 2015).

Outros termos relevantes para a compreensão dos modelos de sobrevivência são:

Função de sobrevivência/survival function - S(t): É a probabilidade de que uma observação sobreviva mais tempo do que um algum período específico *t*. Ela é representada pela seguinte fórmula:

$$S(t) = P(T > t) \quad (4)$$

Função de risco/hazard function – h(t): A função de risco apresenta o potencial de falha em cada unidade de tempo, considerando que a observação sobreviveu até o tempo sendo observado *t*. Diferentemente da função de sobrevivência, ela não é uma probabilidade, e sim

uma taxa, podendo assumir qualquer valor maior do que zero. Pode ser representado como:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (5)$$

Em termos visuais, a função de sobrevivência apresenta uma tendência de queda, indicando que a probabilidade de sobrevier cai à medida que o tempo passa (KLEINBAUM & KLEIN, 2005). Já a função de risco pode ser constante ou aumentar/diminuir com o tempo. A função de risco deve ter mais relevância porque os modelos de regressão em uma análise de sobrevivência analisam justamente o risco de um dado evento acontecer. Isso ficará claro mais à frente, quando tratarmos especificamente dos modelos de Cox.

Verossimilhança parcial/partial likelihood. Os coeficientes dos modelos de Cox são estimados a partir de uma verossimilhança parcial[12]. Diferentemente do formato tradicional[13], a função do risco basal[14] (*baseline hazard*) é retirada do cálculo para o formato parcial por não ser parametrizado (BOX-STEFFENSMEIER & JONES, 2004). São consideras apenas as informações das observações que estão sob risco a cada tempo *t*.

$$PL = \prod_{j=1}^{K} L_j \quad (6)$$

onde K é o número total de eventos na amostra e Lj denota a verossimilhança de falhar no tempo j, dado que a observação sobreviveu até este ponto (KLEINBAUM & KLEIN, 2005).

Estrutura dos dados. Finalmente, também é importante falar sobre a estrutura dos dados de uma análise de sobrevivência. A construção correta do banco desempenha papel importante na hora da execução da análise. Quatro variáveis são obrigatórias ou esperadas[15]. A variável *início* e a variável *fim* indicam, respectivamente, o momento em que a observação entrou na pesquisa e o momento em que sofreu o evento ou foi censurado. Além disso, é recomendável a inserção da variável *tempo*, que indica justamente o período em que a observação esteve presente no banco. De forma simples, ela é a subtração da variável *fim* pela variável *início*. Finalmente, a variável binária *status*, que indica se a observação sofreu o evento ou não. Dentro do modelo tradicional, que é o utilizado em uma regressão de Cox, cada observação ocupará apenas uma linha dentro do banco. A tabela abaixo ilustra um banco tradicional. Ela indica um estudo hipotético que começa no tempo 0 e termina no tempo 30.

12 Uma apresentação básica de máxima verossimilhança é dada por Crawley (2015): "given the data, and given the model, what values of the parameters of that model make the observed data most likely?" (CRAWLEY, 2015, p. 6). No entanto, por ser considerado um método "intrincado" (GUJARATI & PORTER, 2011), não me aprofundarei aqui nos conceitos de verossimilhança.

13 Verossimilhança usual é utilizada no cálculo do coeficiente para os modelos paramétricos de sobrevivência. Ver apêndice A de Allison (1984) e capítulo 3 de Box-Steffensmeier e Jones (2004).

14 A função basal, ou função de base, é assim chamada pois nos casos em que as variáveis independentes têm valor 0, o valor da função de risco é igual ao valor da função basal (COLLETT, 1994).

15 O número exato das variáveis necessárias varia de acordo com o modelo. Para trabalhar com o modelo tradicional de Cox no software R, apenas a variável tempo e status são necessárias. Já para trabalhar com o modelo estendido, são necessárias início, fim e status. Por isso, o recomendável é que todas as quatro variáveis estejam presentes.

Tabela 1 – Banco de dados

ID	Início	Fim	Tempo	Status	X1	X2
A	0	30	30	0	1	1
B	0	27	27	1	1	0
C	1	15	14	1	0	0
D	2	30	28	0	1	1
E	3	15	12	1	0	1

Fonte: autor.

Pode-se perceber que temos cinco observações e uma linha para cada. A primeira observação faz parte do estudo desde o começo, pois seu início é 0. Chegou ao tempo 30 e não sofreu o evento, pois seu status também é 0. Já a terceira observação entrou no tempo 1 do estudo e saiu no tempo 15, tendo passado um total de 14 unidades de tempo na pesquisa. Sabemos que saída se deu por ter sofrido o evento, como indicado pela variável status de valor 1. As observações 1 e 4 servem de exemplo de censuras à direta, pois chegaram ao fim do estudo sem ter experimentado o evento. As variáveis X1 e X2 da tabela 1 representam variáveis explicativas16 com valores dicotômicos e constantes no tempo.

Veremos mais adiante que o banco exemplificado na tabela 1 serve para as regressões tradicionais de Cox. No entanto, se entre as variáveis explicativas houver alguma que varie no tempo, a estrutura deve ser modificada. Neste caso, a observação ganhará uma nova linha sempre que o valor da variável independente mudar. A tabela 2 ilustra tal situação.

Tabela 2 – Banco de dados
(Variáveis dependente do tempo)

ID	Início	Fim	Status	X1	X2
A	0	11	0	1	100
A	11	22	1	1	120
B	0	15	0	0	40
B	12	23	0	0	42
B	23	30	0	0	47

Fonte: autor

Mais uma vez, a tabela indica um estudo hipotético com duração de 30 anos. Pela variável status, sabemos que a observação 1 sofreu o evento enquanto a observação 2 é censurado à direita. A variável independente X1 continua sendo dicotômica e constante no tempo. Mas agora a variável X2 apresenta mudanças. E a cada mudança, uma nova linha é inserida. Para observação 1, ela apresenta apenas uma alteração de valor durante a estada da observação no banco. Por isso, só uma linha a mais é inserida. Já para a observação 2, existem duas mudanças no valor de X2. Dessa forma, duas novas linhas são inseridas. A quebra sempre acontece no tempo em que a alteração ocorre. Uma outra diferença importante é que a variável tempo não faz mais parte do banco. Ela

16 Variáveis explicativas são aquelas teorizadas como causa de variação na variável dependente (KELLSTEDT & WHITTEN, 2018).

contabilizaria cada quebra de linha como a inserção de uma nova observação, zerando o valor sempre que isso acontecesse. Dessa forma, ela é excluída do banco[17]. Esse banco é construído pelo chamado processo de contagem.

Existe ainda uma outra forma de estruturar os dados, caso o pesquisador esteja lidando com eventos recorrentes ou múltiplos. Neste caso, uma vez que uma determinada observação experimenta o evento de interesse, ele deve sair do banco de dados, voltando uma vez que retorna ao grupo de risco. Na ciência política, um exemplo claro – que apresentarei em maior detalhe na seção prática – se dá nos estudos de transição de regime. Se o evento de interesse for a mudança de regime democrático para regime autocrático, o país sairá do banco assim que sofrer o revés institucional e permanecerá fora do grupo de risco enquanto for uma ditadura, pois não há como ele sofrer o evento novamente. No entanto, caso o país transite de volta para a democracia, ele é recolocado no grupo de risco e entra novamente no banco de dados. A tabela a seguir mostra como exemplo um excerto de um banco de dados com informações para a Argentina, construído a partir dos dados de Cheibub, Gandhi e Vreeland (2010).

Tabela 3 – Banco de dados
(Eventos recorrentes)

ID	Início	Fim	Status
ARG	1946	1955	1
ARG	1958	1962	1
ARG	1963	1966	1
ARG	1973	1976	1
ARG	1983	2008	0

Fonte: autor.

Percebe-se que entre 1966 e 1973 e entre 1976 e 1983, por exemplo, a Argentina esteve fora do banco de dados. Isso aconteceu porque durante esses períodos o país era uma ditadura. Dessa forma, não poderia estar em risco de se tornar um país autoritário. Ainda que tenha acontecido golpes dentro de golpes durante o período, todos os eventos ocorreram dentro de regimes autoritários. A tabela mostra que sempre que a observação sofre o evento, uma nova linha é criada. Mas ele só retorna ao banco se as condições iniciais forem restabelecidas. Também é importante ressaltar que para que os anos fossem apresentados da maneira que estão, o banco de dados deveria começar no ano 0.

Tendo compreendido o que são dados censurados e truncados, o que são as funções de sobrevivência e de risco, e visto como os dados devem ser estruturados, passo para a próxima seção: modelos de regressão de Cox.

4. Um modelo eclético: aplicações em diferentes áreas

Modelos de análise de sobrevivência são bastante ecléticos no que diz respeito às disciplinas em que podem ser aplicados. Apesar do conteúdo poder ser bastante diferente ao cruzar fronteiras de conhecimento, a técnica em si se adapta bem. Para ver como ela funciona em diferentes disciplinas, apresento trabalhos em três áreas diferentes: medicina, economia e ciência política.

17 A rigor, ela pode ser mantida. Mas o pesquisador precisa ficar atento ao fato de que ela contabilizaria o tempo de permanência da observação apenas para os intervalos em que não há alteração na variável que tem seu valor alterado. Para que isso não aconteça, o pesquisador precisa fazer as alterações "manualmente". No entanto, como será visto na seção prática, ela pode ser descartada por completo por não ser necessária na construção de um modelo de Cox estendido no R.

Nesta última área, por ser o foco deste trabalho, apresentarei três exemplos. Procurei apresentar a hipótese e o modelo utilizado para testá-la. Por ser o foco deste trabalho, todos utilizaram o modelo de Cox. O quadro no fim da seção sumariza as variáveis de tempo e de status de cada estudo.

Medicina

A própria terminologia do nome da técnica já indica que ela surgiu na área de saúde. Especificamente, na biomedicina. O intuito inicial era estudar o tempo de sobrevivência de pacientes entre o momento do diagnóstico de uma dada doença até a morte do indivíduo (GUO, 2010). Assim, não é surpresa que o número de pesquisas que fazem uso da técnica na área de medicina seja elevado (COLOSIMO & GIOLO, 2006; MACHIN, CHEUNG & PARMAR, 2006).

Para exemplificar a utilização do modelo de Cox na área, vamos olhar o trabalho de Khambaty, Callahan e Stewart (2018). Os autores partem do pressuposto de que existe uma relação entre depressão e diabetes. Mais especificamente, afirmam que depressão poderia atuar como um fator de risco independente para incidência de diabetes. Dessa forma, eles levantam a seguinte pergunta: ao tratar pacientes para depressão também, os médicos também estariam reduzindo o excesso de risco de incidência de diabete. Para tanto, eles testaram o impacto que um programa de cuidados colaborativos com duração de 12 meses teria sobre um grupo de idosos deprimidos e inicialmente livres de diabete. O estudo fez *follow-up* por nove anos e contou com 119 pacientes.

É possível perceber, portanto, a variável que indica falha: desenvolvimento da doença. E ela é instrumentalizada com 0 se o paciente não desenvolver a doença e 1 caso desenvolva. A variável independente também é binária, indicando se o participante fez parte do grupo de tratamento (1) ou do grupo de controle (0). Os autores também rodaram um segundo modelo, incluindo algumas variáveis pessoais como controles: idade, sexo, raça/etnia, histórico de hipertensão, de fumo, e o índice de massa corporal (IMC). Os resultados, no entanto, indicaram que apesar do grupo de tratamento indicar uma queda no nível de depressão, quando comparado com o controle, os grupos não divergiram em relação à incidência de diabete. Assim, os autores argumentam que tratamento para depressão não é, por si só, suficiente para diminuir o risco dos pacientes desenvolverem a doença. Como uma das variáveis de controle tem seus valores alterados no tempo (IMC), os autores utilizaram um modelo de Cox estendido, como visto na fórmula (5).

Economia

A análise de sobrevivência em estudos econômicos pode ter várias aplicações. Desde estudos sobre falência de empresas (LEE, 2014) a pesquisas sobre estabilidade de bancos (PAPPAS ET AL., 2017), duração de desemprego (BIESZK-STOLORZ & MARKOWICZ, 2015) e processo de gentrificação (GRUBE-CAVERS & PATTERSON, 2015). No entanto, irei focar no trabalho de Gury (2011). Ele analisa os determinantes da taxa de abandono dos estudantes de ensino superior na França, i.e., o tempo de duração dos alunos no ensino superior. Assim, a variável status já pode ser desenhada: 0 se os estudantes não abandonam o ensino superior e 1 se abandonam. Para tanto, Gury (2011) utilizou dados do Ministério da Educação da França para acompanhar 5.383 estudantes que ingressaram no ensino médio em 1989 e, posteriormente, se matricularam em uma instituição de ensino superior.

Com o objetivo de explicar o comportamento dos estudantes, ele utiliza três grandes grupos de variáveis independentes: a de características pessoas (sexo e educação formal dos pais, por exemplo); as relacio-

nadas à educação secundária das observações (o tipo de diploma obtido[18], por exemplo); e relacionadas ao ensino superior (se o estudante declarou ter algum projeto profissional).

Para testar as hipóteses, Gury (2011) desenvolve três modelos, indo do modelo mais simples (com apenas o primeiro grupo de variáveis independentes) ao mais completo (com todos os três grupos de variáveis explicativas). Como resultado, o autor encontra um padrão não-linear, indicando que existem dois grupos de estudantes. Os chamados *early leavers*, que abandonam o curso nos dois primeiros anos; e os *late leavers*, que abandonam a partir do terceiro ano. Entre os que sofrem o evento – abandono do ensino superior –, as mulheres tendem a sair já no primeiro ano enquanto os homens abandonam a universidade já no fim do curso. Outras características socioeconômicas, como o nível de educação formal dois pais, só têm efeito no início dos estudos. À medida que os anos passam, as diferenças tendem a desaparecer. Nesse estudo, as variáveis não mudam com o tempo, proporcionando ao autor a possibilidade de trabalhar com o modelo tradicional de Cox, como visto na fórmula (4).

Ciência Política

Para o terceiro exemplo da utilização de modelos de sobrevivência, apresento um estudo na ciência política. Vale destacar que a utilização de tais modelos na disciplina vem ocupando um espaço costumeiramente povoado por modelos logit e probit. Várias áreas dentro da disciplina servem como terreno fértil para os modelos de duração. Por exemplo, eles já foram usados em estudos de transição de regime político (SVOLIK, 2008); para estudo de sobrevivência de governos (DIERMEIER & STEVENSON, 1999; SOMER-TOPCU & WILLIAMS, 2008); e a análise de difusão de políticas públicas (JONES & BRANTON, 2005; BATISTA, 2017).

O primeiro (1) exemplo que vamos olhar é o trabalho de Tansey, Koehler e Schmotz (2017), que estudam o efeito de ligações autocráticas e a sobrevivência de regimes ditatoriais. A hipótese é apresentada da seguinte maneira: quanto maior for o nível de ligações internacionais autocráticas, menor será o risco de quebra do regime autocrático. Tansey et al. (2017) instrumentalizam quatro dimensões da ligação autocrática. Elas são: comércio, migração, ligações diplomáticas e proximidade geográfica. Já a variável status é representada por (0) se um regime democrático sobreviveu ou (1) se quebrou e se tornou democrático. Entre os controles, estão PIB per capita, crescimento do PIB e uma variável *dummy* para indicar se era um período de Guerra Fria ou não. Os dados dos autores incluem 250 regimes autocráticos, entre 1949 e 2008. A variável tempo é ano, assim como nos exemplos anteriores.

A hipótese nula, de que as ligações autocráticas internacionais não tinham efeito na duração de regimes autocráticos, pôde ser rejeitada. Assim, o risco de sofrer uma transformação institucional é dirimido à medida que o grau de relações entre estados autocráticos se fortalecem. Levando em consideração que o trabalho utilizou o PIB per capita e o crescimento do PIB como variáveis de controle, os autores utilizaram o modelo de Cox estendido, como na fórmula (5).

O segundo (2) exemplo é o trabalho de Shipan e Shannon (2003). Os autores analisam o tempo que leva entre a indicação de um candidato à Suprema Corte americana e a confirmação do candidato pelo senado norte-americano. A variável que indica status, portanto, é: 0 quando a confirmação não ocorre e 1 quando ela acontece. O tempo, diferente dos trabalhos vistos anteriormente nessa seção, é mensurado em dias. Fazendo uso

18 De acordo com o autor, existem três tipos de diplomas possíveis no ensino secundário francês: o tecnológico, o profissional e o geral. O autor afirma que os detentores de diploma geral têm tendência maior a concluir o ensino superior, já que o diploma geral do ensino secundário teria baixo valor no mercado de trabalho.

da teoria espacial, a hipótese principal do trabalho é o de que quanto maior for a distância ideológica entre o presidente e o senado, maior também vai ser o tempo entre indicação e confirmação.

Apesar de a distância ideológica ser a principal hipótese que os autores pretendem testar, eles também consideram características do candidato, como: se serve ou já serviu no senado ou se tem experiência prévia com tribunais a nível estadual ou federal. Os resultados encontrados por Shipan e Shannon confirmam a hipótese levantada. Maior distância ideológica aumenta o tempo até a confirmação da indicação feita pelo presidente. Os autores também acharam resultados estatisticamente significantes para a experiência do candidato no senado. E como as variáveis explicativas não variam com o tempo, o modelo utilizado foi o Cox tradicional.

Finalmente, o terceiro (3) exemplo é o trabalho de Crespo-Tenorio, Jensen e Rosas (2014). O objetivo dos autores é compreender o impacto que crises bancárias têm sobre a chance de sobrevivência de incumbentes. A hipótese principal dos autores é a de que as crises bancárias terão impactos negativos sobre os incumbentes em regimes democráticos. O impacto tende a ser ainda maior em países que possuem poucos atores de veto. O evento de interesse, portanto, é a falha na manutenção de um mesmo partido no poder. De forma mais direta: quando eleições levam um outro partido ao poder. O tempo aqui volta a ser mensurado anualmente e a hipótese dos autores também se confirma.

Por fim, o quadro sumariza aspectos dos três estudos vistos aqui.

Quadro 1 – Resumo dos estudos

Disciplina	Variável status	Variável tempo	Variação no tempo	Modelo
Medicina	Desenvolver doença	Anos	Sim	Cox estendido
Economia	Abandonar faculdade	Anos	Não	Cox tradicional
Ciência Política (1)	Queda de regime	Anos	Sim	Cox estendido
Ciência Política (2)	Confirmar indicação	Dias	Não	Cox tradicional
Ciência Política (3)	Derrota do incumbente	Anos	Sim	Cox estendido

Fonte: autor.

Quatro dos cinco estudos analisados aqui utilizam o ano como mensuração do tempo. Três deles possuem variáveis independentes ou de controles cujo valores variam no tempo. Isso implica dizer que eles utilizaram o modelo de Cox estendido e, assim, construíram bancos com a estrutura de contagem, como visto no exemplo da tabela 2. Nesses três casos, a construção estatística do modelo segue a fórmula (5). O estudo de economia e de ciência política sobre confirmação das indicações à Suprema Corte, que não possuíam variável tempo-dependente, utilizaram o modelo tradicional de Cox. Dessa forma, seus bancos devem possuir uma estrutura similar à exemplificada na tabela 1, sendo realizado via fórmula (4).

5. Como fazer? Um passo a passo computacional[19]

Após a abordagem teórica dos modelos de sobrevivência[20] e depois de apresentar três exemplos de utilização desses modelos, chegou a hora de um exemplo prático. Vamos seguir o passo a passo computacional com enfoque no *software* R[21]. O exemplo que utilizarei trata da sobrevivência de regimes democráticos e será elaborado a partir do banco de dados do projeto *Varieties of Democracy* (V-Dem) [22]. A análise focará na América Latina e terá um corte temporal de 1900 a 2016, e trabalhará com seguinte questão de pesquisa: qual o impacto da institucionalização de partidos políticos sobre a sobrevivência de regimes democráticos?

Duas variáveis são obrigatórias em qualquer regressão com modelos de sobrevivência: o tempo de sobrevivência e o evento (status). Esta última é, necessariamente, uma variável dicotômica, em que 1 indica a ocorrência do evento estudado e, por exclusão, 0 significa a manutenção do status. Já a variável tempo indica o tempo desde ou o início do estudo (ou inclusão da observação) até que uma das duas situações ocorra: (1) a observação experimentar o evento ou (2) o período do estudo chegar ao fim. Qualquer mensuração de tempo (dias, meses ou anos, por exemplo) pode ser utilizada. Aqui, utilizaremos ano.

Já temos, portanto, clara a construção da variável dependente. Seguindo a explicação de Guo (2010), temos as duas informações necessárias: a de caráter dicotômico (a democracia sobreviveu (0) ou sofreu revés institucional (1)); e a variável contínua representando tempo. No entanto, o modelo possui variáveis tempo-dependentes. Assim, o banco é construído no formato de contagem, com uma variável indicando o início do período e outra indicando o fim, para a elaboração de uma regressão do modelo de Cox estendido.

Finalmente, a última característica do modelo é a de que os países podem entrar e sair do grupo de risco. Isso acontece porque uma democracia, ao se tornar uma ditadura, sai do grupo de risco (e, consequentemente, do banco de dados). No entanto, não existe impedimento para que esse mesmo país volte a ser uma democracia. Caso isso aconteça, ele volta ao grupo de risco formado por países democráticos que podem, eventualmente, se tornar ditaduras (voltando também ao banco de dados).

Dessa forma, o desenho final é:

Quadro 2 – Desenho do modelo

Modelo	Evento recorrente	Tempo	Evento	Unidade
Cox estendido	Sim	Ano	Quebra de regime	País-ano

Fonte: autor.

19 Com o objetivo de ser transparente e replicável, o script da análise está disponível no OSF (https://osf.io/xp6mg/) e no GitLab. Como o único banco utilizado foi o do V-Dem, recomendo que o download seja feito diretamente na página da iniciativa. Ver nota 21.

20 Para uma visão teórica mais ampla de análise de sobrevivência, é possível aproveitar o material disponível on-line de uma disciplina pela professora Ronghui Xu em: http://www.math.ucsd.edu/~rxu/math284/.

21 O software R não é o único que realiza tal análise. Existe, por exemplo, vasto material sobre como realizar a análise de sobrevivência em STATA ou SPSS.

22 A versão do V-Dem utilizada neste trabalho é a 10, lançada em março de 2020. O download pode ser feito em: https://www.v-dem.net/en/data/data/v-dem-dataset/

Tendo compreendido a estrutura do modelo, apresento agora as variáveis incluídas. Elas são:

Quadro 3 – Variáveis

Variável	No banco
Regime político	democracy
Início	start
Fim	stop
Institucionalização partidária	v2xps_party
Presidencialismo	v2xnp_pres
PIB per capita (ln)	e_migdppcln

Fonte: autor.

As três primeiras linhas são os componentes que formam a variável dependente. O regime político é a variável de status, indicando quando um país é democrático (0) ou quando sofre o evento (1). Ela é elaborada pelo pesquisador a partir de uma variável já existente no V-Dem (v2x_regime). A variável original é construída a partir da medida contínua de democracia do V-Dem. Lührmann, Lindberg e Tannenberg (2017), antes de construir quatro categorias de regime, primeiro os divide em dois grandes grupos: democracias e não-democracias. Essa é a divisão utilizada aqui para a variável dicotômica de status.

Como estou lidando com variáveis que mudam no tempo, preciso construir também as variáveis início e fim. O início indica o começo do intervalo de tempo a partir do qual a variável assume um determinado valor; por sua vez, o fim indica justamente que o intervalo chegou ao fim, pois houve mais uma mudança de valor. Como a variável tempo-dependente deste trabalho muda anualmente, a estrutura fica parecida com a de um banco com dados em painel.

Antes de iniciar a seção prática do trabalho, uma última observação sobre o exemplo utilizado faz-se necessária. O número de eventos por variável (*events per variable,* EPV) é baixo. Entre 1900 e 2016, a região experimentou 14 eventos, tendo três variáveis independentes no modelo. Isso se explica pelo fato da região ter passado grande parcela desse período como regimes ditatoriais e, portanto, fora do grupo de risco[23]. Como consequência, os resultados poderiam ser enviesados. No entanto, por se tratar de um exercício, não me preocuparei com isso aqui[24].

Aplicação no R

Finalmente com o R aberto e com o banco do V-Dem já no ambiente do software, começamos instalando os pacotes que precisaremos utilizar e, em seguida, fazendo o upload da biblioteca virtual. O survival[25] é o pacote mais completo para elaboração de modelos de sobrevivência no R. É com ele que faremos nossa regressão de Cox e também a análise dos resíduos de Schoenfeld.

23 Das 2.337 observações de país-ano, 1.713 foram sob regimes autoritários e 596 sob regimes democráticos.

24 De acordo com Peduzzi et al. (1995), um valor prudente de EPV é 10. Ver: Concato et al. (1995) e Peduzzi (1995).

25 Para uma visão mais panorâmica do pacote, ver o documento elaborado pelo seu idealizador, Terry Thernau, em: https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf.

```
## Pacotes e opções ####
options(scipen = 999)
install.packages("survival")
library("survival")
```

A primeira linha indica a opção scipen = 999. Ela não é obrigatória, apenas evita que o software ofereça outputs em formato de notação científica. Isso é comum quando os valores são muito baixos, por exemplo. A segunda linha é o comando para instalar o pacote survival. Ele permite realizar praticamente todos os modelos dentro da família de análise de sobrevivência. A última linha faz o upload do pacote.

Tratando os dados do V-Dem, podemos fazer um subgrupo para manter apenas aquelas variáveis que serão utilizadas. Também é possível fazer o recorte temporal, além de selecionar especificamente os países da América Latina. Neste exemplo, também é preciso criar a variável *dummy* que representa o evento. Tudo isso pode ser visto no bloco a seguir:

```
## Selecionando as variáveis que farão parte do banco ##
vdem <- `V-Dem-CY-Full+Others-v10`[ , c("country_name", "country_text_id", "country_id",
                                        "year", "v2x_polyarchy", "v2x_regime", "v2xps_party",
                                        "v2xnp_pres", "e_migdppcln", "e_regionpol")]


vdem <- vdem[ which(vdem$year> '1899'), ]  ## Selecinando anos > 1899 ##
vdem <- vdem[ which(vdem$year< '2017'), ]  ## Selecionando anos < 2017 ##
vdem <- vdem[ which(vdem$e_regionpol== '2'), ]  ## Selecionando a América Latina ##
vdem$democracy <- ifelse(vdem$v2x_regime> 1, 0, 1) ## Criando variável 'democracy' ##
colnames(vdem)[1] <- "country"  ## Renomeando a coluna "country_name" como "country" ##
```

Com isso feito, agora é a hora de estruturar o banco no formato de contagem para que ele esteja apto a ser utilizado em um modelo de Cox estendido. Nesse momento é preciso que o pesquisador esteja bastante atento aos dados. Como já dito na subseção anterior, esse trabalho lida com eventos recorrentes. Então o país precisará ser excluído do banco sempre que fizer a transição para um regime ditatorial, podendo retornar caso volte a ser uma democracia.

Vamos olhar um excerto da Argentina para facilitar a compreensão. O bloco mostra cada passo especificamente. De acordo com os dados do V-Dem, entre 1920 e 1929 a Argentina era uma democracia e, portanto, estava no banco. Sofreu o evento em 1930 e, por isso, foi excluído a partir de 1931. Em 1964 volta a ser uma democracia por um breve período e, dessa forma, volta aos grupos de países em risco. Mas já sofre uma nova falha em 1966 e é excluído da análise a partir de 1967[26]. Em 1974 volta a ser democrático e a figurar no banco de dados.

Com isso feito, é hora de criar as variáveis de início e fim:

```
## Criando a variável 'start' ##
vdem$start <- as.numeric(as.factor(vdem$year))
## Criando a variável 'stop' ##
vdem$stop <- vdem$start + 1
```

26 O país sempre é excluído a partir do ano subsequente à falha, pois a observação do evento – por motivos óbvios – precisa estar no banco. Se o país deixar de ser uma democracia em um dado ano, mas já retornar ao estado inicial no ano seguinte, nenhuma alteração deve ser feita.

Depois dos ajustes, é possível usar o comando head() para olhar como ficaram as primeiras linhas do banco.

```
> head(vdem)
    country country_text_id country_id year democracy start stop v2x_polyarchy v2x_regime
207  Mexico             MEX          3 1995         0    73   74         0.514          2
208  Mexico             MEX          3 1996         0    74   75         0.528          2
209  Mexico             MEX          3 1997         0    75   76         0.546          2
210  Mexico             MEX          3 1998         0    76   77         0.596          2
211  Mexico             MEX          3 1999         0    77   78         0.606          2
212  Mexico             MEX          3 2000         0    78   79         0.619          2
    v2xps_party v2xnp_pres e_migdppcln e_regionpol
207       0.869      0.530       9.116           2
208       0.879      0.404       9.121           2
209       0.894      0.381       9.177           2
210       0.894      0.381       9.232           2
211       0.894      0.381       9.249           2
212       0.918      0.211       9.336           2
>
```

Finalmente, o comando para rodar a regressão de Cox:

```
## Regressao de Cox ####
modelo <- coxph(Surv(start, stop, democracy) ~ v2xps_party + v2xnp_pres + e_migdppcln,
                data=vdem)
modelo
```

A função coxph faz parte do pacote survival e indica justamente o modelo de riscos proporcionais de Cox. A partir de Surv(), é possível combinar os dois componentes necessários da variável dependente: tempo e evento. Como o modelo é estendido, ele segue a seguinte lógica: Surv(início,fim,status). No entanto, caso não houvesse uma variável tempo-dependente e o banco fosse elaborado da maneira clássica, seria Surv(tempo,status).

Resultados e interpretação

Depois de construir as variáveis necessárias para o nosso exemplo, ver uma amostra de como ficou o banco de dados e rodar o modelo de sobrevivência, agora é hora de olhar para os resultados encontrados e interpretá-los. O output gerado diretamente pelo comando do último bloco é:

```
Call:
coxph(formula = Surv(start, stop, democracy) ~ v2xps_party +
    v2xnp_pres + e_migdppcln, data = vdem)

                coef  exp(coef)   se(coef)      z       p
v2xps_party   1.5569     4.7442     3.5509  0.438 0.66106
v2xnp_pres   10.1272 25014.0175     3.1730  3.192 0.00141
e_migdppcln  -0.6959     0.4986     1.0544 -0.660 0.50925

Likelihood ratio test=22.99  on 3 df, p=0.00004055
n= 609, number of events= 14
   (1 observation deleted due to missingness)
```

O resultado vem com cinco informações: o coeficiente, o exponencial do coeficiente, erro padrão, a estatística z e o p-valor. Olhando para esta última coluna, vemos que apenas o índice de presidencialismo (v2xnp_pres) é estatisticamente significativo ao nível de 5% (p<0,05).

Os coeficientes são estimados por meio da verossimilhança parcial e sua intepretação direta costuma ser difícil. No entanto, existe uma forma bastante intuitiva de olhar para eles. Valores positivos indicam que as variáveis levam ao aumento no risco do evento; enquanto um valor negativo indica variável que ajuda na redução de risco. Se ignoramos o grau de significância e nos atermos apenas aos coeficientes, podemos afirmar que o PIB per capita atua como uma espécie de protetor de regimes democráticos. Em contrapartida, níveis altos de presidencialismo e de institucio-

nalização partidária podem prejudicar a estabilidade das democracias.

O output do R ainda apresenta o exponencial do coeficiente, que pode ser interpretado como a razão de risco. Valores entre 0 e 1 indicam proteção; acima de 1, maior risco. Ele ainda pode ser lido como percentual ao se aplicar a simples fórmula (1-exp(-coef))*100. Assim, voltando para o exemplo, cada unidade acrescida no valor do logaritmo natural do PIB per capita reduz a chance de quebra da democracia em 50,1%.

Resíduos de Schoenfeld

A proporcionalidade dos riscos é um pressuposto válido para o modelo de Cox tradicional, ou simples, e para o modelo estendido (BATISTA, 2017); e a interpretação dos resíduos de Schoenfeld ocorre da mesma forma tanto para a abordagem simples quanto para a que é formulado via processo de contagem (CARVALHO ET AL., 2011). A primeira etapa da análise é observar se há alguma correlação entre os coeficientes e o tempo. Para fazer isso no R, se utiliza a função cox.zph, também do pacote survival.

```
## Residuos de Schoenfeld ####
res.modelo <- cox.zph(modelo, transform = "identity")
res.modelo
```

O output gerado pelo bloco anterior é:

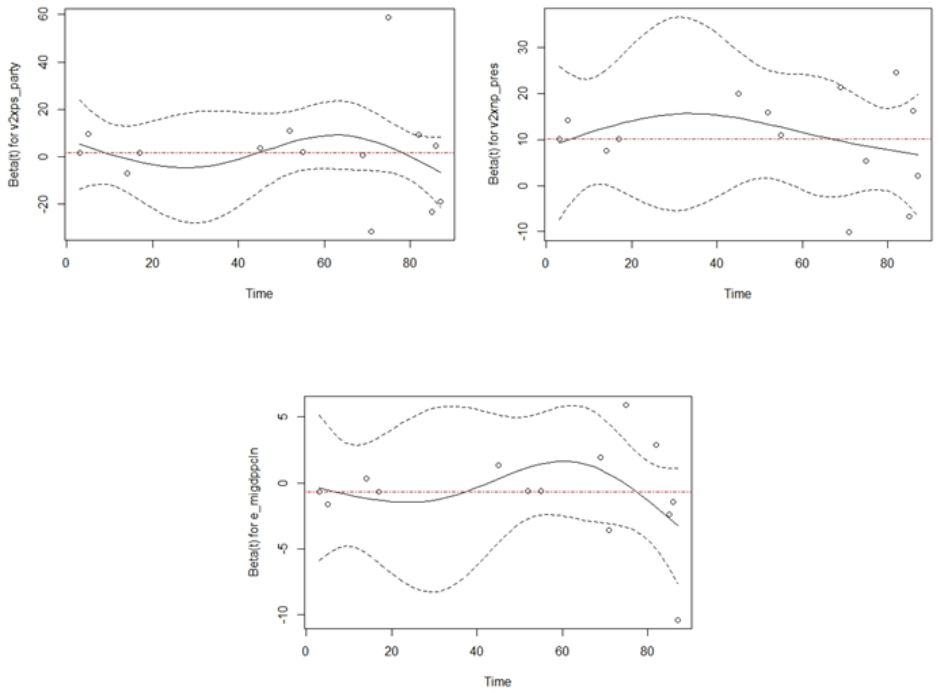
```
> res.modelo
              chisq df    p
v2xps_party 0.0137  1 0.91
v2xnp_pres  0.9335  1 0.33
e_migdppcln 0.2364  1 0.63
GLOBAL      2.0072  3 0.57
```

Olhando o bloco acima, é possível notar que nenhuma variável apresentou uma correlação significativa entre seus coeficientes e o tempo. Assim, não se exclui a hipótese nula que afirma não haver relação entre tempo e variáveis. No entanto, gráficos de dispersão permitem um formato visual para observar se o pressuposto dos riscos proporcionais se mantém. Será preciso criar um gráfico para quantas variáveis existirem no modelo. Neste caso, tenho três variáveis. O bloco a seguir mostra como gerar as figuras.

```
plot(res.modelo[1])
abline(h=modelo$coefficients[1], lty=4, col=2) ## Institucioalização (partidos) ##
plot(res.modelo[2])
abline(h=modelo$coefficients[2], lty=4, col=2) ## Presidencialismo ##
plot(res.modelo[3])
abline(h=modelo$coefficients[3], lty=4, col=2) ## Ln PIB per capita ##
```

A primeira linha de cada parâmetro constrói o gráfico e a segunda linha adiciona uma linha horizontal vermelha que representa os coeficientes de cada uma das variáveis independentes. Os três gráficos gerados são:

Figura 1 – Resíduos de Schoenfeld



Fonte: autor.

Se essa linha, assim como a linha suavizada, estiver dentro do intervalo de confiança, é possível atestar que o pressuposto foi observado. Nos três casos podemos afirmar que os parâmetros se encontram dentro dos intervalos de confiança. Além disso, não existe uma forte tendência positiva ou negativa.

6. Considerações finais

Apesar de ser uma técnica com um grande leque de possibilidades para problemas de pesquisa dentro das ciências sociais, a análise de sobrevivência ainda não decolou como uma ferramenta popular na realidade das sociais brasileira. Portanto, este trabalho tem três objetivos que se complementam: (1) difundir a técnica dentro da grande área das ciências sociais; (2) servir como material introdutório para aqueles que enxergaram valor na técnica para o desenvolvimento de suas pesquisas; e (3) apresentar um passo a passo de sua aplicação no R.

Para tanto, este trabalho apresentou um esforço para mostrar a parte teórica da análise de sobrevivência, junto com um pequeno dicionário de termos comuns para a técnica, para que seja possível compreender os diferentes tipos de desenhos de pesquisa que podem se valer da técnica. Em termos práticos, apresentei uma seção "mãos na massa". Tomando o software R como base, mostrei o passo a passo de como elaborar um modelo de duração, desde a construção, passando pelo modelo propriamente dito, até a análise dos resíduos.

REFERÊNCIAS BIBLIOGRÁFICAS

ALLISON, Paul. (1984), *Event History Analysis: regression for longitudinal event data*. Newbury Park, CA: SAGE Publications Inc.

ATA, Nihal; SÖZER, M. Tekin. (2007), "Cox Regression Models with Nonparametrical Hazard Applied to Lung Cancer Survival Data". *Hacettepe Journal of Mathematics and Statistics*, vol. 3, nº 2: 157-167.

BATISTA, Mariana. (2017), *A Difusão da Lei de Acesso à Informação nos Municípios Brasileiros: fatores internos e externos*. Brasília, DF: Enap.

BOX-STEFFENSMEIER, Janet M.; JONES, Bradford S. (2004), *Event History Modeling: a guide for social scientists*. New York, NY: Cambridge University Press.

BOX-STEFFENSMEIER, Janet M.; ZORN, Christopher J. W. (2001), "Duration Models and Proportional Hazards in Political Science". *American Journal of Political Science*, vol. 45, nº 4: 972-988.

BIESZK-STOLORZ, Beata; MARKOWICZ, Iwona (2015), "Influence of Unemployment Benefit on Duration of Registered Unemployment Spells". *Equilibrium. Quarterly Journal of Economics and Economic Policy*, vol. 10, nº 3: 167-183.

BORUCKA, Jadwiga. (2013), "Extensions of Cox Model for Non-Proportional Hazards Purpose". Disponível em: https://www.lexjansen.com/phuse/2013/sp/SP07_ppt.pdf. Acessado em: 10/06/2020.

CARVALHO, Marilia Sá; ANDREOZZI, Valeska Lima; CODEÇO, Claudia Torres; CAMPOS, Dayse Pereira; BARBOSA, Maria Tereza; SHIMAKURA, Silvia Emiko. (2011), *Análise de Sobrevivência: Teoria e aplicações à saúde*. 2ª ed. Rio de Janeiro: Editora Fiocruz.

CHEIBUB, Jose Antonio; GANDHI, Jennifer; VREELAND, James R. (2010), "Democracy and Dictatorship revisited". *Public Choice*, vol. 143, nº 2-1: 67-101.

COLLETT, D. (1994), *Modelling Survival Data in Medical Research*. Springer-Science +Business Media.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. (2006), *Análise de Sobrevivência Aplicada*. São Paulo, SP: Blucher.

CONCATO, John; PEDUZZI, Peter; HOLFORD, Theodore R.; FEINSTEIN, Alvan R. (1995), "Importance of Event per Independent Variable in Proportional Hazard Analysis I. Background, Goals, and General Strategy". *Journal of Clinical Epidemiology*, vol. 48, nº 12: 1495-1501.

COX, David. (1972), "Regression Models and Life-Tables". *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, nº 2: 187-220.

CRAWLEY, Michael J. (2015) *Statistics: An introduction using R*. 2nd ed. West Sussex, UK: John Wiley & Sons.

CRESPO-TENORIO, Adriana; JENSEN, Nathan M.; ROSAS, Guillermo. "Political Liabilities: surviving banking crises". *Comparative Political Studies*, vol. 47, nº 7: 1047-1074.

DIERMEIER, Daniel; STEVENSON, Randy T. (1999), "Cabinet Survival and Competing Risks". *American Journal of Political Science*, vol. 43, nº 4: 1051-1068.

GRUBE-CAVERS, Annelise; PATTERSON, Zachary. (2015), "Urban rapid rail transit and gentrification in Canadian urban centres: a survival analysis approach". *Urban Studies*, vol. 52, nº 1: 178-194.

GUJARATI, Damodar. (2011), *Econometrics by example*. Palgrave Macmillan.

GUJARATI, Damodar N.; PORTER, Dawn C. (2011), *Econometria Básica* 5ª ed., Porto Alegre, RS: AMGH.

GURY, Nicolas. (2011), "Dropping Out of Higher Education in France: a micro-economic approach using survival analysis". *Education Economics*, vol. 19, nº 1: 51-64.

GUO, Shenyang. (2010), *Survival Analysis*. New York, NY: Oxford University Press.

HARRELL Jr.; Frank E. (2015), *Regression Modeling Strategies: with application to linear models, logistic and ordinal regression, and survival analysis*. Cham, Switzerland: Springer.

JONES, Bradford S.; BRANTON, Regina P. (2005), "Beyond Logit and Probit: Cox duration models of single, repeating, and competing events for state policy adoption". *State Politics and Policy Quarterly*, vol. 5, nº 4: 420-443.

KELLSTEDT, Paul M.; WHITTEN, Guy D. (2018), *The Fundamentals of Political Science Research*. 3rd ed. New York, NY: Cambridge University Press.

KHAMBATY, Tasneem; CALLAHAN, Christopher M.; STEWART, Jesse C. (2018), "Effect of collaborative depression treatment on risk of diabetes: a 9-year follow-up of the IMPACT randomized controlled trial". *PloS one*, vol. 13, nº 8.

KLEINBAUM, David G.; KLEIN, Mitchel. (2005), *Survival Analysis: a self-learning text*. 2ª ed. New York, NY: Springer.

LEE, Elisa T.; WANG, John Weny. (2003), *Statistical Methods for Survival Data Analysis*. 3rd ed. Hoboken, NJ: Wiley.

LEE, Ming-Chang. (2014), "Business Bankruptcy Prediction Based on Survival Analysis Approach". *International Journal of Computer Science & Information Technology*, vol. 6, nº 2: 103-119.

LIN, D. Y.; WEI, L. J.; YING, Z. (1993), "Checking the Cox model with cumulative sums of martingale-based residuals". *Biometrika*, vol. 80, nº 3: 557-572.

LINS, Rodrigo; FIGUEIREDO FILHO, Dalson; ROCHA, Enivaldo. (2017), "Não temos tempo a perder: uma introdução à análise de sobrevivência". *Revista Política Hoje*, vol. 26, nº 1: 279-298.

LÜHRMANN, Anna; LINDBERG, Staffan I.; TANNENBERG, Marcus. (2017), "Regimes in the World (RIW): a robust regime type measure based on V-Dem". *V-Dem Working Paper*, Series 2017:47.

MACHIN, David; CHEUNG, Yin Bun; PARMAR, Mahesh K. B. (2006), *Survival Analysis: a practical approach*. 2nd ed. West Sussex, UK: John Wiley & Sons.

PAPPAS, Vasileios; ONGENA, Steven; IZZELDIN, Marwan; FUERTES, Ana-Maria (2017), "A Survival Analysis of Islamic and Conventional Banks". *Journal of Financial Services Research*, vol. 51: 221-256.

PEDUZZI, Peter; CONCATO, John; FEINSTEIN, Alvan R.; HOLFORD, Theodore R. (1995), "Importance of Event per Independent Variable in Proportional Hazard Analysis II. Accuracy and Precision of Regression Estimates". *Journal of Clinical Epidemiology*, vol. 48, nº 12: 1503-1510.

SCHOENFELD, David. (1982), "Partial Residuals for the Proportional Hazard Regression Model". *Biometrika*, vol. 69, nº1: 239-241.

SHIPAN, Charles R.; SHANNON, Megan L. (2003), "Delaying Justice(s): A duration analysis of Supreme Court Confirmations". *American Journal of Political Science*, vol. 47, nº 4: 654-668.

SOMER-TOPCU, Zeynep; WILLIAMS, Laron K. (2008), "Survival of the Fittest? Cabinet duration in postcommunist Europe". *Comparative Politics*, vol. 40, nº 3: 313-329.

SVOLIK, Milan. (2008), "Authoritarian Reversals and Democratic Consolidation". *American Political Science Review*, vol. 102, nº 2: 153-168.

TANSEY, Oisín; KOEHLER, Kevin; SCHMOTZ, Alexander. (2017), "Ties to the Rest: autocratic linkage and regime survival". *Comparative Politics Studies*, vol. 50, nº 9: 1221-1254.

WALTERS, Stephen J. (2009), *What is a Cox Model?* 2nd ed. Newmark, UK: Hayward Medical Communications.

All kinds of time: Cox regression in political science

Rodrigo Lins - Universidade Federal de Sergipe

Resumo

Este artigo tem como objetivo apresentar o modelo de riscos proporcional de Cox. Para tanto, será feito um debate teórico geral dos modelos de sobrevivência, assim como a apresentação de termos comuns para a técnica. Além disso, será feita uma apresentação passo a passo de como elaborar um modelo de Cox no software R. Espera-se que este trabalho ajude na difusão da técnica dentro das ciências sociais.

Palavras-chave: Análise de Sobrevivência; Modelo de Cox; R.

Abstract

This article presents the Cox Proportional-Hazards Model. To do so, a general theoretical debate on survival models is made, as well as a presentation of common terms for the technique. In addition, there is a step-by-step presentation of how to develop a Cox model in R software. It is hoped that this work will help to disseminate the technique within the social sciences.

Keywords: Survival Analysis; Cox Model; R.

1. Introduction[1]

Survival analysis is a technique with several names: event history analysis, duration model, and time-to-event are just a few of the names given to this statistical tool. Despite the name variation, all have a key feature: time. Therefore, it is no surprise that survival models are concerned with not only "if" but also "when" a given event[2] occurs (GUO, 2010). That is, the researcher's interest is the observation of a qualitative change in the dependent variable in a specific point in time (ALLISON, 1984).

Within the possible approaches for a survival analysis, the Cox proportional hazards model is highlighted (BOX-STEFFENSMEIER & JONES, 2004). Conceptualized as a semiparametric model, it is very malleable. It does not take on a statistical distribution. It only assumes that the action of the independent variables on the hazard is multiplicative (CARVALHO et al., 2011), that is, for each unit increase in the explanatory variable, the hazard is multiplied by its exponentiated coefficient (ALLISON, 1984).

And how goes the use of survival analysis – and more specifically, of the Cox model - in Brazilian political science? It is noticeable that its use has not taken off. While analyzing abstracts of available papers in the archives of the five main Brazilian political science journals[3], there was only one result that indicated the use of survival analysis. Consequently, this work aims to demonstrate the technique to the larger audience of the social sciences so that it more widely considered.

The remainder of this paper is divided thusly: first I focus on the Cox model, with a more theoretical approach; then, I present a brief glossary of common terms in survival analysis, but that are not so widely used in other statistical techniques. The following section shows possible applications in different areas of knowledge. Finally, in the last section, I demonstrate a step-by-step process of how to conduct a survival analysis in R and interpret its results.

2. Cox model: what is it, what is it for, when to use it

What is the Cox proportional hazards model? Beforehand, it is important to understand that survival analysis modelling can be divided into two groups: parametric and semiparametric[4]. Parametric is the one that presuppose a given distribution, such as the Weibull model or its exponential format (BOX-STEFFENSMEIER & JONES, 2004). While the latter expects that the hazard is constant overtime, Weibull's allows for the hazard to increase or lower linearly over time (ALLISON, 1984).

The semiparametric model – the Cox model – does not assume any distribution for the baseline hazard function. That would be a feature that could put it in a non-parametric group. Nevertheless, it is expected that its variables always act with a multiplicative fact on the hazard. That is the feature that would put it in the parametric group. That is why the Cox model is considered a semiparametric model.

Therefore, to answer more directly the question at the beginning of the section, we can affirm, regarding the Cox model: (1) like parametric models, it allows for

1 I would like to thank the anonymous peer reviewers who helped me to substantially improve this paper. All remaining errors are fully my responsibility.

2 In survival analysis literature, the term "event" can also be named "failure". However, since a qualitative change does not always mean an actual failure, and could be a government transition for example, I have opted to use event.

3 DADOS was the only journal with a result, a paper published in 2019. The other four journals included in the search were: the Brazilian Political Science Review; Opinião Pública; Revista Brasileira de Ciências Sociais; and Revista de Sociologia e Política. The search terms were Cox and survival. The search was conducted on 2 June 2020.

4 It is still possible to find non-parametric estimations. However, it is not possible to estimate the effects of explanatory variables in these models. The Kaplan-Meier estimate product, for example, compares the survival curve of two or more groups, not considering the independent variables. See Lee and Wang (2003), chapters four and five, in particular.

the inclusion of explanatory variables; (2) the effect of these variables is multiplicative; and (3) it does not hold assumptions when it comes to the statistical distribution of the baseline hazard function. The Cox model is, consequently, a statistical technique to explore the relationship of a specific event to a series of explanatory variables (WALTERS, 2009).

It was developed by Cox (1972), from whom it inherited the name, and arose as a way to soften the aforementioned assumptions of parametric models. It is called proportional hazards model because for any two observations at any point in time, the ratio between its hazards must be constant. In other words, the ratio of the hazards from observation A to observation B must be the same both on t_0 and on t_{100}, for example (ALLISON, 1984). It can be symbolised in the following equation:

$$\log h(t) = a(t) + b_1 x_1 + b_n x_n \quad (1)$$

The left side of the equation is the natural logarithm of the hazard function. On the other side, the baseline hazard function a(t) can be any function of time. As we have seen, it is specifically because it is unspecified that the Cox model is seen as partially parametric or semiparametric (ALLISON, 1984). In b_1x_1, b represents the coefficient and x, the value of independent variable 1. Equation (1) is used on the type of data seen on Table 1, in which each observation occupies a line, because variables X1 and X_n are constant in time.

The extended Cox model[5]

It is common, when working with longitudinal data, that some of the explanatory variables vary over time. Let's say the GDP per capita is one of the independent variables. It is hard to imagine that it stays constant over the years. Therefore, how to include it in a model? It is precisely for these cases, when there are time-dependent variables (or variables that vary in time) that the extended Cox model must be used. It can be represented in the following way:

$$\log h(t) = a(t) + b_1 x_1 + b_n x_n(t) \quad (2)$$

Here, a(t) continues to be any function of time and variable X1 continues to be constant over time. However, variable X_n does vary as a function of time. For this kind of design, the researcher must construct a database using a counting process, similar to what is seen on Table 2, where a new line is added whenever there are changes in the values of X_n.

The Schoenfeld residuals

There are different ways of testing the fit of a Cox regression model. However, considering that the proportionality of hazards is one of the defining features of the model, it is important to evaluate if this assumption is being met. That is the purpose of Schoenfeld residuals[6]. They give the proportionality of each covariate defined for each individual (CARVALHO et al., 2011). The equation that illustrates the calculation of residuals is:

$$r_{ik} = \delta_i (x_{ik} - a_{ik}) \quad (3)$$

where the occurrence of the event for observation i is symbolized by δ_i. The value of the specific variable is given by x_{ik} and, finally, the weighted average of the hazard covariates in a given time t is a_{ik}.

Although there are different ways of analyzing the residuals, the most intuitive way of understanding the analysis, as argued by Schoenfeld (1982), is that the residuals must be independent of time. Thus, if a correlation between the residuals and time is not statistically significant, that means that the coefficients are not

5 In addition to its extended form, the Cox model can also be stratified. For a more complete overview of the stratified Cox model, see chapter six in Carvalho et al. (2011).

6 There are other ways of analysing the proportionality of hazard. A common way is via martingale residuals (see Lin et al.,1993). There are still the possibilities of deviance residuals and score residuals (see Carvalho et al., 2011).

dependent on time and, consequently, the assumption is met. If it is statistically significant, it indicates that the proportional hazards assumption has not been met.

As the practical section shows, the analysis of Schonfeld residuals can also be done graphically, making its interpretation easier. As it is expected that the correlation is random, the line in the dispersion graph should be horizontal. If there are any trends, positive or negative, the proportionality of hazards has not been met.

In case of not meeting the proportionality of hazards, the first step is to verify if this happens naturally or artificially. Thus, it is important to check for outliers. If the break in the assumption is natural, two approaches are possible: (1) to try a new model, such as the stratified Cox model (ATA & SÖZER, 2007; BORUCKA, 2013); or (2) to add an interaction effect between the variable that does not satisfy the assumption of proportionality and some function of time[7] (BOX-STEFFENSMEIER & ZORN, 2001).

3. Small survival glossary

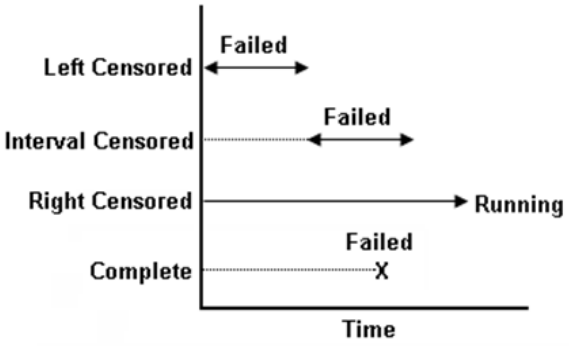
Before moving on, it is important to write out a brief glossary. Given that survival models deal with vocabulary that are not common to other statistical techniques, I believe it is justified to show at least the main terms. The concepts presented in this section are valid not only for Cox models, but for survival analysis in general. The small glossary does not intend to be exhaustive. This means that only the most common and relevant terms are indicated here, to enable the understanding of survival models. First off, these models can deal with two types of incomplete data, censored and truncated.

Censored: An observation is considered censored when its beginning or end are not observed, making it difficult to determine a specific point in time. There are three different types of censoring: (1) right -censored; (2) left-censored; and (3) interval-censored. The first occurs when the observation arrives at the end of the study without experiencing the event ($T>t^-$)[8]. Thus, the researcher only knows if this observation survived up to a given point. But they do not know when or if the event was ever experienced (GUJARATI, 2011). We say that data is censored to the left when we do not know when the event happened, but we do know that it happened before a given moment ($T<t+$). Finally, interval-censored ($t^-<T<t+$) indicates that the event happened within a period of time, but the researcher does not know exactly when (CARVALHO et al., 2011; KLEINBAUM & KLEIN, 2005). Figure 1 illustrates these three types of censoring.

7 Generally, the natural logarithm is used. See Box-Steffensmeier and Zorn (2001).

8 T represents the survival time or time to completion (CARVALHO et al., 2011). It is the interval between the insertion of the observation in the study until the moment it experiences the failure or is censored. t^- indicates that the time of the study is shorter than the survival of the observation and t+ simbolizes that the time of the study surpasses the moment in which the observation was censored.

Figure 1 – Censoring



Source: Lins et al. (2017).

There are still three other types of censoring: type I, type II, and random, common in clinical studies (COLOSIMO & GIOLO, 2006). Type I censoring, similarly to right-censoring, occurs when the study is ended after a pre-established timeframe and, at that point, some observations had not yet gone through the event of interest. Type II censoring occurs when the study is ended after a pre-established number of events takes place. Finally, random censoring is the one where an observation is removed from the study without the event having happened (idem).

Truncated: The data can also be truncated to the right and to the left. Truncation to the right occurs when cases that have gone through the event are selected. This means all observations that went through the event of interest before the final limit of the study. Truncation to the left includes cases that experienced the event after the inferior limit of the study's timeframe (CARVALHO et al., 2011). It is also worth highlighting observations with delayed entries. These are the cases that were included in a database once the research had already started. These cases are usually mistakenly treated as truncated to the left (HARRELL, 2015).

Other terms relevant for comprehending survival models are:

Survival function - S(t): It is the probability that an observation survives longer than a specific time *t*. It is represented by the following equation:

$$S(t) = P(T > t) \tag{4}$$

Hazard function – h(t): The hazard function represents the potential of failure for each unit of time, considering that the observation survived until the time of observation *t*. Unlike the survival function, it is not a probability but a rate, taking any value higher than zero. It can be represented as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{5}$$

Visually, the survival function has declining trend, indicating that the probability of survival decreases as time passes (KLEINBAUM & KLEIN, 2005). The hazard function can be constant or increase/decrease over time. The hazard function should take precedence because regression models in a survival analysis analyze the risk of a given event occurring. This will become clearer as we deal specifically with Cox models.

Partial likelihood: The coefficients of Cox models are estimated based on a partial likelihood[9]. Unlike the

9 A basic presentation of maximum likelihood is given by Crawley (2015): "given the data, and given the model, what values of the parameters of that model make the observed data most likely?" (CRAWLEY, 2015, p. 6). However, because it is considered an "intricate" method (GUJARATI & PORTER, 2011), I will not explore the concepts of likelihood here.

traditional form[10], the baseline hazard function[11] is removed from the calculation for the partial form because it is not parameterized (BOX-STEFFENSMEIER & JONES, 2004). Only the information from the observations that are under risk at each time *t* are considered.

$$PL = \prod_{j=1}^{K} L_j \quad (6)$$

where K is the total number of events in the sample and Lj is the likelihood of failure at time j, given that the observation survived until this point (KLEINBAUM & KLEIN, 2005).

Data structure: Finally, it is also important to talk about the data structure in a survival analysis. The proper construction of the database has an important role when executing the analysis. Four variables are mandatory or expected. The *start* variable and the *end* variable respectively indicate the moment in which the observation entered the research and the moment at which it experienced the event or was censored. In addition, it is advisable to insert the *time* variable, which indicates the period in which the variable was present in the dataset. Simply put, it is the subtraction of the *end* variable by the *start* variable. Lastly, the binary variable *status*, which indicates if the observation experienced the event or not. Within the traditional model, which is the one used in a Cox regression, each observation will occupy a single line within the dataset. The table below illustrates a traditional dataset. It demonstrates a hypothetical study that begins in time 0 and ends in time 30.

Table 1 – Database

ID	Start	End	Time	Status	X1	X2
A	0	30	30	0	1	1
B	0	27	27	1	1	0
C	1	15	14	1	0	0
D	2	30	28	0	1	1
E	3	15	12	1	0	1

Source: the author.

We have five observations and one line for each. The first observation is part of the study since the beginning, given that its start is at 0. It reached time 30 without experiencing the event, since its status is also 0. The third observation started at time 1 of the study, and left at time 15, having experienced a total of 14 units of time in the research. We know that its exit happened because it experienced the event, as indicated by the status variable having the value 1. Observations 1 and 4 are right-censored examples, because they reached the end of the study without having experienced the event. Variables X1 and X2 of Table 1 represent explanatory variables[12] with dichotomous values and that are constant in time.

Further we will see that the dataset represented in Table 1 can be used in traditional Cox regressions. However, if among the explanatory variables there is one that varies in time, the structure must be modified. In that case, the observation will gain a new line every time the value of the independent variable changes. Table 2 illustrates this situation.

10 The general likelihood is used in the calculation of the coefficient for the parametric survival models. See Appendix A in Allison (1984) and Chapter 3 of Box-Steffensmeier and Jones (2004).

11 The baseline function is thus named because in the cases where the independent variables take on the value of 0 the value of the hazard function equals the value of the baseline function (COLLETT, 1994).

12 Explanatory variables are the ones theorized as the cause of the variation in the dependent variable (KELLSTEDT & WHITTEN, 2018).

Table 2 – Database
(Time-dependent variables)

ID	Start	End	Status	X1	X2
A	0	11	0	1	100
A	11	22	1	1	120
B	0	15	0	0	40
B	12	23	0	0	42
B	23	30	0	0	47

Source: the author.

Again, the table indicates a hypothetical study lasting 30 years. Given the status variable, we know that observation 1 experienced the event, while observation 2 is right-censored. Independent variable X1 remains dichotomous and constant in time. But now variable X2 shows changes. And for every change, a new line is inserted. For observation 1, it shows only one value alteration during the whole time it is present in the dataset. Therefore, only one additional line is inserted. For observation 2, there are two changes in the value of X2. Consequently, two new lines are inserted. The break always occurs at the time the alteration happens. Another important difference is that the time variable is no longer a part of the database. It would count each line break as the input of a new observation, nullifying the value every time it happened. Thus, it is excluded from the dataset[13]. This dataset is built through the so-called counting process.

There is still another way of structuring the data, for instance if the researcher is dealing with recurring or multiple events. In this case, once a given observation experiences the event of interest, it must leave the dataset, returning once it goes back to the risk set. In political science, a clear example – which I will present with greater detail in the how-to section – is the study of regime transition. If the event of interest is the change from democratic regime to autocratic, the country will leave the dataset as soon as it experiences the institutional setback and it will remain out of the risk set as long as it is a dictatorship, because it cannot experience the event again. However, should it transition back into democracy, it is placed back in the risk set and re-enters the dataset. The following table is an excerpt of a dataset with information regarding Argentina, built from the data by Cheibub, Gandhi, and Vreeland (2010).

Table 3 – Database
(Recurring events)

ID	Start	End	Status
ARG	1946	1955	1
ARG	1958	1962	1
ARG	1963	1966	1
ARG	1973	1976	1
ARG	1983	2008	0

Source: the author.

13 As a rule, it can be kept. But the researcher needs to be cognizant of the fact that it would count the time of the variable's permanence only for the intervals for which there are no alterations in the variable that has its value altered. For that not to happen, the research must make the alterations "manually". However, as the example section shows, it can be completely excluded because it is not necessary to construct an extended Cox model in R.

Between 1966 and 1973 and again between 1976 and 1983, for example, we can see that Argentina was excluded from the dataset. This was done because during those years, the country was a dictatorship. Thus, it was not at risk of becoming an authoritarian country. Although coups-within-coups happened throughout that period, all events occurred during authoritarian regimes. The table shows that every time an observation experiences the event, a new line is created. But it only returns to the dataset if the initial conditions are re-established. It is also important to highlight that for the years to be shown as they are, the database should also begin at year 0.

Having understood what are censored and truncated date, what are survival and hazard functions, and seen how data must be structured, I move on to the next section, showing Cox regression models.

4. An eclectic model: applications in different fields

Survival analysis models are very eclectic when it comes to the disciplines to which they can be applied. Despite the content being very different when crossing knowledge borders, the technique itself adapts well. To see how it works across different disciplines, I present works from three areas: medicine, economics, and political science. For the latter, since it is the focus of this paper, I will show three examples. I have sought to lay out the hypothesis and the model used to test it. Being the focus of this paper, all used the Cox model. The box at the end of the section summarizes the time and status variables of each study.

Medicine

The terminology of the technique's name itself indicates that it arose in the healthcare field. Specifically, in biomedicine. The original goal was to study patients' survival time between the diagnosis of a given illness until death (GUO, 2010). Consequently, it is unsurprising the elevated number of studies in medicine that make use of the technique (COLOSIMO & GIOLO, 2006; MACHIN, CHEUNG & PARMAR, 2006).

To demonstrate the use of the Cox model in the field, we will use the work of Khambaty, Callahan, and Stewart (2018). The authors start from the assumption that there is a relationship between depression and diabetes. More specifically, they argue that depression could be an independent risk factor for the incidence of diabetes. Thus, they wonder if by treating patients for depression, doctors would be reducing the added risk for diabetes. They tested the impact a collaborative treatment program lasting 12 months had on a group of elderly individuals with depression and initially free of diabetes. The study followed-up for nine years and had 119 patients.

It is possible to see, therefore, the variable that indicates failure: developing the illness. It is instrumentalized as 0 if the patient does not develop the illness and 1 if they do. The independent variable is also binary, indicating if the participant was part of the treatment group (1) or the control group (0). The authors also ran a second model, including some personal variables as controls: age, gender, race/ethnicity, history of hypertension, of smoking, and body mass index (BMI). The results, however, indicate that despite the treatment group indicating a drop in depression levels, when compared to the control, the groups did not diverge regarding the incidence of diabetes. Consequently, the authors argue that treatment for depression is not, on its own, sufficient to lower the risk of patients developing the illness. Since one of the control variables has its values altered over time (BMI), the authors used an extended Cox model, as seen in equation (5).

Economics

Survival analysis in economics may have several applications. From studies on the failure of businesses (LEE, 2014), to studies on the stability of banks (PAPPAS et al., 2017), the duration of unemployment (BIESZKSTOLORZ & MARKOWICZ, 2015), and gentrification processes (GRUBE-CAVERS & PATTERSON, 2015). However, I will focus on the work of Gury (2011). He analyzes the determining factors in the dropout rate of higher education students in France, i.e., the duration time of students in higher education. Thus, the status variable can already be designed: 0 if the students do not abandon higher education and 1 if they do. To conduct the study, Gury (2011) use data from France's Ministry of National Education, to follow the 5383 students that started secondary school in 1989 and later registered in a higher education institution.

With the goal of explaining the students' behaviour, he uses three large groups of independent variables: one related to personal characteristics (gender and formal education of parents, for instance); one related to the observations' secondary education (such as the type of degree obtained[14]); and one related to higher education (if the student declared having some professional project).

To test the hypotheses, Gury (2011) developed three models, going from the simplest (with only the first group of independent variables) to the most complete (with all three groups of explanatory variables). As a result, the author finds a non-linear pattern, indicating that there are two groups of students. The *early leavers*, who abandoned the course in the first two years; and the *late leavers*, who abandoned it from the third year on. Among the ones who experience the event – abandon higher education –, women tend to leave in the first year, while men abandoned university towards the end of the course. Other social economic characteristics, such as the parents' level of formal education, only show effect in the first years of schooling. As years go by, differences tend to disappear. In this study, the variables do not change with time, allowing the author to work with the traditional Cox model, as seen in equation (4).

Political Science

For the third example on using survival models, I demonstrate a study in political science. it is important to highlight that these models have been used in the discipline where logit and probit models would usually be found. Several areas in the discipline are ideal for duration models. For example, they have been used in political regime transition studies (SVOLIK, 2008); in the study of government survival (DIERMEIER & STEVENSON, 1999; SOMER-TOPCU & WILLIAMS, 2008); and in the analysis of public policy diffusion (JONES & BRANTON, 2005; BATISTA, 2017).

The first (1) example we will be looking at is the work by Tansey, Koehler, and Schmotz (2017), who studied the effect of autocratic links and the survival of dictatorships. The hypothesis is presented thusly: the higher the level of international autocratic links, the lower the risk of autocratic regime breakdown. Tansey et al. (2017, p. 2) instrumentalize four dimensions of autocratic linkage. They are: trade, migration, diplomatic ties, and geographic proximity. The status variable is represented by (0) if a democratic regime survived or (1) if it broke down and became democratic. Among the controls are the GPD per capita, GDP growth, and a dummy to indicate whether it was during the Cold War. The authors' data include 250 autocratic regimes between 1949 and 2008. The time variable is year, like in the previous examples.

14 According to the author, there are three types of possible degrees in French secondary education: technical, professional, and general. The author argues that holders of the general degree are more likely to finish higher education since the secondary education general degree is of low value in the labour market.

The null hypothesis, that international autocratic linkages did not have an effect on the duration about autocratic regimes, was rejected. Thus, the risk of suffering an institutional transformation is lessened as the degree of relationships among autocratic states is strengthened. Considering that the work used GDP per capita and GDP growth as control variables, the authors used the extended Cox model as in equation (5).

The second (2) example is the work of Shipan and Shannon (2003). The authors analyzed the time it takes between nominating a candidate to the United States Supreme Court and the confirmation of that candidate by the Senate. The status variable, therefore, is: 0 when the confirmation does not occur and 1 when it does. Time, unlike the studies seen previously in this section, is measured in days. Using spatial theory, the main hypothesis in the study is that the larger the ideological distance between the president and the Senate, the longer will be the time between nomination and confirmation.

Despite ideological distance being the main hypothesis the authors intended to test, they also consider candidate characteristics, such as: if they serve or have served in the Senate or if they had previous experience with state or federal courts. The results found by Shipan and Shannon confirmed the hypothesis. A larger ideological distance increases the time of confirmation of the president's nomination. The authors also find statistically significant results for the candidates experience in the Senate. And since the explanatory variables do not vary with time, the traditional Cox model was used.

Finally, the third (3) example is the work by Crespo-Tenorio, Jensen, and Rosas (2014). The authors' goal is to understand the impact banking crises have on the chance of incumbent survival. The authors' main hypothesis is that banking crises will have a negative impact on incumbents in democratic regimes. That impact tends to be even stronger in countries where there are few veto players. The event of interest, therefore, is failure in maintaining the same party in power. To put it more directly: when elections lead another party to power. Time is again measured annually, and the authors' hypothesis is also confirmed.

Lastly, the box below summarizes the features of the three studies seen here.

Box 1 – Summary of studies

Disciplina	Variável status	Variável tempo	Variação no tempo	Modelo
Medicine	Develop the illness	Years	Yes	Extended Cox
Economics	Abandon university	Years	No	Traditional Cox
Political Science (1)	Regime breakdown	Years	Yes	Extended Cox
Political Science (2)	Nominee confirmed	Days	No	Traditional Cox
Political Science (3)	Incumbent defeat	Years	Yes	Extended Cox

Source: the author.

Four out of the five studies analyzed here used year as a measure of time. Three of them have independent or control variables with values varying over time. This means that they used the extended Cox model and thus constructed datasets with the counting structure, as seen in Table 2. In these three cases, the statistical construction of the model follows equation (5). The economics study and the political science study on the confirmation of nominees to the Supreme Court, which did not have a time-dependent variable, used the

traditional Cox model. Therefore, their datasets should possess a similar structure to the one shown end Table 1 and would have used equation (4).

5. How to – a computational step-by-step guide[15]

After the theoretical approach on survival models[16] and presenting three examples on using those models, it is time for a practical example. We will follow a computational step-by-step using R[17]. The example is about the survival of democratic regimes, elaborated with a database from the *Varieties of Democracy* (V-Dem) project[18]. The analysis will focus on Latin America and with a timeframe between 1900 and 2016, working with the following research question: what is the impact of the institutionalization of political parties on the survival of democratic regimes?

Two variables are mandatory in any regression with survival models: the survival time and the event (status). The latter is, necessarily, a dichotomous variable, in which 1 indicates the occurrence of the event and, by exclusion, 0 means status quo. The time variable indicates the time since the beginning of the study (or inclusion of the observation) until one of two situations occurs: (1) the observation experiences the event or (2) the period of the study reaches an end. Any measure of time (days, months, or years, for example) can be used. Here, we will use year.

We have, therefore, a clear construction of the dependent variable. Following the explanation of Guo (2010), we have the two pieces of information needed: one dichotomous (democracy survived (0) or suffered an institutional setback (1)); and one continuous variable representing time. However, the model has time-dependent variables. Thus, the dataset is constructed in the counting format, with a variable indicating the start of the period and another indicating the end, for the elaboration of an extended Cox regression model.

Finally, the last feature of the model is that countries can enter and exit the risk set. This happens because a democracy, by becoming a dictatorship, leaves the risk set (and, consequently, the database). However, there is no impediment for that same country to become a democracy again. Should this happen, it returns to the risk set comprised of democratic countries that can, eventually, become dictatorships (returning again to the database).

Thus, the final design is:

Box 2 – Design of the model

Model	Recurring event	Time	Event	Unit
Extended Cox	Yes	Year	Regime breakdown	Country-year

Source: the author.

15 Aiming to be transparent and replicable, the script for the analysis is available at OSF (https://osf.io/xp6mg/) and GitLab. Given that the only data set used was V-Dem, I recommend that the download be done directly from their webpage. See footnote 21.

16 For a broader theoretical perspective on survival analysis, one can use the material available online from professor Ronghui Xu's class: http://www.math.ucsd.edu/~rxu/math284/.

17 R is not the only software that can conduct this analysis. There is, for instance, ample material on how to run survival analyses on STATA or SPSS.

18 The V-Dem dataset used here is the version 10, released in March 2020. The download can be made here: https://www.v-dem.net/en/data/data/v-dem-dataset/

Having understood the structure of the model, I now present the variables included. They are:

Box 3 – Variables

Variable	In the dataset
Political regime	democracy
Start	start
End	stop
Party institutionalisation	v2xps_party
Presidentialism	v2xnp_pres
GDP per capita (ln)	e_migdppcln

Source: the author.

The first three lines are the components that form the dependent variable. The political regime is the status variable, indicating when a country is democratic (0) or when it experiences the event (1). It is elaborated by the researcher from a variable already present in V-Dem (v2x_regime). The original variable is constructed from the continuous measure of democracy from V-Dem. Lührmann, Lindberg, and Tannenberg (2017), before constructing the four regime categories, first divide them into two large groups: democratic and non-democratic. That is the division used here for the dichotomous status variable.

As I am dealing with variables that change over time, I also need to construct the variables start and end. The start indicates the beginning of the time interval from which the variable takes on a determined value; the end indicates precisely that the interval has reached an end, because there was another change in value. Since the time-dependent variable in this work changes annually, the structure looks like that of a panel dataset.

Before beginning the practical section of the work, one final observation on the example is needed. The number of events per variable (EPV) is low. Between 1900 and 2016, the region experienced 14 events, with three independent variables in the model. This is due to the fact that the region spent a large part of that period as dictatorships and, therefore, outside of the risk set[19]. Consequently, the results could be biased. Nevertheless, because this is an exercise, it will not be a concern[20].

Application in R

At last, with R open and the V-Dem database already loaded in the environment, we begin by installing the packages we will need and, next, uploading the virtual library. The survival[21] package is the most complete for elaborating survival models in R. That is the one we will use to run our Cox regression and our Schoenfeld residuals analysis.

19 From the 2337 country-year observations, 1713 were under authoritarian regimes and 596 under democractic regimes.

20 According to Peduzzi et al. (1995), a prudente EPV value is 10. See Concato et al. (1995) and Peduzzi (1995).

21 For a broader view of the package, see the document elaborated by its creator, Terry Thernau, at: https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf.

```
## Pacotes e opções ####
options(scipen = 999)
install.packages("survival")
library("survival")
```

The first line indicates the option scipen = 999. It is not mandatory; it simply avoids the software from offering outputs in the format of scientific notation. It is common when values are too low, for example. The second line is the command to install the survival package. It makes possible to run practically all the models within the family of survival analysis. The last line uploads the package.

While treating the V-Dem data, we can make a subgroup to keep only those variables that will be used. It is also possible to choose a timeframe, as well as specifically select the Latin American countries. In this example, it is also necessary to create the dummy variable that represents the event. All of this can be seen below:

```
## Selecionando as variáveis que farão parte do banco ##
vdem <- `V-Dem-CY-Full+Others-v10`[ , c("country_name", "country_text_id", "country_id",
                                        "year", "v2x_polyarchy", "v2x_regime", "v2xps_party",
                                        "v2xnp_pres", "e_migdppcln", "e_regionpol")]


vdem <- vdem[ which(vdem$year> '1899'), ]  ## Selecinando anos > 1899 ##
vdem <- vdem[ which(vdem$year< '2017'), ]  ## Selecionando anos < 2017 ##
vdem <- vdem[ which(vdem$e_regionpol== '2'), ]  ## Selecionando a América Latina ##
vdem$democracy <- ifelse(vdem$v2x_regime> 1, 0, 1) ## Criando variável 'democracy' ##
colnames(vdem)[1] <- "country"  ## Renomeando a coluna "country_name" como "country" ##
```

This being done, it is now time to structure the database in the counting format, so it is possible to use it in an extended Cox model. At this point it is necessary that the researcher pays close attention to the data. As mentioned in the previous subsection, this work deals with recurring events. Therefore, the country will need to be excluded from the database every time it transitions into a dictatorship, with the permission to return in case it goes back to being a democracy.

Let's look at an excerpt of Argentina to make understanding this easier. The code shows each step specifically. According to the V-Dem data, between 1920 and 1929, Argentina was a democracy and, therefore, was in the database. It experienced an event in 1930 and, consequently, was excluded from 1931. In 1964, it goes back to being a democracy for a brief period and thus, returns to the group of countries at risk. But it experiences a new breakdown in 1966 and it is excluded from the analysis from 1967[22]. In 1974, it returns to being a democracy and to the database.

With this done, it is time to create the start and end variables:

```
## Criando a variável 'start' ##
vdem$start <- as.numeric(as.factor(vdem$year))
## Criando a variável 'stop' ##
vdem$stop <- vdem$start + 1
```

22 The country is always excluded from the subsequent year to the failure since the observation of an event – for obvious reasons – needs to be in the database. if the country stops being a democracy in a given year, but already returns to its initial status in the following year, no alteration must be made.

After adjustments, it is possible to use the head()command to see how the first lines in the dataset look.

```
> head(vdem)
    country country_text_id country_id year democracy start stop v2x_polyarchy v2x_regime
207  Mexico             MEX          3 1995         0    73   74         0.514          2
208  Mexico             MEX          3 1996         0    74   75         0.528          2
209  Mexico             MEX          3 1997         0    75   76         0.546          2
210  Mexico             MEX          3 1998         0    76   77         0.596          2
211  Mexico             MEX          3 1999         0    77   78         0.606          2
212  Mexico             MEX          3 2000         0    78   79         0.619          2
    v2xps_party v2xnp_pres e_migdppcln e_regionpol
207       0.869      0.530       9.116           2
208       0.879      0.404       9.121           2
209       0.894      0.381       9.177           2
210       0.894      0.381       9.232           2
211       0.894      0.381       9.249           2
212       0.918      0.211       9.336           2
>
```

Finally, the command to run the Cox regression:

```
## Regressao de Cox ####
modelo <- coxph(Surv(start, stop, democracy) ~ v2xps_party + v2xnp_pres + e_migdppcln,
                data=vdem)
modelo
```

The coxph function is part of the survival package and it precisely results in the Cox proportional hazards model. Using Surv(), it is possible to combine the two components needed in the dependent variable: time and event. As the model is extended, it uses the following logic: Surv(start,end,status). However, if there were no time-dependent variable and the database were elaborated in the traditional way, it would be Surv(time,status).

Results and interpretation

After construction all the necessary variables for our example, seeing a sample of how the dataset looks, and running the survival model, it is now time to look at the results found and interpret them. The output directly generated by the command in the last code section is:

```
Call:
coxph(formula = Surv(start, stop, democracy) ~ v2xps_party +
    v2xnp_pres + e_migdppcln, data = vdem)

                coef  exp(coef)   se(coef)      z       p
v2xps_party   1.5569     4.7442     3.5509  0.438 0.66106
v2xnp_pres   10.1272 25014.0175     3.1730  3.192 0.00141
e_migdppcln  -0.6959     0.4986     1.0544 -0.660 0.50925

Likelihood ratio test=22.99  on 3 df, p=0.00004055
n= 609, number of events= 14
   (1 observation deleted due to missingness)
```

The output comes with five pieces of information: the coefficient, exponentiated coefficient, the standard error, the z-score, and the p-value. Looking at the latter, we see that the presidentialism index (v2xnp_pres) is statistically significant at 5% ($p<0.05$).

The coefficients are estimated by partial likelihood and their direct interpretation can be difficult. Nevertheless, there is a very intuitive way of looking at them. Positive values indicate that the variables lead to an increase in risk of the event, while a negative value indicates that the variable aids in reducing risk. If we ignore the level of significance and stick only with the coefficients, we can state that the GDP per capita acts as a form of shield in democratic regimes. On the other hand, high levels of presidentialism

and party institutionalization can harm the stability of democracies.

R's output also shows the exponentiated coefficient, which can be interpreted as the risk ratio. Values between 0 and 1 indicate protection; above 1, greater risk. It can also be read as a percentage by applying the simple equation (1-exp(coef))*100. Therefore, returning to the example, each unit added in the value of the natural logarithm of the GDP per capita reduces the chance of a breakdown in democracy in 50.1%.

Schoenfeld residuals

The proportional hazards assumption is valid for the traditional or simple Cox model, and for the extended model (BATISTA, 2017); and the interpretation of the Schoenfeld residuals occurs in the same way for both the simple approach and for when it is formulated via the counting process (CARVALHO et al., 2011). the first step in the analysis is to observe if there is any correlation between the coefficients and time. To do that in R, the cox.zph function is used, also in the survival package.

```
## Residuos de Schoenfeld ####
res.modelo <- cox.zph(modelo, transform = "identity")
res.modelo
```

The output generated by the previous code is:

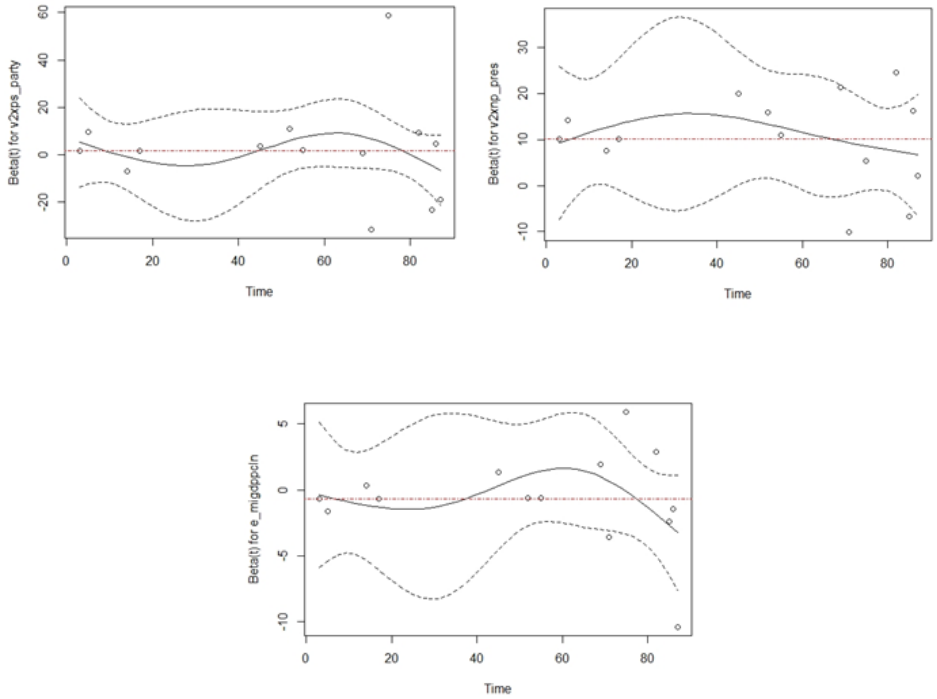
```
> res.modelo
             chisq df    p
v2xps_party 0.0137  1 0.91
v2xnp_pres  0.9335  1 0.33
e_migdppcln 0.2364  1 0.63
GLOBAL      2.0072  3 0.57
```

Looking at the code above, it is possible to see that no variable presented a significant correlation between their coefficients and time. Thus, the null hypothesis that states that there is no relationship between time and the variables is not excluded. However, dispersion graphs allow a visual way to observe if the proportional hazards assumption is upheld. It is necessary to create a graph for however many variables exist in the model. In this case, I have three variables. The following code shows how to generate the figures.

```
plot(res.modelo[1])
abline(h=modelo$coefficients[1], lty=4, col=2) ## Institucioalização (partidos) ##
plot(res.modelo[2])
abline(h=modelo$coefficients[2], lty=4, col=2) ## Presidencialismo ##
plot(res.modelo[3])
abline(h=modelo$coefficients[3], lty=4, col=2) ## Ln PIB per capita ##
```

The first line of each parameter constructs the graph, and the second line adds a horizontal, red line that represents the coefficients of each of the independent variables. The three graphs generated are:

Figure 1 –Schoenfeld Residuals



Source: the author.

If that line, similarly to the softened line, is within the confidence interval, it is possible to affirm that the assumption is upheld. In the three cases we can state that the parameters can be found within the confidence intervals. In addition, there are no strong positive or negative trends.

6. Discussion

Despite being a technique with a wide range of possibilities for research problems within the social sciences, survival analysis has not yet taken off as a popular tool in Brazilian social sciences. Therefore, this paper has three complementary goals: (1) to share the technique within the large field of the social sciences; (2) to serve as introductory material for those who see value in the technique in the development of their research; and (3) to present a step-by-step of its application in R.

To do so, this paper demonstrated the theoretical background of survival analysis, along with a small glossary of common terms for the technique, so that it is possible to understand the different types of research design that might use it. In a practical perspective, I presented a "hands on" section. Using R, I demonstrated a step-by-step on how to elaborate a duration model, from construction, through the model, until the residuals analysis.

BIBLIOGRAPHICAL REFERENCES

ALLISON, Paul. (1984), *Event History Analysis: regression for longitudinal event data*. Newbury Park, CA: SAGE Publications Inc.

ATA, Nihal; SÖZER, M. Tekin. (2007), "Cox Regression Models with Nonparametrical Hazard Applied to Lung Cancer Survival Data". *Hacettepe Journal of Mathematics and Statistics*, vol. 3, nº 2: 157-167.

BATISTA, Mariana. (2017), *A Difusão da Lei de Acesso à Informação nos Municípios Brasileiros: fatores internos e externos*. Brasília, DF: Enap.

BOX-STEFFENSMEIER, Janet M.; JONES, Bradford S. (2004), *Event History Modeling: a guide for social scientists*. New York, NY: Cambridge University Press.

BOX-STEFFENSMEIER, Janet M.; ZORN, Christopher J. W. (2001), "Duration Models and Proportional Hazards in Political Science". *American Journal of Political Science*, vol. 45, nº 4: 972-988.

BIESZK-STOLORZ, Beata; MARKOWICZ, Iwona (2015), "Influence of Unemployment Benefit on Duration of Registered Unemployment Spells". *Equilibrium. Quarterly Journal of Economics and Economic Policy*, vol. 10, nº 3: 167-183.

BORUCKA, Jadwiga. (2013), "Extensions of Cox Model for Non-Proportional Hazards Purpose". Disponível em: https://www.lexjansen.com/phuse/2013/sp/SP07_ppt.pdf. Acessado em: 10/06/2020.

CARVALHO, Marilia Sá; ANDREOZZI, Valeska Lima; CODEÇO, Claudia Torres; CAMPOS, Dayse Pereira; BARBOSA, Maria Tereza; SHIMAKURA, Silvia Emiko. (2011), *Análise de Sobrevivência: Teoria e aplicações à saúde*. 2ª ed. Rio de Janeiro: Editora Fiocruz.

CHEIBUB, Jose Antonio; GANDHI, Jennifer; VREELAND, James R. (2010), "Democracy and Dictatorship revisited". *Public Choice*, vol. 143, nº 2-1: 67-101.

COLLETT, D. (1994), *Modelling Survival Data in Medical Research*. Springer-Science +Business Media.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. (2006), *Análise de Sobrevivência Aplicada*. São Paulo, SP: Blucher.

CONCATO, John; PEDUZZI, Peter; HOLFORD, Theodore R.; FEINSTEIN, Alvan R. (1995), "Importance of Event per Independent Variable in Proportional Hazard Analysis I. Background, Goals, and General Strategy". *Journal of Clinical Epidemiology*, vol. 48, nº 12: 1495-1501.

COX, David. (1972), "Regression Models and Life-Tables". *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, nº 2: 187-220.

CRAWLEY, Michael J. (2015) *Statistics: An introduction using R*. 2nd ed. West Sussex, UK: John Wiley & Sons.

CRESPO-TENORIO, Adriana; JENSEN, Nathan M.; ROSAS, Guillermo. "Political Liabilities: surviving banking crises". *Comparative Political Studies*, vol. 47, nº 7: 1047-1074.

DIERMEIER, Daniel; STEVENSON, Randy T. (1999), "Cabinet Survival and Competing Risks". *American Journal of Political Science*, vol. 43, nº 4: 1051-1068.

GRUBE-CAVERS, Annelise; PATTERSON, Zachary. (2015), "Urban rapid rail transit and gentrification in Canadian urban centres: a survival analysis approach". *Urban Studies*, vol. 52, nº 1: 178-194.

GUJARATI, Damodar. (2011), *Econometrics by example*. Palgrave Macmillan.

GUJARATI, Damodar N.; PORTER, Dawn C. (2011), *Econometria Básica* 5ª ed., Porto Alegre, RS: AMGH.

GURY, Nicolas. (2011), "Dropping Out of Higher Education in France: a micro-economic approach using survival analysis". *Education Economics*, vol. 19, nº 1: 51-64.

GUO, Shenyang. (2010), *Survival Analysis*. New York, NY: Oxford University Press.

HARRELL Jr.; Frank E. (2015), *Regression Modeling Strategies: with application to linear models, logistic and ordinal regression, and survival analysis*. Cham, Switzerland: Springer.

JONES, Bradford S.; BRANTON, Regina P. (2005), "Beyond Logit and Probit: Cox duration models of single, repeating, and competing events for state policy adoption". *State Politics and Policy Quarterly*, vol. 5, nº 4: 420-443.

KELLSTEDT, Paul M.; WHITTEN, Guy D. (2018), *The Fundamentals of Political Science Research*. 3rd ed. New York, NY: Cambridge University Press.

KHAMBATY, Tasneem; CALLAHAN, Christopher M.; STEWART, Jesse C. (2018), "Effect of collaborative depression treatment on risk of diabetes: a 9-year follow-up of the IMPACT randomized controlled trial". *PloS one*, vol. 13, nº 8.

KLEINBAUM, David G.; KLEIN, Mitchel. (2005), *Survival Analysis: a self-learning text*. 2ª ed. New York, NY: Springer.

LEE, Elisa T.; WANG, John Weny. (2003), *Statistical Methods for Survival Data Analysis*. 3rd ed. Hoboken, NJ: Wiley.

LEE, Ming-Chang. (2014), "Business Bankruptcy Prediction Based on Survival Analysis Approach". *International Journal of Computer Science & Information Technology*, vol. 6, nº 2: 103-119.

LIN, D. Y.; WEI, L. J.; YING, Z. (1993), "Checking the Cox model with cumulative sums of martingale-based residuals". *Biometrika*, vol. 80, nº 3: 557-572.

LINS, Rodrigo; FIGUEIREDO FILHO, Dalson; ROCHA, Enivaldo. (2017), "Não temos tempo a perder: uma introdução à análise de sobrevivência". *Revista Política Hoje*, vol. 26, nº 1: 279-298.

LÜHRMANN, Anna; LINDBERG, Staffan I.; TANNENBERG, Marcus. (2017), "Regimes in the World (RIW): a robust regime type measure based on V-Dem". *V-Dem Working Paper*, Series 2017:47.

MACHIN, David; CHEUNG, Yin Bun; PARMAR, Mahesh K. B. (2006), *Survival Analysis: a practical approach*. 2nd ed. West Sussex, UK: John Wiley & Sons.

PAPPAS, Vasileios; ONGENA, Steven; IZZELDIN, Marwan; FUERTES, Ana-Maria (2017), "A Survival Analysis of Islamic and Conventional Banks". *Journal of Financial Services Research*, vol. 51: 221-256.

PEDUZZI, Peter; CONCATO, John; FEINSTEIN, Alvan R.; HOLFORD, Theodore R. (1995), "Importance of Event per Independent Variable in Proportional Hazard Analysis II. Accuracy and Precision of Regression Estimates". *Journal of Clinical Epidemiology*, vol. 48, nº 12: 1503-1510.

SCHOENFELD, David. (1982), "Partial Residuals for the Proportional Hazard Regression Model". *Biometrika*, vol. 69, nº1: 239-241.

SHIPAN, Charles R.; SHANNON, Megan L. (2003), "Delaying Justice(s): A duration analysis of Supreme Court Confirmations". *American Journal of Political Science*, vol. 47, nº 4: 654-668.

SOMER-TOPCU, Zeynep; WILLIAMS, Laron K. (2008), "Survival of the Fittest? Cabinet duration in postcommunist Europe". *Comparative Politics*, vol. 40, nº 3: 313-329.

SVOLIK, Milan. (2008), "Authoritarian Reversals and Democratic Consolidation". *American Political Science Review*, vol. 102, nº 2: 153-168.

TANSEY, Oisín; KOEHLER, Kevin; SCHMOTZ, Alexander. (2017), "Ties to the Rest: autocratic linkage and regime survival". *Comparative Politics Studies*, vol. 50, nº 9: 1221-1254.

WALTERS, Stephen J. (2009), *What is a Cox Model?* 2nd ed. Newmark, UK: Hayward Medical Communications.

Uma introdução à regressão com dados de painel[1]

Rafael Mesquita - Universidade Federal de Pernambuco
Antônio Alves Tôrres Fernandes - Universidade Federal de Pernambuco
Dalson Britto Figueiredo Filho - Universidade Federal de Pernambuco

Resumo

Apesar da crescente oferta de dados em formato de painel, ainda são raros os estudos no Brasil que combinam as dimensões transversal e longitudinal na mesma análise. Por exemplo, em uma amostra de mais de 7 mil artigos publicados entre 2000 e 2018 em periódicos de CPRI, apenas 45 casos citavam técnicas específicas para lidar com observações de unidades espaciais (países, estados, pessoas) repetidas em intervalos regulares do tempo (anos, meses, dias). Diante dos benefícios inferenciais que tal abordagem pode proporcionar e da escassez de pesquisas sobre o tema, este artigo apresenta uma introdução à regressão de painel. Metodologicamente, sintetizamos as principais recomendações da literatura e mostramos a implementação no *R Statistical* com o pacote *plm*, indo desde a seleção de modelos até o tratamento dos dados e apresentação de resultados. Para aumentar o potencial pedagógico do trabalho, disponibilizamos os materiais de replicação, incluindo dados originais e scripts computacionais. Com este artigo esperamos difundir a utilização de análises longitudinais na pesquisa empírica em CPRI no Brasil.

Palavras-chave: Dados em painel; TSCS; CPRI no Brasil; Regressão linear; Metodologia política.

1 Esse trabalho se beneficiou dos comentários dos membros do grupo de Métodos de Pesquisa do Departamento de Ciência Política da Universidade Federal de Pernambuco (DCP – UFPE). Uma versão preliminar foi apresentada sob o título "Regressão com dados de painel: balanço da utilização na CPRI brasileira e estratégias para maior difusão do método", na seção de "Ensino e Pesquisa em Ciência Política e Relações Internacionais" do 12º encontro da Associação Brasileira de Ciência Política, João Pessoa, Paraíba (2020). Agradecemos os comentários e sugestões dos participantes, em particular os apontamentos da professora Elia Alves e Jakson Aquino. Parte do conteúdo advém da tese de doutorado de um dos autores (Mesquita, 2018). Materiais de replicação estão disponíveis em: <https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b>. Agradecemos a Ivan Fernandes (UFABC) e Gustavo Fernandes (FGV) pela disponibilização do banco de dados utilizado no trabalho.

Abstract

In spite of the growing offer of data organized in panel format, there are still few studies in Brazil combining cross-sectional and longitudinal dimensions. For instance, out of a sample of over 7 thousand articles published between 2000 and 2018 in Brazilian Political Science and International Relations journals, only 45 articles cited specific techniques to handle observations comprising spatial units (countries, states, persons) repeated in regular time intervals (years, months, days). Given the inferential gains this approach can afford and the scarcity of research on the topic, this article presents an introduction to panel regression. Methodologically, we summarize the main recommendations of the literature and show how to implement them via *R Statistical* and *plm* package, going from model selection to data treatment and presentation of results. To improve the pedagogical potential of this work, we share replication materials, including the raw data and computational scripts. With this article, we aim to broaden the use of longitudinal analysis in empirical research in Political Science and International Relations in Brazil.

Keywords: Panel Data; TSCS; Political Science and International Relations in Brazil; Linear Regression; Political Methodology.

1. Introdução

Muitos dados em Ciência Política e Relações Internacionais (CPRI) estão dispostos em formato de painel, ou seja, observações de unidades espaciais mensuradas em intervalos regulares do tempo (Beck, 2001). A notação *Times-Series Cross-Section* (TSCS) é recorrente na literatura e revela que a base de dados combina uma dimensão espacial (países, estados, municípios, pessoas, etc.) e outra temporal (anos, meses, dias, etc.) (Beck, 2008). Para explicar, a expressão *time-series* informa que os casos estão dispostos ao longo do tempo enquanto a denominação *cross-section* indica que a base é formada por múltiplas observações transversais (Gujarati, 2004).

Para determinadas perguntas de pesquisa, essa combinação de informações temporais e espaciais é de alta importância para obtenção de respostas mais confiáveis do que seria produzido pela utilização de uma única abordagem. Se quiséssemos, por exemplo, verificar se o crescimento econômico afeta o resultado das eleições nos municípios brasileiros (Fernandes e Fernandes 2017), poderíamos de início fazer apenas um estudo *cross-section* comparando vários municípios na última eleição. Esse desenho de pesquisa, porém, será limitado, pois sabemos que nos últimos anos o crescimento econômico em praticamente todo o território nacional foi baixo, de modo que aprenderemos apenas sobre os efeitos da contração econômica e não do crescimento. Outras circunstâncias peculiares do período recente com impacto eleitoral (e.g.: denúncias de corrupção) tampouco poderão ser controladas se tiverem abrangência nacional. Contudo, superar essas dificuldades é possível se reunirmos dados não somente da última eleição, mas de repetidas eleições para todos os municípios de interesse.

Além disso, dados em painel são valiosos pois permitem ganhos importantes em termos de mensuração, controle de *confounders*, aprimoramento de inferências causais e tamanho da amostra (Baltagi, 2005; Fortin-Rittberger, 2013). Obviamente que essas vantagens vêm com custos. Por exemplo, a estrutura do termo de erro requer procedimentos específicos para que os resultados sejam consistentes e válidos (Worrall, 2008). Apesar de propício aos fenômenos estudados pela CPRI, o método continua pouco utilizado: em uma amostra de mais de 7 mil artigos publicados entre 2000 e 2018 em periódicos de CPRI, apenas 45 casos citavam técnicas específicas de regressão de painel.

Assim, diante dos benefícios analíticos de tal abordagem e da sua limitada utilização na área de CPRI, o presente artigo apresenta soluções didáticas sobre o que fazer e não fazer com dados de painel (Beck e Katz, 1995). Nosso público alvo são estudantes de graduação e pós-graduação em fases iniciais de treinamento. Essa contribuição pedagógica se torna tanto mais importante dada a ausência de material didático especialmente voltado para CPRI.[2] A maior parte dos manuais ou é direcionado a outras disciplinas, como economia (Wooldridge, 2002; Kennedy, 2003; Stock e Watson, 2004), ou apresenta um grau proibitivo de complexidade.[3] Pensando nisso, incluímos também uma seção de perguntas e respostas com dúvidas frequentes sobre o assunto, além dos scripts computacionais que permitem que os leitores adaptem as rotinas computacionais a seus respectivos interesses de pesquisa.

O artigo está organizado da seguinte forma: a primeira seção apresenta os resultados de uma análise bibliométrica sobre o uso da regressão de painel em CPRI; a segunda fornece uma introdução à lógica de análise de dados longitudinais; a terceira seção replica os dados de Fernandes e Fernandes (2017) para mostrar o passo a

2 Para trabalhos em português com o mesmo espírito didático, consultar Fávero (2013) e Marques (2000). Para um tutorial em inglês com ênfase computacional no pacote *plm*, ver Henningsen e Henningsen (2019).

3 Ver por exemplo o grau de dificuldade do exame final do curso de econometria para dados de painel do professor William Greene <http://people.stern.nyu.edu/wgreene/Econometrics/PanelDataEconometricsFinalExam.pdf>

passo da implementação computacional da regressão de painel; e a última seção sumariza nossas recomendações sobre como difundir análises longitudinais na pesquisa empírica em CPRI no Brasil.

2. Balanço do uso de painel de dados na CPRI brasileira

Quão populares são dados em painel na CPRI brasileira? Para responder essa pergunta, examinamos 7.764 artigos publicados entre 2000 e 2018 em 17 periódicos hospedados na *Scientific Electronic Library Online Platform* (Scielo)[4] (ver Tabela 1 nos Anexos). Semelhante a Medeiros et al. (2016), utilizamos análise automatizada de conteúdo para detectar palavras-chave no texto integral dos artigos. A partir dos pacotes *rscielo* (Meireles et al., 2019) e *quanteda* (Benoit et al., 2018), contabilizamos a frequência de expressões ligadas aos métodos de painel de dados. De toda a amostra, apenas 45, ou seja 0,58%, mencionaram expressões ligadas a dados em painel.[5] Embora o intervalo se inicie em 2000, somente em 2008 surgem artigos mencionando esses termos, o que significa que a utilização dessa técnica em CPRI é bastante recente.

Comparamos a frequência de dados em painel com o número de citações ao Método de Mínimos Quadrados Ordinários (MQO, em inglês *Ordinary Least Squares, OLS*), tendo em vista ser este último um dos métodos mais difundidos em CPRI (Krueger e Lewis-Beck, 2008).[6] Para cada artigo mencionando dados em painel (110 citações em 45 artigos) existem outros cinco que usam MQO (758 citações em 263 artigos). A Figura 1 ilustra essa comparação.

4 O *rscielo* automatiza a coleta de informações do Scielo. No momento da redação deste artigo, o pacote identifica um total de 374 periódicos, referentes a todas as áreas do conhecimento. Destes, selecionamos 65 que receberam conceito Qualis Capes no campo de CPRI em 2016. Como periódicos podem ser pontuados em mais de uma área, alguns destes títulos, apesar de avaliados no campo de CPRI, possuíam uma área de concentração distinta, como economia, educação ou saúde pública. Inspecionando a lista de periódicos e os sites de cada revista, retivemos 17 títulos cujo foco é declaradamente CPRI. Deste subconjunto, coletamos o texto de 8.829 artigos ao todo. Após eliminação de duplicatas e falsos-positivos, chegou-se a 8.789. Filtrando em seguida apenas publicações entre 2000 e 2018, obtivemos a quantidade final de 7.764.

5 Termos utilizados para o dicionário de dados em painel: "panel data", "dados em painel", "painel de dados", "tscs", "*time series cross section*".

6 Termos utilizados para o dicionário de MQO: "*ols*", "mmq", "mqo", "*ordinary least squares*", "mínimos quadrados", "*linear regression*", "regressão linear", "*regression analysis*", "análise de regressão".

Figura 1 – Prevalência de MQO/OLS vs. painel de dados



Fonte: elaborado pelos autores. O gráfico à esquerda contabiliza ocorrências das palavras-chaves ligadas a MQO e dados em painel e o da direita conta o total de artigos. As linhas azuis mostram os totais para dados em painel (PD) e as vermelhas para MQO (OLS).

Apesar da tendência de crescimento, a popularidade dos métodos de dados em painel hoje é equivalente à de MQO há uma década. O uso de dados em painel é, ademais, fortemente concentrado: 42% dos artigos foram publicados na Revista de Administração Pública. Das 17 revistas, seis não contabilizaram sequer uma menção ao método. Esses resultados apontam para uma baixa difusão dos dados em painel na produção nacional em CPRI.

No restante do artigo, introduziremos os conceitos essenciais para regressão de painel a partir da revisão bibliográfica dos principais manuais e mostraremos o passo a passo da implementação computacional a partir da replicação de Fernandes e Fernandes (2017).

3. Entendendo dados longitudinais

Esta seção apresenta os fundamentos da análise de dados longitudinais. Descrevemos primeiramente as diferentes formas de tabular informações ao longo do tempo. Discutimos ainda os principais tipos de painel (curto x longo; balanceado x desbalanceado) (seção 2.1). Em seguida, explicamos a notação do modelo de regressão de painel e as quatro principais modalidades de estimação (2.2). Por fim, apresentamos as vantagens e desvantagens de cada abordagem (2.3 e 2.4).

Dados longitudinais podem ser compreendidos a partir de uma comparação com dados transversais (Menard,

2002). Em um desenho transversal (*cross-section*), as informações sobre casos e variáveis são coletadas em um ponto específico do tempo. Por exemplo, taxa de homicídios (variável) para todas as unidades da federação (casos) em 2013 (tempo). Por outro lado, dados longitudinais são coletados em pelo menos dois períodos diferentes de tempo, o que permite observar a mudança dos valores em cada caso para cada variável de interesse.[7] Por exemplo, a taxa de homicídios (variável) para todas as unidades da federação (casos) entre 2000 e 2010 (tempo). A Figura 2 ilustra a diferença entre dados transversais, séries temporais, e sua combinação na forma de painel. Como exemplo, utilizamos a taxa de homicídios por 100 mil habitantes de três estados brasileiros.

Figura 2 – Estrutura dos dados

Transversal / *Cross-Section*

UF	Tempo	Taxa de homicídios	Média Geral	Desvio Entre Casos
i	t	X_i	$\bar{X}$	$X_i - \bar{X}$
Acre	2011	22,0	24,7	-2,7
Bahia	2011	39,4	24,7	14,7
Santa Catarina	2011	12,8	24,7	-11,9

Série Temporal / *Time-Series*

UF	Tempo	Taxa de homicídios	Média Individual	Desvio Intra-Caso
i	t	X_t	$\bar{X}$	$X_t - \bar{X}$
Bahia	2011	39,4	40,2	-0,8
Bahia	2012	43,4	40,2	3,2
Bahia	2013	37,8	40,2	-2,4

Dados em Painel / *Time-Series Cross-Section*

UF	Tempo	Taxa de homicídios	Média Geral	Média Individual	Desvio Geral	Desvio Intra-Caso	Desvio Entre Casos
i	t	X_{it}	$\bar{X}$	$\bar{X}_i$	$X_{it} - \bar{X}$	$X_{it} - \bar{X}_i$	$\bar{X}_i - \bar{X}$
Acre	2011	22	26,4	26,5	-4,4	-4,5	0,1
Acre	2012	27,4	26,4	26,5	1,0	0,9	0,1
Acre	2013	30,1	26,4	26,5	3,7	3,6	0,1
Bahia	2011	39,4	26,4	40,2	13,0	-0,8	13,8
Bahia	2012	43,4	26,4	40,2	17,0	3,2	13,8
Bahia	2013	37,8	26,4	40,2	11,4	-2,4	13,8
Santa Catarina	2011	12,8	26,4	12,5	-13,6	0,3	-13,9
Santa Catarina	2012	12,9	26,4	12,5	-13,5	0,4	-13,9
Santa Catarina	2013	11,9	26,4	12,5	-14,5	-0,6	-13,9

Fonte: elaboração própria, com base em dados de Cerqueira (2017).

7 Menard (2002) identifica quatro desenhos de pesquisa longitudinal: (1) *total population designs*; (2) *repeated cross-sectional designs*; (3) *revolving panel designs* e (4) *longitudinal panel designs*. Wooldridge (2013, p.448) também diferencia entre "*independently pooled cross section*" e "*panel data*"/"*longitudinal data*"; no primeiro caso, trata-se de amostras aleatórias de indivíduos diferentes em sucessivos momentos no tempo (e.g.: amostragem do censo), ao passo que no último os sujeitos observados permanecem os mesmos ao longo do tempo.

Em nosso exemplo, a variável *UF* representa a unidade de análise transversal. A variável *Tempo* indica a dimensão longitudinal das informações, no caso, o ano. X_{it} representa o valor da variável de interesse para o caso *i* no tempo *t*. Por exemplo, o Acre teve uma taxa de homicídios de 22 em 2011. As demais colunas apresentam medidas agregadas derivadas de X_{it}; embora estas normalmente não figurem nas planilhas de dados propriamente, nós as reproduzimos para fins didáticos.

A média geral (por vezes chamada *grand mean*) é a soma total dos casos (237,7) dividida pelo tamanho da amostra (9), ou seja, 26,4. O desvio geral é a diferença de cada observação da média geral. Por exemplo, como em 2011 a Bahia teve taxa 39,4, ela estava 13 pontos acima da média geral (39,4 - 26,4). O desvio entre os casos é a diferença entre a média de cada caso e a média geral. Por exemplo, Santa Catarina dista 12,5 - 26,4 = -13,9 pontos da média geral. Por fim, o desvio intra-caso representa a diferença entre cada observação e a média individual daquele estado, de modo que, em 2013, Santa Catarina esteve 11,9 - 12,5 = -0,6 pontos abaixo da sua média histórica.

Esta configuração em que cada linha da planilha representa uma combinação única de caso no tempo é também conhecida como formato "*long*" e é o arranjo padrão para a maior parte dos programas estatísticos. Em diversos cenários, pode haver variação na disponibilidade de informações transversais ou longitudinais. Como dados em painel são o produto da combinação de ambas, resulta que diferentes tipos de painéis podem ser formados, a depender do número de casos ou de unidades temporais disponíveis, como explicamos a seguir.

3.1 Tipos de painel: curtos x longos, balanceados x não-balanceados

Painéis assumem duas principais configurações. Há um painel curto, "*stacked*" ou "*cross-section dominant*" quando o número de casos supera a quantidade de períodos temporais, ou seja, N>T. Por sua vez, dizemos que o painel é longo ("*temporally dominant*") quando o número de períodos temporais é maior do que a quantidade de casos (T>N). Alguns autores convencionaram chamar apenas os primeiros de painel de dados e os últimos de dados *Time-Series Cross-Section* (TSCS) (Beck, 2001; Fortin-Rittberger, 2013).

Para cada configuração do painel (curto ou longo), assume-se diferentes expectativas em relação à distribuição dos dados. Beck (2001) argumenta que, para dados em painel (N>T), as unidades selecionadas são encaradas como amostras de uma população, observadas em um curto intervalo de tempo. Dessa forma, as inferências não são específicas às unidades e podem ser generalizáveis. Já em TSCS (T>N), unidades são julgadas como fixas e observadas por um longo período de tempo. Desse modo, as inferências são orientadas às unidades e menos generalizáveis. O exemplo tradicional da primeira são pesquisas de censo, enquanto que da última são comparações macroeconômicas entre países.[8]

Outra característica importante diz respeito à disponibilidade da informação. Temos o painel balanceado ou equilibrado quando há informação para todos os casos em todos os períodos de tempo. Por outro lado, dizemos que o painel é não-balanceado ou desequilibrado quando alguns casos estão ausentes em determinados períodos de tempo. Uma complicação adicional, que pode afetar a consistência das estimativas, diz respeito à natureza do desequilíbrio do painel. A ausência de dados pode ser aleatória, o que tende a prejudicar a eficiência dos coeficientes (erro padrão, p-valor, intervalos de confiança). Todavia, o desbalanceamento pode ser relacionado com alguma variável de interesse

8 Semelhantemente, as propriedades assintóticas de TSCS supõem que N é fixo e T pode crescer até o infinito. Já as de dados em painel, que T é fixo e N pode crescer ao infinito (Fortin-Rittberger, 2013). Daí deduz-se que, quanto maior N em termos proporcionais, melhor os dados se adequam a modelos de dados em painel e, inversamente, quanto maior o T, melhor a adequação aos modelos TSCS.

do estudo. Por exemplo, pode ser mais fácil encontrar informações detalhadas para países mais ricos do que nações mais pobres. Quando a ausência de informação é sistemática por qualquer motivo, as estimativas tendem a apresentar viés. De acordo com Baltagi (2005), painéis desequilibrados tendem a ser mais frequentes do que painéis completos, o que reforça a importância de descrever exatamente o processo de coleta, tratamento e análise dos dados.

Explicamos agora como dados em painel podem ser utilizados para análise de regressão.

3.2 Notação e tipos de regressão de painel

Para os propósitos do artigo, é importante apresentar a notação clássica do modelo de regressão linear, assumindo dados apenas transversais:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad (1)$$

Em que Y representa a variável dependente, α indica o intercepto, β representa a variação observada em Y quando a variável independente, X, aumenta em uma unidade, e ε indica a natureza estocástica do modelo. O subscrito *i* indica que as observações são indexadas por caso. Continuando com o exemplo da Figura 2 (taxa de homicídios por UF ao longo do tempo), essa equação pode representar o efeito do desemprego (X) sobre a taxa de homicídios (Y). A regressão de painel é uma extensão do modelo anterior, em que:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \qquad (2)$$

Os subscritos *i* e *t* informam que as observações são indexadas, respectivamente, por caso e tempo. Nosso modelo se torna mais poderoso, pois acumulamos informações sobre a relação entre desemprego e homicídios para várias UF e anos. Como prosseguir, entretanto, com a regressão? Descrevemos abaixo as quatro abordagens tradicionais para estimar modelos de painel.

Pooled OLS

Ao reunir todas as observações em uma mesma regressão linear ("*pooling*"), tem-se *Pooled OLS* (POLS). Essa abordagem pressupõe que todas as unidades podem ser contempladas pela mesma reta de regressão. Mesmo que as unidades *i* apresentem diferenças umas das outras, considera-se que o conjunto de variáveis independentes já carrega toda a informação importante. Ou seja, os principais fatores que diferenciam entre as observações já estão explicitados no modelo, de modo que não é necessário controlar por outros fatores não-observados. Em nosso exemplo, seria como se a taxa de homicídios, qualquer que fosse a UF, fosse principalmente determinada pelo desemprego, e não houvesse outras peculiaridades nas UF afetando o resultado: se a Bahia tivesse o nível de desemprego de Santa Catarina, teria a mesma taxa de homicídios desta. Como heterogeneidades individuais não são modeladas, temos a Equação (2), em que há apenas um intercepto (α) para toda a população e todo o mais que não é observável é alocado como ruído no fator de erro (ε).

Se houver fatores não-observados em ε que, entretanto, não forem aleatórios, mas sistemáticos, POLS já não será apropriado. Se tivermos elementos não-observados correlacionados com X, temos um caso tradicional de viés de variável omitida e a estimativa do parâmetro β ($\hat{\beta}$) será inconsistente.[9] Como as observações se repetem no tempo, a independência das observações está em risco. Suponha que para os três estados do exemplo haja algum fator que não podemos observar e que, contudo, impacta na taxa de homicídios, digamos a influência do

9 Se o fenômeno estudado é melhor descrito por um conjunto de variáveis e alguma dessas é deixada de fora da equação, o efeito das variáveis preservadas no modelo será ou sobre ou subestimado. Essa distância sistemática para mais ou para menos no parâmetro caracteriza um viés.

crime organizado. Se a influência do crime organizado for correlacionada com desemprego, haverá viés; se for específica em cada estado (e.g.: a influência criminosa no Acre é a mesma ao longo do tempo), os erros não serão independentes (ε_{it} será correlacionado com ε_{it-1}). Por essa razão, a primeira motivação para se afastar do modelo POLS é a necessidade de trabalhar a heterogeneidade oculta dentro do termo de erro de uma maneira mais sofisticada.

Os fatores não-observados se dividem em elementos que permanecem sempre os mesmos para cada unidade e elementos que variam segundo a unidade e o tempo. Então, o "erro composto" (ε_{it}) da Equação (2) pode ser destrinchado em duas partes: uma fixa para cada unidade (μ_i) e outra variável (ν_{it}).

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it} \qquad (3)$$

Os erros persistentes às unidades (μ_i) também são conhecidos como heterogeneidade não-observada, efeito individual, ou efeito fixo do termo de erro, enquanto que ν_{it} é usualmente chamado de efeito variável ou idiossincrático. Diferente das outras três abordagens que serão apresentadas aqui, POLS não dispõe de estratégias para lidar com μ_i, pois não desagrega ε_{it}. Todas as demais especificações decompõem o fator de erro em fixo e variável.

Primeiras Diferenças / First Differences

Na abordagem de Primeiras Diferenças ou *First Differences* (FD), todas as variáveis são diferenciadas, isto é, subtraídas de seu valor no período temporal anterior. A subtração é eficiente em lidar sobretudo com problemas dinâmicos dos dados, como correlação serial e não-estacionariedade, já que, ao subtrair uma observação em *t* da observação em *t*-1, tudo o que não houver variado de um instante para o outro (inclusive a heterogeneidade não-observada) será apagado.

Por subtrair observações aos pares, FD traz alguns custos: reduz a variabilidade dos dados e assim aumenta os erros-padrão, sendo portanto mais recomendado para bases grandes. Ademais, a diferenciação remove todos fatores constantes ao longo do tempo ou variando monotonicamente (i.e.: sempre aumentando à mesma taxa todo ano), o que impossibilita o uso de variáveis como *dummies* para países, distâncias geográficas e variáveis de crescimento constante (Wooldridge, 2013, p. 462–463). O uso de FD é uma das soluções drásticas recomendadas por Beck (2008) diante de dependências temporais persistentes no modelo.

Efeitos Fixos / Fixed Effects

Fixed Effects (FE) ou Efeitos Fixos, assim como a FD, buscam livrar-se do componente fixo do erro (μ_i), porém, ao invés da diferenciação, utilizam a *within transformation*. Dá-se o nome de *within* a essa operação pois ela retém apenas o desvio intra-caso (ver Figura 2). Essa remoção do valor da média de cada observação ("*time demeaning*" ou "*group-mean centering*") terá como resultado a eliminação de α e μ_i, cujos valores para cada unidade não mudam com a passagem do tempo.[10] Portanto, assim como FD, o estimador FE omite fatores constantes no tempo.

Por lidar com o problema da heterogeneidade e ser bastante robusta, a especificação FE é vista por Croissant e Millo (2018, p.5) como um *benchmark* para dados em painel. Isso é especialmente o caso para disciplinas que lidam com dados observacionais, o que também se verificou em nosso levantamento bibliométrico: efeitos fixos foram o método mais recorrente para CPRI (26% dos artigos).

Um dos benefícios de FE em comparação com POLS é que, enquanto este postula apenas um

10 Os componentes α e μ_i são removidos para estimação consistente de β. Repare, não obstante, que alguns *softwares* estatísticos reportam valores para interceptos individuais (μ_i) ou populacionais (média de todos os μ_i). Vide Wooldridge (2002, seção 10.5.3) para detalhes.

intercepto para toda a amostra, aquele leva em consideração que cada unidade pode ter seu próprio intercepto. Em comparação com FE, POLS terá estimadores viesados, já que não lida com μ_i. Já sob FE, o modelo se torna menos viesado, porém também menos eficiente.

Computacionalmente, obtêm-se resultados semelhantes em aplicar um modelo FE e um POLS acrescido de *dummies* para todos os casos. Nesta situação, POLS geralmente atende pelo nome de *Least Square Dummy Variable* (LSDV). A Figura 3 ilustra os procedimentos de diferenciação e de subtração da média.

Figura 3 - Ilustração dos procedimentos de diferenciação e *within transformation*

Primeiras Diferenças / First Differences

ID	Tempo	Taxa de homicídios	Taxa de desemprego
i	*t*	Y_{it}	X_{it}
Acre	2011	22	5,4
Acre	2012	27,4	8
Acre	2013	30,1	9,6
Acre	2014	29,4	9,8
Bahia	2011	39,4	10,5
Bahia	2012	43,4	10,2
Bahia	2013	37,8	9,9
Bahia	2014	40	10
Santa Catarina	2011	12,8	3,6
Santa Catarina	2012	12,9	3,1
Santa Catarina	2013	11,9	3,4
Santa Catarina	2014	13,5	3,1

$$Y_{it-1} = \alpha + \beta X_{it-1} + \mu_i + \nu_{it-1}$$
$$- \left(Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it} \right)$$
$$\Delta Y_{it} = \beta \Delta X_{it} + \Delta \nu_{it}$$

ID	Tempo	Taxa de homicídios	Taxa de desemprego
i	*t*	ΔY_{it}	ΔX_{it}
Acre	d1	5,4	2,6
Acre	d2	-0,7	0,2
Bahia	d1	4	-0,3
Bahia	d2	2,2	0,1
Santa Catarina	d1	0,1	-0,5
Santa Catarina	d2	1,6	-0,3

Efeitos Fixos / Fixed Effects

ID	Tempo	Taxa de homicídios	Taxa de desemprego	Média Individual (homic.)	Média Individual (desemp.)
i	*t*	Y_{it}	X_{it}	$\bar{Y}_i$	$\bar{X}_i$
Acre	2011	22	5,4	27,2	8,2
Acre	2012	27,4	8	27,2	8,2
Acre	2013	30,1	9,6	27,2	8,2
Acre	2014	29,4	9,8	27,2	8,2
Bahia	2011	39,4	10,5	40,2	10,2
Bahia	2012	43,4	10,2	40,2	10,2
Bahia	2013	37,8	9,9	40,2	10,2
Bahia	2014	40	10	40,2	10,2
Santa Catarina	2011	12,8	3,6	12,8	3,3
Santa Catarina	2012	12,9	3,1	12,8	3,3
Santa Catarina	2013	11,9	3,4	12,8	3,3
Santa Catarina	2014	13,5	3,1	12,8	3,3

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it}$$
$$- \left(\bar{Y}_i = \alpha + \beta \bar{X}_i + \mu_i + \bar{\nu}_i \right)$$
$$\ddot{Y}_i = \beta \ddot{X}_i + \ddot{\nu}_i$$

ID	Tempo	Desvio Intra-Caso (homic.)	Desvio Intra-Caso (desemp.)
i	*t*	$Y_{it} - \bar{Y}_i$	$X_{it} - \bar{X}_i$
Acre	2011	-5,2	-2,8
Acre	2012	0,2	-0,2
Acre	2013	2,9	1,4
Acre	2014	2,2	1,6
Bahia	2011	-0,8	0,4
Bahia	2012	3,3	0,0
Bahia	2013	-2,4	-0,3
Bahia	2014	-0,1	-0,2
Santa Catarina	2011	0,0	0,3
Santa Catarina	2012	0,1	-0,2
Santa Catarina	2013	-0,9	0,1
Santa Catarina	2014	0,7	-0,2

Fonte: elaborado pelos autores, a partir de dados do Cerqueira et al. (2017) e IBGE (2017).

O esquema ilustra os efeitos da diferenciação e da *within transformation* apenas para as variáveis observadas: homicídios e desemprego. Mesmo assim, a subtração abrange também os fatores não-observados (indicados em itálico nas equações da imagem), resultando na supressão do intercepto geral (α) e do efeito individual do fator de erro (μ_i). Em nosso exemplo, significa que os modelos resultantes, ao estimarem o efeito do desemprego sobre a taxa de homicídios, eliminam a influência fixa do crime organizado em cada estado. Não só este, mas todo fator não-observado que for invariante no tempo será removido, e a estimação irá se basear apenas nos elementos que variam.

Efeitos Aleatórios / Random Effects

Como foi visto, uma deficiência de FD e FE é a impossibilidade de testar variáveis que não mudam com o tempo. Mesmo as que variam lentamente tendem a ser subestimadas (Clark e Linzer 2015). Todavia, há muitas variáveis de interesse para CPRI que são estáticas, como geografia, gênero, nacionalidade e outras. A especificação dos Efeitos Aleatórios ou *Random Effects* (RE) contorna essa dificuldade, permitindo a inclusão de variáveis que não mudam com a passagem do tempo.

Assim como para FE, supõe-se que o erro composto possui parte fixa e parte idiossincrática, ou seja, $\varepsilon_{it} = \mu_i + \nu_{it}$. Contudo, ao invés de eliminar μ_i por inteiro, supõe-se que ele não tenha correlação com X (Bell e Jones 2015, p. 142). Se o efeito fixo individual não tem correlação com as variáveis independentes, não é mais fonte de viés e portanto não precisa mais ser eliminado; e.g.: se acreditamos que crime organizado é independente do desemprego, não há motivo para removê-lo.

O segundo pressuposto distintivo de RE diz respeito aos efeitos individuais. Efeitos *fixos* supõem que o efeito de cada indivíduo pode ser calculado como um valor único. Por isso, FE calcula um intercepto para cada indivíduo na forma de um parâmetro. RE, em contraste, considera μ_i e ν_{it} como realizações *aleatórias* de uma distribuição. No exemplo envolvendo o crime organizado não-observado nas UFs, FE irá estimar esse impacto individual caso a caso, sem supor que há alguma relação entre o nível de crime organizado na Bahia e no Acre. Já RE imprime mais estrutura ao conjunto, pois supõe que os efeitos do crime tendem a uma distribuição normal, consideradas todas as UFs (Bell et al., 2019, p. 1061).

Alguns autores classificam RE como um meio-termo entre POLS e FE (Clark e Linzer 2015, p. 402). POLS não elimina μ_i e FE o apaga completamente. RE executa parcialmente este procedimento, pois subtrai das observações uma parcela da média intra-grupo ("*quasi-demeaning*"). A proporção dessa parcela é dada por θ (*theta*, no output do *plm*) que assume valores de 0 a 1. *Theta* pode ser interpretado como um diagnóstico de qual componente mais contribui com a variância de ε: se fixo (μ_i) ou idiossincrático (ν_{it}). Quando há pouca variância nos efeitos fixos ($\theta \to 0$) as estimativas de RE se aproximam de POLS e, no caso contrário, ($\theta \to 1$) de FE (Wooldridge, 2013, p. 494).

Em contraponto, a suposição que não há relação entre μ_i e X é difícil de sustentar em boa parte das situações reais. Outra diferença significativa é que estimadores RE sempre terão algum viés, podendo só ser eficientes. A escolha entre FE e RE é, portanto, um *tradeoff* entre entre eficiência (RE) e não-viesamento (FE).

A comparação direta entre RE e FE é recorrente na literatura. Disciplinas mais afeitas ao uso de experimentos randomizados tendem a preferir RE (Dieleman e Templin, 2014). Beck (2001, p. 284) argumenta que estudos observacionais preferem FE a RE, pois este último é mais recomendado quando se consideram as unidades como uma amostra de uma população mais ampla, e aquele quando se busca fazer inferências restritas às unidades. A ideia de "intercambialidade" captura essa noção de forma intuitiva: quando para um conjunto de observações podemos substituir um caso A por outro B sem grande perda de informação - ou seja, o *nome do caso* não importa - temos algo próximo de uma amostragem aleatória (Hausman 1978, p. 1262; King 2001, p. 498). A expectativa de RE de uma distribuição normal dos efeitos individuais se torna então cabível. Em alguns cenários de CPRI, essa intercambialidade é plausível (e.g.: deputados em um congresso), em outros não (e.g.: países membros do Conselho de Segurança da ONU).

Por fim, ao fazer uma análise empírica, é sempre informativo comparar lado a lado os resultados das quatro formas de regressão: POLS, FD, FE e RE. O contraste

aponta para os diferentes resultados obtidos a depender do que se faz com μ_i: POLS mantém o fator fixo totalmente inserido no componente de erro, RE mantém uma parcela, e FE e FD o eliminam completamente (Wooldridge, 2013, p. 494).

3.3 Benefícios dos dados em painel

O Quadro 1 sintetiza os principais benefícios técnicos e problemas operacionais associados aos dados em painel.

Quadro 1 – Benefícios e problemas dos dados em painel

Benefícios	Problemas
(1) Facilitar a detecção de causalidade; (2) Mensuração de variação individual; (3) Reduzir erros de mensuração; (4) Aumentar o tamanho da amostra; (5) Controlar por problemas de variáveis omitidas.	(1) Correlação serial dos resíduos; (2) Não-estacionariedade; (3) Heterogeneidade; (4) Heteroscedasticidade de painel; (5) Correlação contemporânea de erros; (6) Estruturas complexas de dependência.

Fonte: elaborado pelos autores, com base em Fortin-Rittberger (2013), Hsiao (2003) e Kennedy (2003).

A primeira vantagem dos painéis é facilitar a detecção de relações causais entre X e Y. Pearl (2000) aponta três pressupostos para identificar uma relação causal: (1) associação entre as variáveis (correlação); (2) precedência temporal e (3) não-espuriedade da relação. Diferente dos dados transversais, em que os valores de X e Y são mensurados no mesmo período, dados em painel permitem observar se a variação temporal de X está correlacionada com a variação temporal de Y. Dessa forma, além de controlar por eventuais variáveis espúrias, os dados longitudinais favorecem a satisfação do pressuposto da precedência temporal entre a variável independente e a dependente.

Além disso, a natureza longitudinal também favorece o estudo da mudança, já que conserva a mesma unidade de análise ao longo do tempo. Assim, é possível examinar a variação individual de diferentes indicadores entre os casos de interesse. Por exemplo, a análise agregada do número de homicídios em uma cidade é informativa. Já uma base de dados que informa o dia do óbito e bairro abre possibilidades adicionais. Será possível observar, por exemplo, se os finais de semana em certo bairro concentram a maior parte das ocorrências. Como aponta Fortin-Rittberger (2013), essa possibilidade de monitorar, no espaço e no tempo, onde e quando surgem certos fenômenos de interesse aproxima os estudos longitudinais, mesmo sendo observacionais, de uma lógica de experimento controlado.

Como terceira vantagem, temos medidas mais confiáveis sobre a mudança das variáveis. De acordo com Blalock (1979), o problema da mensuração é o principal desafio ao desenvolvimento das ciências sociais. Em particular, em CPRI, conceitos importantes, democracia por exemplo, não são diretamente observáveis. Uma das formas de examinar a confiabilidade de uma medida é mensurá-la várias vezes ao longo do tempo. Logo, os dados em painel ajudam a superar eventuais erros de mensuração.[11]

11 Erros podem ser sistemáticos ou aleatórios. O erro sistemático é prejudicial em análises descritivas pois as estimativas produzidas serão viesadas. Por outro lado, quando o erro sistemático atinge todos os casos com a mesma magnitude não afeta a consistência das estimativas em modelos de regressão. Por sua vez, o erro aleatório na variável dependente tende a reduzir a eficiência (maior erro padrão), mas não afeta a consistência das estimativas. Por fim, quando as variáveis independentes também apresentam erro aleatório, as estimativas tendem a ser inconsistentes e, em geral, subestimadas (King, Keohane e Verba, 1994).

Além dos ganhos em mensuração, outro trunfo dos dados em painel é aumentar o tamanho da amostra. Um dos problemas recorrentes em desenhos de pesquisa é a indeterminação, ou seja, "muitas variáveis, N pequeno" (Lijphart, 1971, p. 686). Quando temos poucas observações, é difícil sustentar que a relação causal proposta se explica por um motivo e não qualquer outro. Assim, a determinação de um desenho de pesquisa diz respeito à proporção entre variáveis e observações. Por isso, aumentar o número de casos é uma das recomendações costumeiras para solucionar a indeterminação (King, Keohane e Verba, 1994). Ademais, amostras maiores superam problemas de eficiência (menor erro padrão), multicolinearidade e facilitam a utilização de testes estatísticos mais robustos (Hsiao, 2003).

Por fim, dados em painel controlam os problemas de variáveis omitidas (Kennedy, 2003). A omissão de uma variável importante produzirá viés nas estimativas se ela estiver correlacionada com outras variáveis modeladas. A perspectiva puramente transversal não fornece muitas soluções para lidar com fatores não-observados. Frequentemente, somos simplesmente ignorantes sobre os *unobservables* que incidem sobre Y. Contudo, a partir do momento em que os casos são acompanhados ao longo do tempo, expedientes como FE, FD e RE podem ser usados para remover ou mitigar essas heterogeneidades individuais (Wooldridge 2002, seção 10.1).

Esses benefícios, entretanto, não vêm sem custo, como vemos a seguir.

3.4 Desafios dos dados em painel e como superá-los

Para fins didáticos, propomos uma breve recapitulação dos pressupostos da regressão linear de MQO como ponto de partida para abordar os problemas típicos das regressões de painel.[12] Seguindo Wooldridge (2013, p. 59, 119), os principais pressupostos para MQO podem ser resumidos em seis: (P.1) Há uma relação linear entre X e Y; (P.2) Ausência de autocorrelação (ou amostra aleatória);[13] (P.3) Ausência de multicolinearidade perfeita ("*full rank*");[14] (P.4) Exogeneidade de X (ou "média condicional zero"): o erro não deve ser uma função das variáveis independentes. Ou seja, ele é aleatório e nenhuma das variáveis independentes no modelo porta informação a respeito do seu valor;[15] (P.5) Homocedasticidade: há variância uniforme dos resíduos da regressão e estes não são correlacionados uns com os outros. Se por algum motivo em um segmento dos dados a variância dos resíduos muda em sintonia com os valores das variáveis explicativas, há heteroscedasticidade;[16] e (P.6) Distribuição normal dos erros.

Se as chamadas premissas Gauss-Markov (P.1-P.5) são respeitadas, pode-se dizer que os parâmetros estimados são os *Best Linear Unbiased Estimators* (ou BLUE) dos parâmetros populacionais. "*Best*" porque possuem a menor variância já que há (P.5) homocedasticidade (variância uniforme dos resíduos); "*Linear*" por conta da relação linear (P.1); e "*Unbiased*" porque não há viés no parâmetro, já que os regressores são exógenos e não afetam o erro (P.4).

Porém, modelos de painel facilmente podem infringir essas suposições. Por combinarem tanto séries temporais quanto comparações transversais, herdam os problemas de ambos os tipos de dados. Os problemas

12 A lista de pressupostos da MQO por vezes varia entre autores. Para outras listagens dos pressupostos da regressão linear e do teorema de Gauss-Markov, ver Berry (1993) e Larocca (2012).

13 Amostras aleatórias e ausência de autocorrelação podem parecer expectativas distintas, mas são equivalentes pela seguinte perspectiva: em ambos os casos, se tem que $Corr(\varepsilon i, \varepsilon j = 0)$, para $i \neq j$.

14 O principal dano provocado pela multicolinearidade é o aumento da variância dos coeficientes da regressão, o que aumenta a chance de encontrar resultados não-significativos.

15 A violação da exogeneidade equivale a um viés de variável omitida.

16 Em outras palavras, o modelo se encaixa melhor a certas porções dos dados e não a outras. Heteroscedasticidade não leva a viés ou inconsistência de estimadores MQO, porém por interferir na distribuição das variâncias impossibilita a construção de intervalos de confiança e testes de significância (Wooldridge, 2013, pp. 268–269).

podem ser divididos entre os dinâmicos e os espaciais. A literatura geralmente recomenda atentar em primeiro lugar para os problemas dinâmicos. Estes violam principalmente os requisitos da ausência de autocorrelação (P.2), exogeneidade (P.4) e homocedasticidade (P.5). Os principais são:

(1) **Correlação serial dos resíduos:** há correlação entre os resíduos no momento *t* e *t-1*. Trata-se de uma das características persistentes de painéis longos, pois, como Box et al. (2016) advertem, um aspecto intrínseco das séries temporais é que observações adjacentes são interdependentes. O principal problema gerado pela correlação dos erros é a inconsistência dos testes de significância (p-valor e intervalos de confiança).[17] Por esse motivo, a correlação serial deve ser detectada e solucionada antes de seguir adiante tratando dos problemas espaciais. Há diferentes ferramentas para detecção de processos temporais na estrutura de erros: os gráficos da função de autocorrelação (ACF) e autocorrelação parcial (PACF), os testes de Durbin-Watson, Breusch-Godfrey, Ljung-Box, entre outros. Em sua maioria, eles são direcionados a verificar processos autorregressivos de primeira ordem, ou AR(1), que são mais habituais.[18]

Pode-se testar correlação AR(1) nos erros de modelos POLS de forma simples e direta, fazendo uma regressão dos resíduos em *t* nos resíduos em *t*-1, sem intercepto, conforme $\varepsilon_{it} = \rho\varepsilon_{it-1} + \nu_{it}$. O resultado para o parâmetro ρ indicará se resíduos passados estão ou não correlacionados de forma significativa com os futuros (Wooldridge, 2002, seção 7.8.5; 2013, p. 417).

As especificações FE, RE e FD exigirão testes específicos. Cobrir todo o arsenal de testes existentes ultrapassa o escopo deste artigo (para uma lista, ver Croissant e Millo 2008, p. 22-28; 2018, seção 4.3), porém destacamos que a seleção de testes deve ser guiada pelo tipo de estimação utilizada no modelo (POLS, FE, RE ou FD), a proporção N/T, e o número de observações.

Uma vez que se perceba correlação serial, soluções recomendadas incluem adicionar a variável dependente em *lag*, inserir erros-padrões robustos, ou aplicar FD. Quanto à primeira alternativa: recomenda-se a inserção de uma *lagged dependent variable* (LDV), supondo que o processo seja AR(1) e após isso fazer novo teste para verificar se a autocorrelação foi eliminada (Beck, 2001). Tal solução, não obstante, foi criticada por Achen (2000, p. 14), afirmando que a LDV pode absorver todo o poder explicativo dos demais regressores e assim produzir viés.[19]

(2) **Não-Estacionariedade:** uma série temporal é estacionária quando está em equilíbrio estatístico, isto é, apresenta média e variância constantes ao longo do tempo. Ainda que os valores oscilem, retornarão à média. É não-estacionária quando a média e variância não são constantes, ou seja, a série não tende a retornar a uma média anterior após desvios. Dito de outra forma, a não-estacionariedade pode ser entendida como um efeito persistente de choques anteriores (Box et al., 2016, p. 7; Fortin-Rittberger, 2013). Desvios da estacionariedade podem advir do tipo de fenômeno monitorado, mas também podem resultar da frequência da mensuração: se as informações foram mensuradas por dia ou por mês, por exemplo, é possível encontrar efeitos sazonais e, para amostras com T grande, comportamento cíclicos. A inspeção visual da série temporal é uma das maneiras mais simples de identificar tendências ou sazonalidade. Diferentes testes oriundos da literatura sobre séries temporais

17 Tecnicamente, a autocorrelação afeta a eficiência das estimativas. Em particular, quando a autocorrelação for positiva, os erros padrão dos coeficientes serão subestimados, o que aumenta a chance de incorretamente rejeitar a hipótese nula (erro do tipo I).

18 Embora não seja habitual esperar processos mais complexos que AR(1) para dados em painel (Fortin-Rittberger, 2013), é possível que haja correlações de ordem superior, o que irá motivar remédios mais complexos (e.g.: a inclusão de variáveis com múltiplos *lags*).

19 Quando os regressores possuem alguma tendência, a introdução de uma LDV irá dominar a regressão. O coeficiente da LDV será viesado para cima e o dos demais regressores para baixo. Também não é recomendável quando o T é pequeno, pois sacrifica um período de análise (Achen, 2000).

podem ser empregados para verificar não-estacionaridade: ACF, PACF e testes de raiz unitária (e.g.: *Augmented Dickey-Fuller,* ADF). Mais recentemente, testes próprios ao contexto de dados em painel, "*Panel Unit Root Tests*", têm sido propostos (ver Croissant e Millo 2018, cap.8). A aplicação desses testes será dificultada se o painel tiver poucos períodos temporais. Se se verificar que a série é não-estacionária, soluções incluem utilizar FD ou, em casos de não-estacionariedade por tendência, incluir uma variável da passagem do tempo (*time trend*) (Fortin-Rittberger, 2013). Observemos agora os aspectos espaciais.

(3) **Heterogeneidade:** regressões supõem que os casos são homogêneos. Quaisquer características peculiares são explicadas pelas variáveis independentes, como denotado anteriormente através do conceito de "intercambialidade" (King, 2001, p. 498). Há heterogeneidade quando um ou alguns dos casos observados possuem características distintivas que, não tendo sido modeladas, irão alojar-se no termo de erro e assim provocar correlações entre ε_{it} e X_{it}. Heterogeneidade viola, desse modo, os pressupostos de exogeneidade (P.4) e homoscedasticidade (P.5). Se houver heterogeneidade, normalmente será indevido supor que há um só intercepto para toda a população, como nos modelos POLS. Se cada unidade possui características fixas impactantes, a ausência de interceptos próprios a cada uma levará a uma linha de regressão equivocada (Fortin-Rittberger, 2013). Por sua vez, a violação do pressuposto da homocedasticidade implica na produção de testes de significância inconsistentes, o que aumenta a chance de incorrer em inferências errôneas.

Como discutimos na seção 3.2, abandonar estimações POLS em favor de outras especificações é uma das formas de sanar a heterogeneidade. As últimas três dificuldades espaciais comprometem principalmente a exogeneidade (P.4), homocedasticidade (P.5) e a normalidade dos erros (P.6).

(4) **Heteroscedasticidade de painel:** os resíduos precisam ter variância constante em ambos os sentidos, isto é, entre unidades *i* e entre períodos temporais *t*. Há heteroscedasticidade de painel quando resíduos têm variância constante ao longo do tempo para cada unidade (*within unit*), mas inconstante entre as unidades (*across units*). Dito de outra forma, cada unidade terá sua própria variância de resíduo. Ela pode decorrer de má especificação do modelo ou quando um ou dois casos não se encaixarem bem na especificação.

(5) **Correlação contemporânea de erros:** o erro de um caso está correlacionado com o erro de outros casos para o mesmo momento no tempo. Se, por exemplo, um choque externo incide em um ano e afeta várias unidades ao mesmo tempo, os resíduos das unidades neste período terão variância diferente dos demais períodos.

(6) **Estruturas complexas de dependência:** há mais casos especiais que, de alguma forma, impõem dependência entre observações. Um dos mais explorados em estudos observacionais é a correlação espacial ou geográfica entre unidades. Ou seja, a contiguidade ou proximidade entre dois países os expõe a mais eventos em comum e, por não ser algo aleatório, pode levar a vieses (Fortin-Rittberger, 2013).

De modo geral, testes como o de Pesaran (2004) oferecem evidência se há dependências entre unidades tais como as descritas nos itens (4) a (6), embora haja variações especializadas para diferentes tipos de painel e de estrutura de dependência (ver Croissant e Millo 2008, p. 28-31; 2018, cap. 10).

Problemas espaciais em modelos POLS podem exigir uso de estimadores FE, RE ou FD. Também se recomenda, para qualquer estimador, a revisão do conjunto de variáveis, no intuito de contemplar novos fatores que possam abarcar a heterogeneidade entre as unidades; além da aplicação de erros-padrão robustos.

Concluindo, é útil lembrar que a probabilidade de ocorrência desses problemas segue o tipo de painel em questão. Painéis do tipo T>N tendem a sofrer mais de problemas decorrentes de heterocedasticidade e correlação serial de erros, ao passo que painéis onde N>T padecem frequentemente de heterogeneidade não-mensurada entre os casos.

A Figura 4 resume a lista de pressupostos e violações discutida nesta seção.

Figura 4 - Resumo dos pressupostos MQO e problemas dinâmicos e temporais

	Pressuposto	Formalização	Por que violá-lo é um problema?
P.1	Relação linear	$Y = \alpha + \beta X + \varepsilon$	Viés
P.2	Ausência de autocorrelação, Amostra aleatória	$Corr(\varepsilon i, \varepsilon j) = 0$ $i \neq j$	Viés
P.3	Ausência de colinearidade perfeita	$Corr(X1,X2) \neq \pm 1$	Impossível estimar $\hat{\beta}$; Viés; Infla Erro-Padrão($\hat{\beta}$), p-valor
P.4	Exogeneidade	$E(\varepsilon \mid x) = 0$	Viés
P.5	Homoscedasticidade	$Var(\varepsilon \mid x) = \sigma^2$	Invalida Erro-Padrão($\hat{\beta}$), p-valor
P.6	Distribuição normal dos erros	$\varepsilon \sim Normal(0, \sigma^2)$	Invalida Erro-Padrão($\hat{\beta}$), p-valor

Normalmente afetam:

	Problemas dinâmicos
1	Correlação serial dos resíduos
2	Não-estacionariedade
	Problemas espaciais
3	Heterogeneidade
4	Heteroscedasticidade de painel
5	Correlação contemporânea dos erros
6	Estruturas complexas de dependência

Fonte: elaborado pelos autores.

4. Aplicação

Nesta seção, reproduzimos os dados de Fernandes e Fernandes (2017) para ilustrar como se faz uma análise de painel. Embora seja costumeiro utilizar bases clássicas dos pacotes estatísticos para ensino (e.g.: "Grunfeld", "EmplUK", etc.), consideramos esta uma valiosa oportunidade para um *replication* que permita incentivar ao mesmo tempo o aprendizado e a cultura de transparência e reprodutibilidade nas ciências sociais (King, 1995).

O principal objetivo do artigo "A importância do crescimento econômico local na escolha do chefe do Executivo no Brasil" é analisar o impacto do crescimento econômico local sobre o percentual de votos recebidos pelo candidato incumbente nas eleições presidenciais e municipais brasileiras entre 2000 e 2010.

Para responder à pergunta de pesquisa, Fernandes e Fernandes (2017) mobilizam 17 variáveis independentes para explicar a variável dependente: a proporção de votos recebida pelo candidato mandatário. Conforme os autores: "Estimamos a relação por meio de técnicas de dados em painel que permitem analisar se as idiossincrasias não observáveis afetam ou não a estimação. Para isso, usamos os três modelos tradicionais de painel estático: *painel empilhado, efeitos aleatórios (EA) e efeitos fixos (EF)*" (Fernandes e Fernandes,

2017, p. 660). Para replicar os dados do trabalho, utilizamos o *software R*, *RStudio* e o pacote *plm* (Croissant e Millo, 2008). Para facilitar a compreensão do passo a passo computacional, reportaremos os scripts ao longo do texto.

Para os testes a seguir, é útil já ter preparado no ambiente do *RStudio* os quatro tipos de especificação a serem comparadas: POLS, FE, RE e FD. Assim, é necessário, em primeiro lugar, carregar o banco de dados e adequá-lo para execução dos modelos.

```
##### ABRINDO BANCO DE DADOS ############

library(haven) # pacote para ler dados do Stata
library(plm) # pacote para executar os modelos de dados em painel

BANCO <- read_dta("fernandes_2017.dta") # Arquivo da base de dados no diretório

BANCO <- pdata.frame(BANCO, index = c("codibge", "ano")) # aqui o banco é convertido
ao formato pdata.frame
# para executar os modelos. No 'index', São acrescentadas as dimensões espacial e
temporal dos dados.
# Nesse caso, 'codibge' representa a dimensão espacial (cód. do munic.) e 'ano' o
ano da eleição.
```

Aplicando o comando *pdim()* podemos ver que a base de dados possui N= 5.565 municípios, e T=2-6.[20] Há, ao todo, 28.945 observações. Ou seja, estamos diante de um painel curto (N>T) e não-balanceado. Ao transformar o banco em um objeto do tipo *pdata.frame*, é criado um índice contendo as dimensões espacial e temporal, o que viabiliza a estimação dos modelos.

20 Note que o período observado é "ano eleitoral". As observações começam em 2000 e vão até 2010, registrando dados a cada dois anos. Por esse motivo a base possui T=6 e não T=10.

```
##### FÓRMULA FERNANDES 2017 ##################

form <- as.numeric(fracaovotos) ~ cresc+crescuf+crescbr+lpibreal+lpibuf+
  lpibbrasil+prefeitobasepresidente+persaude+lpop+leec+
  lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespes

##### POOLED OLS ###############################

POLS <- plm(form, data = BANCO, model = "pooling")  # modelo pooled

##### EFEITOS FIXOS ############################

mode_fe <- plm(form, data = BANCO, model = "within") # O modelo de EF é executado
com o model = 'within'

##### EFEITOS ALEATÓRIOS #######################

mode_re <- plm(form, data = BANCO, model = "random") # o modelo de EA é executado
com o model = "random".

##### PRIMEIRAS DIFERENÇAS #####################

mode_fd <- plm(form, data = BANCO, model = "fd") # o modelo de FD é executado com o
model = "fd".
```

Executamos quatro modelos: POLS, FE, RE e FD. O modelo de POLS é executado por meio da especificação em *model* = "*pooling*" na função *plm*.[21] Do mesmo modo, os modelos FE, RE e FD também são especificados em *model*.

Qual dos modelos acima devemos utilizar? Para respondermos, é necessário verificarmos os pressupostos de cada um.

4.1 Pooled OLS?

A primeira e mais simples escolha é POLS, que aplica uma única linha de regressão, com o mesmo intercepto, para toda a população. Assim, para julgar a adequabilidade desta especificação, precisamos estimar o nível de homogeneidade de população, e se o nosso conjunto de regressores esgota ou não essa heterogeneidade. A relação entre votos em incumbentes e crescimento econômico (e outras variáveis de controle) é similar nos 5.565 municípios? Ou há variação significativa entre as unidades, mais do que foi possível explicar apenas pelas nossas variáveis? Para decidir, pode-se recorrer a (a) conhecimento contextual, (b) análise gráfica e (c) testes.

(a) O conhecimento de causa do pesquisador é o primeiro teste de plausibilidade. O quanto sabemos sobre o objeto estudado por vezes já é informativo sobre o grau de heterogeneidade ou homogeneidade entre unidades. Por exemplo, segundo a literatura, a dinâmica eleitoral na capital paulista é comparável ao que ocorre em Serra da Saudade (MG) que possui população de 781 habitantes? Assim, consultar a teoria sobre o assunto pode fornecer pistas sobre a plausibilidade de um *pooling*.

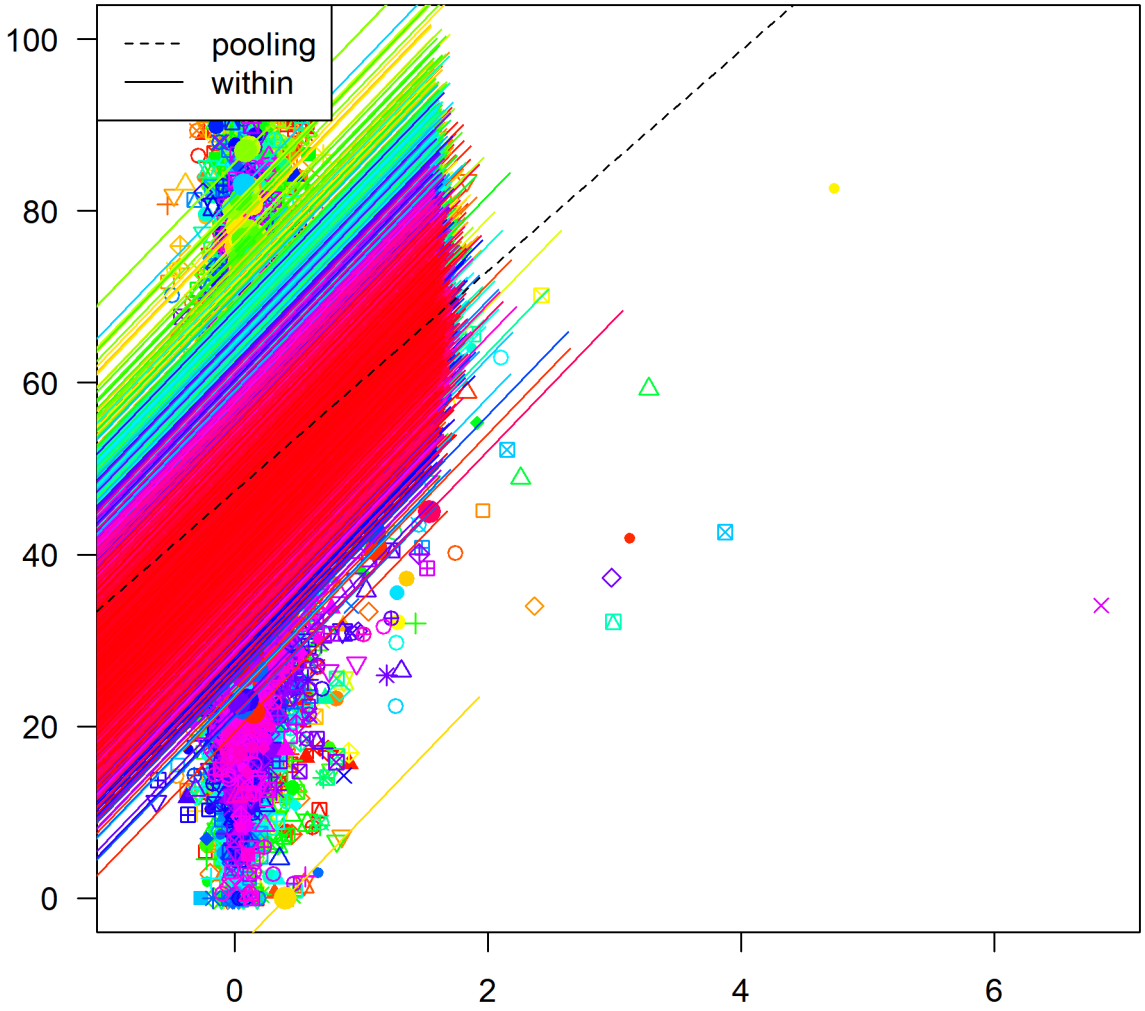
21 O modelo *pooled* também pode ser executado por meio da função *lm*. Entretanto, para execução dos testes para verificação de pressupostos é necessário que o banco de dados esteja no formato produzido via *plm*.

(**b**) Os gráficos são úteis para análise exploratória, principalmente quando N é pequeno. Para visualizar se podemos abarcar todos os municípios por uma única reta de regressão, podemos comparar, em um mesmo gráfico de dispersão, uma linha única aplicada à toda a população vs. linhas separadas para cada município. Exemplificamos a seguir com um gráfico bivariado com a principal variável independente, "crescimento econômico" (eixo horizontal), e a variável dependente "fração de votos no incumbente" (eixo vertical).

```
##### COMPARAÇÃO GRÁFICA INTERCEPTOS POLS VS FE #######

mode_eff <- plm(fracaovotos ~ cresc,
                data = BANCO, index = c("codibge", "ano", group = "codibge"),
                model = "within")
plot(mode_eff)
```

Figura 5 - Comparação Pooled OLS e FE para modelo bivariado fração de votos no incumbente x crescimento econômico



Fonte: elaborado pelos autores, baseado em dados de Fernandes e Fernandes (2017).

A solução *pooling* (linha pontilhada) propõe uma linha situada aproximadamente no meio das observações, ao passo que o modelo *within* (linhas sólidas coloridas) melhor representa a variação dos dados e indica que há larga diferença entre os interceptos de cada município.

Outra ferramenta de análise gráfica é o *plot* dos resíduos do modelo final. Como POLS pressupõe homogeneidade, seus resíduos devem ser semelhantemente aleatórios para todos os casos. Podemos agrupar os resíduos da regressão por município e por ano como forma de observar se a variância no termo de erro foi mais pronunciada para certos municípios ou períodos, como exemplificado adiante.

```
##### ANALISANDO OS RESIDUOS POOLED OLS ###########

library(broom)
library(ggplot2)

POLS_lm <- lm(fracaovotos ~ cresc+crescuf+crescbr+lpibreal+lpibuf+
               lpibbrasil+prefeitobasepresidente+persaude+lpop+leec+
               lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespes,
             data = BANCO, na.action = "na.exclude")
# é necessario rodar o modelo POLS com lm para acrescentar
# os resíduos do modelo ao data.frame original

base <- augment(POLS_lm, data=BANCO)
# com o comando 'augment' do pacote broom, podemos juntar
# os resíduos do modelo com as variaveis do banco original.

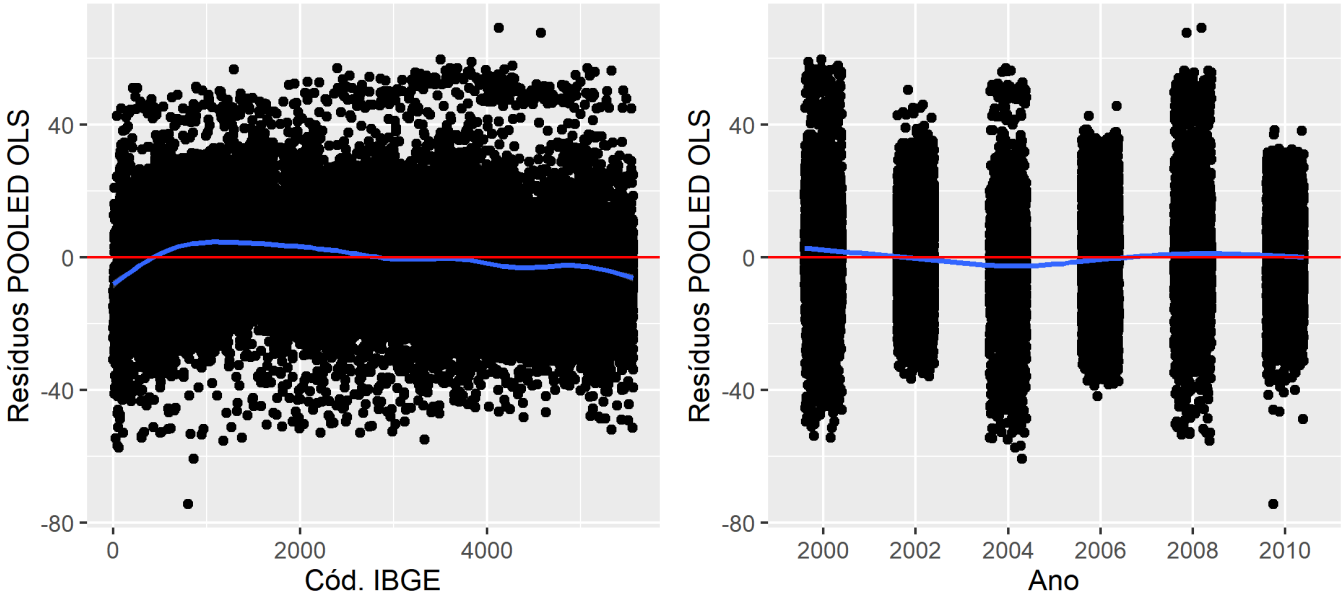
                  #### Município ####

ggplot(base, aes(x=as.numeric(as.factor(codibge)), y=.resid))+
  geom_point()+
  geom_smooth()+
  geom_hline(yintercept = 0, color="red") + ylab("Resíduos POOLED OLS") +
  xlab("Cód. IBGE")
                  ####   Ano  ####

base$ano <- as.numeric(base$ano)

ggplot(base, aes(x=jitter(ano), y=.resid))+
  geom_point()+
  geom_smooth()+
  geom_hline(yintercept = 0, color="red") + ylab("Resíduos POOLED OLS") +
  xlab("Ano") + scale_x_discrete(limits=c("2000", "2002", "2004",
                                          "2006", "2008", "2010" ))
```

Figura 6 - Resíduos do modelo POLS agrupados por município e ano



Fonte: elaborado pelos autores, com base em Fernandes e Fernandes (2017).

Os resíduos dos municípios não são semelhantes: há algum fator sistematicamente diferenciando os municípios de código IBGE inferior a 3.000 dos demais com numeração mais alta. A variância entre anos eleitorais também é diferente. O motivo rápido se percebe: a variância é maior para anos de disputas municipais (2000, 2004 e 2008) e menor para as presidenciais (2002, 2006 e 2010). Esta situação ilustra o problema do *pooling* e da heterogeneidade, pois estamos tentando incluir num mesmo modelo preditivo eleições para prefeito e presidente. Esses resultados nos motivam a considerar a inclusão de *dummies* para cada ano ou tipo de eleição.

(c) o pacote *plm* permite implementar diferentes testes com o objetivo de verificar, sob a forma de hipótese nula/alternativa, se há efeitos individuais significativos. Podem ser aplicados o teste-F e o teste Breusch-Pagan LM. O teste-F compara se há diferenças significativas entre o POLS e o modelo FE, sendo a hipótese nula de que POLS é superior. Como podemos observar pelo resultado do p-valor (<0,05), o modelo FE é significativamente superior.

```
##### F-TEST POLS ################################

pFtest(mode_fe, POLS) # F-test com o modelo de efeitos fixos e Pooled
##
##  F test for individual effects
##
## data:  form
## F = 1.4214, df1 = 5548, df2 = 20786, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Já o teste Breusch-Pagan LM nos ajuda a escolher entre POLS e RE.[22] A hipótese nula do teste prevê que a variância específica por município é zero, o que recomendaria POLS. Conforme o resultado do teste, H0 é rejeitada e verifica-se então que POLS pode ser descartado (p-valor < 0,05).

```
##### BREUSCH-PAGAN TEST ########################

plmtest(POLS, type="bp", effect = "individual") # teste de Breusch-Pagan + efeitos
individuais
##
##  Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels
##
## data:  form
## chisq = 107.15, df = 1, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Os comentários sobre efeitos individuais também se aplicam aos efeitos temporais. Seguindo a mesma lógica, trata-se de verificar se cada ano do estudo possui idiossincrasias específicas que fazem com que as observações naquele intervalo possuam alguma heterogeneidade peculiar. Isso corresponde à noção de choques no tempo. Assumindo tanto efeitos temporais como individuais, temos os chamados efeitos "*two ways*". Nos testes descritos acima, a presença desses efeitos pode ser testada modificando o argumento "*effect*", como exibido abaixo:

```
plmtest(POLS, type="bp", effect = "twoways") # teste de Breusch-Pagan + efeitos indi-
viduais e temporais
##
##  Lagrange Multiplier Test - two-ways effects (Breusch-Pagan) for
##  unbalanced panels
##
## data:  form
## chisq = 2762.9, df = 2, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Diante da análise gráfica e do teste de Breusch-Pagan, sabemos que os modelos testados apresentam heterogeneidade espacial e também temporal. Assim, executamos novamente os modelos adicionando novas *dummies*. No estudo original, os autores adicionam ao modelo *dummies* para os anos e para diferenciar eleições municipais das presidenciais (Fernandes e Fernandes 2017, p.665). Embora na maioria das situações as *dummies* anuais sejam o caminho padrão para capturar heterogeneidade temporal, as especificidades desta base nos recomendam usar apenas a *dummy* de tipo de eleição.[23] Ela deverá aprimorar principalmente os modelos POLS e FD, enquanto que, para FE e RE, alteramos o efeito para "*two ways*", o que introduzirá efeitos por município e ano, implicando que *dummies* temporais como essa e variáveis como o PIB nacional possivelmente serão omitidas por colinearidade. Com a adição dessas variáveis, é possível controlar os efeitos heterogêneos identificados anteriormente.

22 O teste de Breusch-Pagan LM é um dos mais tradicionais, porém o comando *plmtest* pode calcular extensões do mesmo. Essas variações são úteis a depender das configurações dos dados em questão (*balanced vs unbalanced, short vs long*). Ver Croissant e Milo (2018, p. 86).

23 Para esses dados, as *dummies* para cada ano terão, muito provavelmente, colinearidade acentuada com algumas das demais variáveis, principalmente variáveis agregadas a nível nacional (como o PIB nacional, que terá um valor fixo a cada ano) e mesmo a nova *dummy* para eleições municipais.

```
##### MODELOS COM VARIÁVEL ELEIÇÃO ###############

# variavel para indicar se a eleição é pres (0) ou mun (1)
BANCO$tip_ele <- ifelse(BANCO$ano == 2000 |
                        BANCO$ano == 2004 |
                        BANCO$ano == 2008, 1, 0)

# formula com dummies para tipo de eleição
form_d <- as.numeric(fracaovotos) ~ cresc+crescuf+crescbr+lpibreal+lpibuf+
  lpibbrasil+prefeitobasepresidente+persaude+lpop+leec+
  lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespes+
  tip_ele

##### POOLED OLS C DUMMIES ###############

POLS_dic <- plm(form_d, data = BANCO, model = "pooling")


##### FD C DUMMIES ########################

mode_fd_dic <- plm(form_d, data = BANCO, model = "fd")


##### FE TWO-WAYS #########################

mode_fe_2w <- plm(form_d, data = BANCO, model = "within", effect = "twoways")


##### RE TWO-WAYS #########################

mode_re_2w <- plm(form_d, data = BANCO, model = "random", effect = "twoways", random.
method = "walhus")
```

4.2 Efeitos fixos ou aleatórios?

Descartado POLS, podemos contemplar FE ou RE. Como já dissemos, nesta escolha é importante avaliar tanto a natureza dos dados quanto a intensidade da correlação entre μ_i e X. Para tanto, podemos recorrer ao (a) conhecimento contextual, (b) relevância ou indispensabilidade de certas variáveis de interesse, e (c) teste de Hausman.

(a) Como explicado na seção 3.2, o pesquisador deve responder à seguinte pergunta sobre a origem de seus dados: posso considerá-los como uma amostra aleatória de uma população mais ampla? Pensando em termos de exogeneidade, temos bons motivos para esperar que alguns municípios, simplesmente por ser, digamos, Salvador ou Sirinhaém, tenham peculiaridades não-observadas (e.g.: hegemonia histórica de certo partido) que se correlacionam com as variáveis selecionadas?

(b) Uma das limitações de FE (também de FD) é impedir o teste de elementos invariantes no tempo. Quando há interesse substantivo em testar variáveis deste tipo, o uso de RE pode ser contemplado, sempre levando em conta o *tradeoff* entre eficiência e viés. O modelo de Fernandes e Fernandes (2017) que replicamos não usa variáveis fixas no tempo, porém se desejássemos ir além e testar, por exemplo, *dummies* de controle para

as regiões brasileiras (Norte, Nordeste, Centro-Oeste, Sudeste e Sul), não poderíamos fazê-lo para FE.

(c) O comando *phtest* aplica o teste de especificação de Hausman (1978), comparando um modelo FE com RE, sob a hipótese nula de que ambos são consistentes. A hipótese nula será rejeitada quando se verificar correlação entre as heterogeneidades individuais não-observadas e as variáveis independentes. Em nosso caso, rejeita-se a hipótese nula que ambos FE e RE são consistentes. Sabemos portanto que há correlação entre a heterogeneidade individual não-observada e os regressores, de modo que apenas FE terá estimadores consistentes. É importante ter em mente que o teste de Hausman, mesmo sendo o padrão para discriminar entre FE e RE, pressupõe modelos bem especificados, livres de viés e homocedásticos.[24]

```
##### HAUSMAN TEST ##############################

phtest(mode_fe_2w, mode_re_2w) # teste de Hausman (modelo fixo e aleat.)
##
##  Hausman Test
##
## data:  form_d
## chisq = 851.98, df = 15, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
phtest(form_d, data=BANCO, method="aux", vcov=vcovHC) # versão robusta do teste
##
##  Regression-based Hausman test, vcov: vcovHC
##
## data:  form_d
## chisq = 810.83, df = 18, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

4.3 Lidando com problemas temporais

Como exposto na seção 3.4, os problemas temporais são tipicamente (a) não-estacionariedade e (b) correlação serial dos resíduos. O primeiro refere-se às variáveis brutas no banco de dados, e o segundo aos resíduos após a aplicação de modelos. Dada essa sequencialidade, deve-se primeiro verificar a estacionariedade.

(a) Painéis longos podem se beneficiar de testes como ACF, PACF, ADF ou *Panel Unit Root Tests* para verificar a estacionariedade. Como nosso painel é curto (T=6), esses testes seriam pouco informativos. Se ainda assim quisermos verificar de alguma maneira se os dados são estacionários ou não, podemos procurar de forma genérica por tendências não reversíveis de crescimento ou retração nas variáveis ao longo do tempo. Abaixo fazemos essa inspeção visual para a variável dependente "fração de votos", mostrando os valores por município e uma média geral de cada ano. Há alguma margem para esperarmos uma tendência de crescimento ao longo do tempo na fração de votos do incumbente.

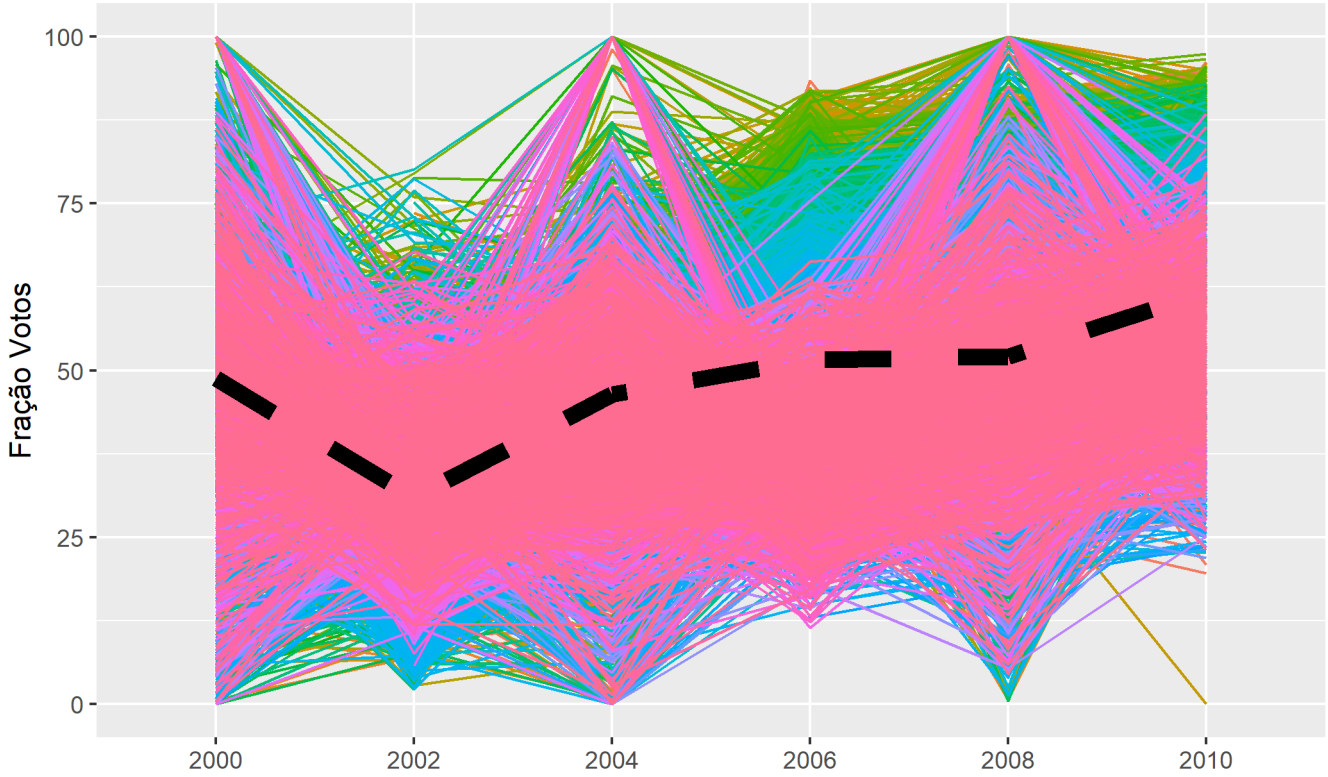
24 Se os requisitos de ausência de *misspecification*, independência entre termo de erro e variáveis independentes, ausência de correlação serial e de heterocedasticidade não forem satisfeitos, recomenda-se, além inicialmente da revisão do modelo, o uso de formas robustas do teste. No *plm*, isso se opera com os argumentos *phtest(..., method="aux")* ou ainda *phtest(..., method="aux", vcov=vcovHC)* para matrizes variância-covariância customizadas.

```
##### VERIFICAR ESTACIONARIEDADE #############

library(dplyr)
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:plm':
##
##     between, lag, lead
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
banco_est <- BANCO %>%
  group_by(ano) %>%
  mutate(mediafvotos = mean(fracaovotos, na.rm=T)) #banco com a media dos votos

ggplot(banco_est, aes(x=ano, y=fracaovotos, group=codibge, color=codibge))+
  geom_line() + geom_line(aes(x=ano, y=mediafvotos, group=1),
  linetype="dashed", size=3, color="black")+ ylab("Fração Votos") +
  xlab("") + guides(color="none")
```

Figura 7 - Fração de votos ao longo do tempo (municípios e média geral)



Fonte: elaborado pelos autores, com base em dados originais de Fernandes e Fernandes (2017).

Se há tendência, recomenda-se a inclusão de uma *time trend*, geralmente uma variável T que assume os valores de 1 para o primeiro ano, 2 para o segundo, e assim por diante.

(b) Até então nossa seleção entre os estimadores POLS, FE, RE e FD se pautou em quão bem cada um lidava com a heterogeneidade não-observada. Agora, o sucesso em eliminar correlações temporais também se torna importante para eleger um estimador. Como explicado em 2.4, há diversos testes para correlação serial que podem ser aplicados aos resíduos, a depender do tipo de painel (curto ou longo) e estimador empregado. O pacote *plm* oferece atualmente oito testes para correlação serial.[25] Como o teste de Hausman já nos aconselhou abandonar RE, nós utilizaremos apenas dois testes, voltados para FE e FD.

Para verificar a presença de correlação serial em modelos de efeito fixo, utilizamos a função *pwartest*. Como podemos observar no *output*, a hipótese alternativa do teste é de que existe correlação serial. O p-valor foi menor que 0,05, indicando a rejeição da hipótese nula (não existe correlação serial). Ou seja, conforme o resultado do teste, o modelo de efeitos fixos apresenta correlação serial.

```
##### TESTAR CORR. SERIAL PARA FE ########################

pwartest(mode_fe_2w) #pwartest para o modelo FE
##
##  Wooldridge's test for serial correlation in FE panels
##
## data:  mode_fe_2w
## F = 71.527, df1 = 1, df2 = 20801, p-value < 2.2e-16
## alternative hypothesis: serial correlation
```

Já o *pwfdtest* é utilizado para verificar a presença de correlação serial nos modelos FD. A hipótese nula é de que não existe correlação serial relacionada aos erros. Dado o p-valor (< 0,05), rejeitamos a hipótese nula. Ou seja, o modelo FD apresenta problemas de correlação serial.

```
##### TESTAR CORR. SERIAL PARA FD ######################

pwfdtest(mode_fd_dic) #pwfdtest para o modelo FD
##
##  Wooldridge's first-difference test for serial correlation in panels
##
## data:  mode_fd_dic
## F = 6004.9, df1 = 1, df2 = 15307, p-value < 2.2e-16
## alternative hypothesis: serial correlation in differenced errors
```

Os resultados dos testes sugerem que tanto FE quanto FD abrigam correlação serial nos termos de erro. Da mesma forma, a inspeção dos dados brutos indicam tendência ascendente, ou seja, não-estacionariedade. Para remediar a situação, poderíamos contemplar a inclusão

25 Os seis testes não cobertos neste artigo são: teste genérico de Wooldridge para efeitos não-observados, aplicável a POLS (*pwtest*); teste de Breusch-Godfrey/Wooldridge para correlação serial, preferencialmente aplicável a RE, porém também a FEs longos (*pbgtest*); teste de Durbin-Watson, também preferencialmente a RE (*pdwtest*) e sua generalização por Bhargava, Narendranathan e Franzini (*pbnftest*); teste localmente robusto de Bera, Sosa-Escudero e Yoon para correlação serial ou efeitos aleatórios (*pbsytest*); teste de Baltagi e Li para correlação serial (*pbltest*).

de uma *time trend* (porém, como vimos em 2.2, essas não são estimáveis por FE ou FD), LDV ou emprego de erros-padrão robustos à correlação serial. Como foi visto anteriormente, problemas temporais tendem a ser mais preocupantes conforme T>N. Em nosso caso, tendo em vista que N é muito superior a T, e para manter a parcimônia do exemplo, vamos nos contentar em preservar as variáveis atuais e ao final empregar robustificação.

4.4 Lidando com problemas espaciais

Uma vez que já lidamos com a heterogeneidade não-observada, é de se esperar que os problemas espaciais restantes digam respeito à heteroscedasticidade de painel e à correlação contemporânea dos erros. A mesma inspeção visual aplicada na seção 4.1, agrupando resíduos por município, também informa esta decisão. A Figura 6 já mostrou que a variância dos resíduos entre municípios é sistemática.

Além da inspeção visual, podemos recorrer a testes específicos. O pacote *plm* oferece um comando geral para testar *cross-sectional dependence* (*pcdtest*), cabendo ao usuário escolher o tipo de teste, segundo as características do painel. Como temos T curto e N longo[26], o teste mais apropriado é o proposto por Pesaran (2004).

```
##### TESTAR HETEROSC. DE PAINEL  ##################################

pcdtest(mode_fe_2w, test = "cd") # modelo Efeitos Fixos
##
##  Pesaran CD test for cross-sectional dependence in panels
##
## data:  as.numeric(fracaovotos) ~ cresc + crescuf + crescbr + lpibreal +     lpi-
buf + lpibbrasil + prefeitobasepresidente + persaude +     lpop + leec + lheu + lses
+ laseps + ldesporc + ldespcor +     linvest + ldespes + tip_ele
## z = -0.66933, p-value = 0.5033
## alternative hypothesis: cross-sectional dependence
```

Os resultados mostram que nosso modelo FE *two ways* está livre de dependência contemporânea, tendo apenas problemas de correlação serial. Devemos aplicar erros-padrão robustos contra correlação serial. Podemos acessar diferentes opções de robustez no *plm*, a depender do tipo de painel, da forma de agrupamento dos dados e do tipo de problema a remediar (ver Croissant e Millo 2018, cap. 5; Henningsen e Henningsen 2019, pp. 382-384). Em nosso caso, como queremos lidar com a correlação serial de um FE com efeitos *two ways*, aplicamos a matriz *vcovDC* para estimar erros-padrão robustos:

```
##### LIDANDO COM CORR. SERIAL ############

library(lmtest)
mode_fe2w_rbst <- coeftest(mode_fe_2w,
                           vcov = vcovDC)
# adicionando erros robustos 'double cluster'
```

26 Breusch-Pagan LM, acionado pelo argumento test="lm" é recomendado para T longo e N curto; scaled-Breusch-Pagan LM, "sclm", para T e N longos; "rho" e "absrho" fornecem os valores do coeficiente de correlação entre resíduos para pares de observações (ver Croissant e Millo, 2018, seção 4.4).

4.5 Resultados

Apresentamos os resultados de todos os modelos na Tabela 1. Para permitir a comparação dos erros-padrão originais do modelo FE com os robustos, colocamos ambos em colunas separadas na tabela.

Sendo nosso interesse principal pedagógico e não substantivo, apresentamos apenas uma versão resumida da lista de variáveis. Assim, das mais de 20 variáveis do modelo original, conservamos 11 para a tabela.

Tabela 1 - Resultados da replicação

	Fração de Votos				
	POLS	**RE (2W)**	**FE (2W)**	**FE (2W, Robust.)**	**FD**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Cresc. Munic.	1.683***	1.098***	1.148*	1.148	0.190
	(0.608)	(0.040)	(0.652)	(1.517)	(0.669)
Cresc. UF	0.095***	0.086***	-0.058*	-0.058	-0.144***
	(0.027)	(0.002)	(0.030)	(0.163)	(0.033)
Cresc. BR	3.660***	3.709***			4.231***
	(0.090)	(0.088)			(0.091)
Prefeito Base Presid.	1.211***	0.742***	1.289***	1.289***	1.136***
	(0.208)	(0.014)	(0.235)	(0.409)	(0.254)
Perc. Orç. Saúde	0.040***	0.121***	0.044*	0.044	0.058***
	(0.014)	(0.001)	(0.023)	(0.035)	(0.019)
Pop. Munic. (log)	0.302	-0.415***	-6.037***	-6.037**	-3.360*
	(0.258)	(0.018)	(1.449)	(2.609)	(1.824)
Desp. Educ. Cult. (log)	6.917***	6.461***	1.543***	1.543	0.162
	(0.358)	(0.025)	(0.525)	(1.284)	(0.553)
Desp. Invest. (log)	2.245***	1.983***	1.430***	1.430	1.644***
	(0.136)	(0.009)	(0.165)	(0.903)	(0.180)
Desp. Pessoal (log)	-1.424***	-0.335***	3.304***	3.304	1.167*
	(0.493)	(0.035)	(0.635)	(2.264)	(0.698)
Eleição Municip.	1.074***	0.828***			-0.915***
	(0.277)	(0.246)			(0.264)
Intercepto	102.026**	152.778***			-4.350***
	(43.701)	(37.792)			(0.373)
N Obs.	**26.352**	**26.352**	**26.352**		**20.803**
R2	**0.321**	**0.319**	**0.039**		**0.254**
R2 Ajust.	**0.320**	**0.318**	**-0.219**		**0.253**
Estat. F	**691.129*** (df = 18; 26333)**	**12.309.190*** **	**55.646*** (df = 15; 20783)**		**392.979*** (df = 18; 20784)**

Fonte: elaborado pelos autores, com base em Fernandes e Fernandes (2017). O valor de theta para o modelo de efeitos aleatórios two ways foi θ = 0,11.

Os testes de especificação nos conduzem a ter mais confiança nos resultados obtidos por meio de FE. Não obstante, a comparação de todos os modelos se revela instrutiva. Os resultados de POLS e RE tendem a estar mais próximos um do outro para muitos dos regressores, enquanto que as estimativas de FE e FD são mais semelhantes entre si. Isso é esperado pois, como vimos, os dois últimos removem a heterogeneidade individual de forma semelhante. Assim, ler os resultados de POLS/RE até FE/FD é uma maneira de verificar como se comportam as estimativas conforme vamos controlando mais e mais as idiossincrasias municipais; o "efeito líquido", por assim dizer, das variáveis, descontadas todas as peculiaridades dos casos.

Em algumas variáveis, a diferença se manifesta na troca de sinal. Por exemplo, os modelos POLS e RE afirmam que haveria efeitos positivos da taxa de crescimento econômico da UF sobre a fração de votos, ou negativos de despesas com pessoal. Temos um caso de viés pois, controlando pelas heterogeneidades de cada município por meio de FE e FD, vemos que seus impactos seriam na verdade o inverso.

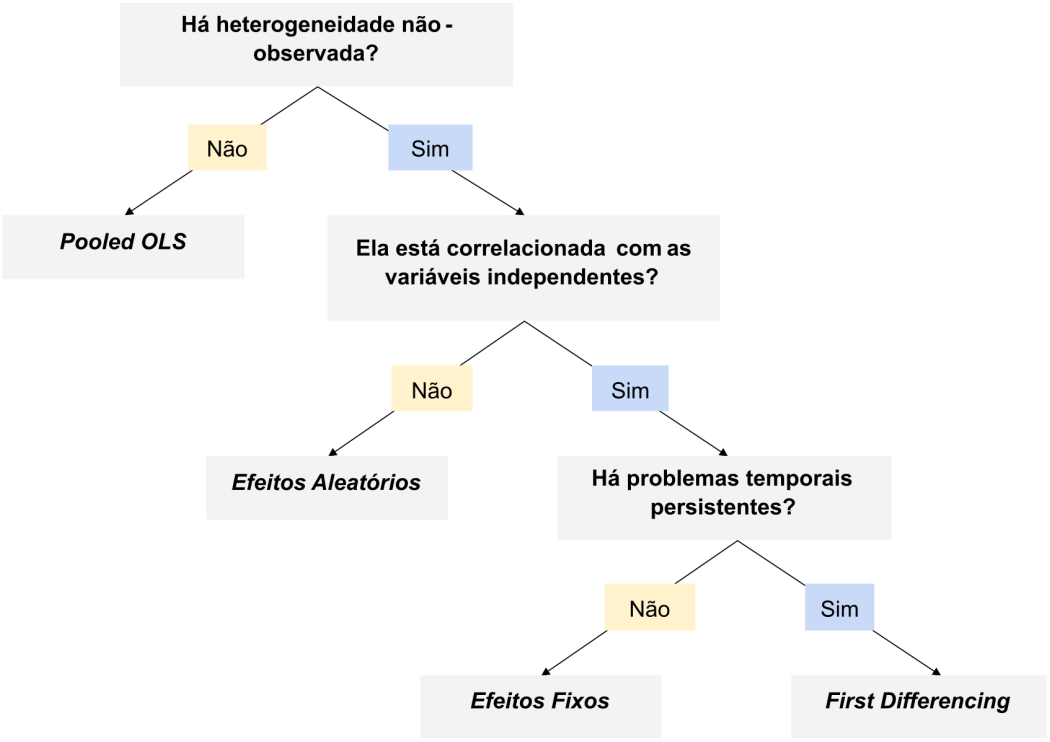
Para outros regressores, percebe-se uma mudança na força da relação. Às vezes a magnitude do efeito se torna mais débil (i.e.: despesa em educação e cultura, despesa em investimento), ocasionalmente mais forte (*dummy* para mesma base prefeito-presidente, população), e por vezes estável (percentual do orçamento de saúde). Ademais, enquanto POLS e RE tendem a acusar mais significância estatística, FD e principalmente FE com erros padrão robustos são mais conservadores. Notadamente, a principal variável de interesse do estudo, crescimento econômico municipal, não retém efeito significativo sobre os votos no incumbente.

Por fim, deve ser apontado que os interceptos POLS e RE apresentam valores irrealistas pelo fato de os dados não terem sido centralizados e que, como antecipamos, FE *two ways* omitiu *dummies* temporais (tipo de eleição) e aquelas fixas ao ano (PIB BR) por colinearidade.

4.6 Resumo do procedimento

Para concluir este tutorial, podemos representar o processo de escolha de um modelo como uma árvore de decisão similar à Figura 8.

Figura 8 - Como escolher o melhor estimador?



Fonte: elaborado pelos autores.[27]

Esquematicamente, primeiro se avalia em que medida existe heterogeneidade não-observada entre as entidades da base de dados. Isso pode ser feito a partir do conhecimento contextual do problema, da análise gráfica dos interceptos e pelo emprego de testes específicos (F e Breusch-Pagan LM). Uma vez descartada a opção POLS, devemos estimar se existe correlação entre a heterogeneidade não-observada e o conjunto de variáveis explicativas, por exemplo através do teste de Hausman. Inexistindo correlação, justifica-se o uso de RE. Por outro lado, na presença de correlação entre os erros e as variáveis independentes, podemos cogitar a utilização de FE ou FD. A opção entre esses dois últimos deve ser equilibrada por considerações de problemas temporais do modelo, bem como tamanho da base. Ao final do processo, testes para dependências temporais e espaciais são necessários para verificar a necessidade de erros-padrão robustos.

Este esquema é uma simplificação para fins didáticos e enfatiza as principais decisões que, se ignoradas, poderiam gerar viés ou ineficiência. Porém, como argumentamos ao longo do artigo, considerações substantivas também devem guiar essa escolha.

5. Conclusão

Este artigo apresentou uma introdução à lógica de dados de painel. Mostramos os principais benefícios advindos da análise de informações longitudinais e descrevemos os problemas mais frequentes que surgem quando a planilha de dados reúne, simultaneamente, dimensões de espaço e tempo. Depois disso, demonstramos como implementar uma análise de regressão com dados de painel no R. Por se tratar de um trabalho com finalidade pedagógica, compartilhamos ainda todos os dados e scripts de modo que alunos, professores e pesquisadores podem facilmente adaptar nossas rotinas computacionais aos seus interesses substantivos.

27 Este esquema é inspirado em um original mais complexo, que os autores creditam à professora Claudia Schwirplies (Uni HH).

No início deste artigo, apontamos como o uso de dados de painel na CPRI brasileira ainda é incipiente. Assim, acreditamos ser relevante explicar o que pode ser feito para mudar esse quadro. Entre as diferentes alternativas disponíveis, destacamos o seguinte: (1) reformular a estrutura curricular dos cursos de graduação e pós-graduação com o objetivo de incluir disciplinas específicas de cálculo, estatística e computação e ofertar disciplinas eletivas sobre como trabalhar com dados longitudinais, tal qual o curso do professor William Greene; (2) na impossibilidade de implementar o item 1, garantir a oferta de cursos especiais de verão, como o *Modeling Dynamics* ministrado pela professora Lorena Barberia (USP), *Modeling Dynamics in Space and Time* ministrado pelos professores Guy Whitten e Lorena Barberia (USP) e Análise de Dados em Painel por Igor Viveiros no curso MQ/UFMG; (3) diversificar, regionalmente, a oferta intermitente de cursos de análise de dados. Na impossibilidade de implementar o item 3, garantir apoio financeiro para que estudantes de mestrado e doutorado, principalmente do Norte e Nordeste, possam frequentar escolas de verão; (4) fomentar o aproveitamento de cursos on-line como créditos currículos válidos durante a pós-graduação; (5) edição, por parte das principais revistas de CPRI, de dossiês especiais sobre métodos de pesquisa, com especial ênfase para artigos intuitivos e pedagógicos, tal qual o volume 142 do Journal of Econometrics. Ainda, reservar, nas melhores revistas de CPRI, espaço para publicação de artigos com ênfase metodológica como Beck e Katz (1995); e (6) incentivar os pesquisadores que se dedicam à divulgação de métodos e técnicas de pesquisa em CPRI a partir de linhas específicas de financiamento em instituições de fomento (CAPES e CNPq).

Sabemos das dificuldades operacionais e logísticas que permeiam cada uma dessas iniciativas. No entanto, considerando os potenciais benefícios que podem ser incorporados em nossos modelos de análise e a limitada utilização de dados longitudinais na pesquisa empírica em CPRI, acreditamos que já passou da hora de levar o tempo a sério (Beck, Katz e Tucker, 1998; Boef e Keele, 2008).

REFERÊNCIAS BIBLIOGRÁFICAS

ACHEN, C. H. (2000), "Why lagged dependent variables can suppress the explanatory power of other independent variables". In: *annual meeting of the political methodology section of the American political science association, UCLA*.

BALTAGI, Badi H. (2005), *Econometric Analysis of Panel Data*. England, John Wiley & Sons Ltd.

BERRY, William D. (1993). *Understanding regression assumptions*. Thousand Oaks, Sage.

BECK, Nathaniel; KATZ, Jonathan N. (1995), "What to do (and not to do) with time-series cross-section data", *American political science review*, v. 89, n. 3: 634-647.

BECK, Nathaniel; KATZ, Jonathan N.; TUCKER, Richard. (1998), "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable", *American Journal of Political Science*, v. 42, n. 4: 1260-1288.

BECK, Nathaniel. (2001), "TIME-SERIES–CROSS-SECTION DATA: What Have We Learned in the Past Few Years?", *Annual Review of Political Science*, v. 4, n. 1: 271–293.

BECK, Nathaniel. (2008), "Time-Series Cross-Section Methods", *in* J. Box-Steffensmeier; H. Brady; D. Collier (Eds.). *The Oxford Handbook of Political Methodology*. New York, OUP: 475–493.

BLALOCK, Hubert M. The presidential address: Measurement and conceptualization problems: The major obstacle to integrating theory and research. *American sociological review*, v. 44, n. 6, p. 881-894, 1979.

DE BOEF, Suzanna; KEELE, Luke. (2008), "Taking time seriously", *American Journal of Political Science*, v. 52, n. 1: 184-200.

BELL, Andrew; JONES, Kelvyn. (2015), "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data", *Political Science Research and Methods*, v. 3, n. 1: 133–153.

BELL, Andrew; FAIRBROTHER, Malcolm; JONES, Kelvyn. (2019), "Fixed and random effects models: making an informed choice", *Quality & Quantity*, v. 53, n. 2: 1051-1074.

BENOIT, K. et al. (2018), "quanteda: An R package for the quantitative analysis of textual data", *Journal of Open Source Software*, v. 3, n. 30: 774.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. (2016), *Time series analysis: forecasting and control*. Hoboken/New Jersey, Wiley.

CLARK, Tom S.; LINZER, Drew A. (2015), "Should I Use Fixed or Random Effects?", *Political Science Research and Methods*, v. 3, n. 2: 399–408.

CROISSANT, Yves; MILLO, Giovanni. (2008), "Panel data econometrics in R: The plm package", *Journal of statistical software*, v. 27, n. 2: 1-43.

CROISSANT, Yves; MILLO, Giovanni. (2018), *Panel Data Econometrics with R*, Wiley.

CERQUEIRA, Daniel et al. (2017), *Atlas da violência 2017*.

DIELEMAN, J. L.; TEMPLIN, T. (2014), "Random-Effects, Fixed-Effects and the within-between Specification for Clustered Data in Observational Health Studies: A Simulation Study", *PLoS ONE*, v. 9, n. 10: e110257.

FÁVERO, L. P. L. (2013), "Dados em painel em contabilidade e finanças: teoria e aplicação", *BBR-Brazilian Business Review*, v. 10, n. 1: 131-156.

FERNANDES, Ivan Filipe de Almeida Lopes; FERNANDES, Gustavo Andrey de Almeida Lopes. (2017), "A importância do crescimento econômico local na escolha do chefe do Executivo no Brasil", *Revista de Administração Pública*, v. 51, n. 4: 653-688.

FINKEL, S. E. (1995), *Causal analysis with panel data* (No. 105). Sage.

FORTIN-RITTBERGER, J. (2013), "Time-Series Cross-Section", *in* H. Best; C. Wolf (Eds.), *The SAGE Handbook of Regression Analysis and Causal Inference*. London, SAGE Publications: 387–408.

GUJARATI, Damodar N. *Basic Econometrics*. McGraw-Hill Companies. New York, 2004.

HAUSMAN, Jerry A. (1978), "Specification tests in econometrics", *Econometrica: Journal of the econometric society*, v. 46, n. 6: 1251-1271.

HENNINGSEN, Arne; HENNINGSEN, Géraldine. (2019), "Analysis of Panel Data Using R", *in* M. Tsionas (Ed.) *Panel Data Econometrics: Theory*. London, Elsevier: 345-396.

HSIAO, C. (2003), *Panel data analysis*, Cambridge University Press.

IBGE. (2017), Pesquisa Nacional por Amostra de Domicílios Contínua. Disponível em: https://www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html?=&t=o-que-e. Acesso em 30-06-2020

KING, Gary. (2001). "Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data", *International Organization*, v. 55, n. 2: 497-507.

KING, Gary. (1995), "Replication, replication", *PS: Political Science and Politics*, v. 28, n. 3: 444-452.

KING, Gary; KEOHANE, Robert O.; VERBA, Sidney. (1994), *Designing social inquiry: Scientific inference in qualitative research*. Princeton University Press.

KENNEDY, Peter. (2003), *A guide to econometrics*. MIT Press.

KRUEGER, James S.; LEWIS-BECK, Michael S. (2008), "Is OLS dead?", *The Political Methodologist*, v. 15, n. 2: 2-4.

LAROCCA, R. (2012). "Gauss–Markov Theorem" in SALKIND, N. J. (ed) *Encyclopedia of Research Design*. Thousand Oaks, Sage, 529-533.

LIJPHART, A. (1971), "Comparative Politics and the Comparative Method", *American Political Science Review*, v. 65, n. 3: 682–693.

MARQUES, Luís D. (2000), "Modelos Dinâmicos com Dados em Painel: revisão de literatura". *Working paper*. Disponível em http://wps.fep.up.pt/wps/wp100.pdf. Acesso em 23-06-2020.

MENARD, Scott. (2002), *Longitudinal research*. Thousand Oaks, SAGE Publications.

MEDEIROS, M.; BARNABE, I.; ALBUQUERQUE, R.; LIMA, R. (2016), "What does the field of International Relations look like in South America?", *Revista Brasileira de Política Internacional*, v. 59, n. 1: e004

MEIRELES, Fernando; SILVA, Denisson; BARBOSA, Rogerio. (2019), *rscielo: A Scraper for Scientific Journals Hosted on Scielo*. R package version 1.0.0. https://CRAN.R-project.org/package=rscielo.

MESQUITA, Rafael. (2018) *Liderança regional em perspectiva comparada: Brasil e Turquia*. Tese de doutorado, UFPE. Disponível em https://repositorio.ufpe.br/handle/123456789/32942. Acesso em 29-05-2021

MUMMOLO, J.; PETERSON, E. (2018). "Improving the interpretation of fixed effects regression results", *Political Science Research and Methods*, v. 6, n. 4: 829-835.

PESARAN, M. H. (2004), "General diagnostic tests for cross section dependence in panels", *CESifo Working Paper Series*, 1229.

PEARL, J. (2000), *Causality: Models, reasoning, and inference*. New York, Cambridge University Press.

SKRONDAL, A., RABE-HESKETH, S. (2008), "Multilevel and Related Models for Longitudinal Data", *in* Leeuw J.., Meijer E. (Eds.) *Handbook of Multilevel Analysis*, New York, Springer: 275-300.

STOCK, James H.; WATSON, Mark W. (2004), *Econometria*. Pearson.

VERBEEK, Marno; NIJMAN, Theo. (1992), "Testing for selectivity bias in panel data models", *International Economic Review*, v. 33, n. 3: 681-703.

WOOLDRIDGE, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MIT Press.

WOOLDRIDGE, J. M. (2013), *Introductory econometrics: a modern approach*. Australia, South-Western Cengage Learning.

WORRALL, John L. (2008), "An introduction to pooling cross-sectional and time series data", *in* S. Menard (Ed.) *Handbook of longitudinal research: Design, measurement, and analysis*, Elsevier: 233-248.

Perguntas e Respostas

1. **Devo usar efeitos fixos ou aleatórios?**

Em geral, efeitos aleatórios podem introduzir viés, isto é, aumentar a chance de subestimar ou superestimar sistematicamente o parâmetro de interesse. Porém, esse método tende a reduzir a variância dos coeficientes, o que significa estimativas mais eficientes (Kennedy, 2003). Assim, modelos com efeitos aleatórios tendem a produzir estimativas mais consistentes quando a amostra é pequena e quando a correlação entre a variável independente e os efeitos das observações é relativamente baixa (Clark e Linzer, 2015). Por sua vez, efeitos fixos tendem a produzir estimativas não-viesadas, principalmente em amostras grandes (Mummolo e Peterson, 2018). Uma desvantagem dos efeitos fixos é a redução do grau de liberdade do modelo por conta da inclusão de *dummies*. Se, por outro lado, a amostra é suficientemente grande, isso não será um problema na estimação dos coeficientes. Para mais informações: Wooldridge (2002) e Clark e Linzer (2015).

2. Qual é a diferença entre um painel equilibrado e um painel desequilibrado?

Stock e Watson (2004) definem painel equilibrado como um conjunto de dados em que não existem casos ausentes (*missing cases*), ou seja, temos informação para todas as unidades observadas (N) em todos os períodos do tempo (T). O desequilíbrio aparece quando, por qualquer motivo, ocorrem descontinuidades na matriz de dados, seja na falta de observações, seja pela ausência de unidades de tempo. Wooldridge (2002) alerta que, de forma geral, o tratamento estatístico será similar para ambos os tipos de painel. Em todos os casos, o pesquisador deve informar a razão do desequilíbrio e descrever detalhadamente os critérios de seleção da amostra. Para uma introdução sobre o assunto, ver Mayer (2010)[1]. Para uma discussão detalhada das especificidades do painel desequilibrado, Wooldridge (2002) e Baltagi (2005). Para um tratamento mais avançado sobre viés de seleção em dados de painel, ver Verbeek e Nijman (1992).

3. Painéis curtos versus painéis longos?

A amplitude do painel depende da razão entre entidades (N) e unidades de tempo (T). No painel curto, temos muitas entidades observadas durante um período relativamente curto de tempo. Imagine um conjunto de dados com informações detalhadas para todos os municípios brasileiros em 2000 e 2010. Sem dúvida, é muita informação, mas como a quantidade de entidades (mais de 5 mil municípios) é muito superior ao número de períodos de tempo (10 anos), estamos diante de um painel curto em que há dominância da dimensão transversal. O painel longo ocorre quando a oferta de dados transversais é significativamente menor do que a amplitude temporal das informações. Por exemplo, imagine uma base de dados das cinco regiões do país (N, NE, CO, SE e S) reunidas anualmente entre 1950 e 2020. Essa composição em que T>N nos aproxima de uma tendência *time series*, exigindo técnicas específicas. Para mais informações, ver Cameron e Trivedi (2005)[2] e Stock e Watson (2004).

1 Ver: <https://homepage.univie.ac.at/robert.kunst/pan2010_pres_mayer.pdf>.

2 Ver: <http://public.econ.duke.edu/~vjh3/e262p_07S/readings/Cameron_&_Trivedi_Microeconometrics_Chapters_2.pdf>.

4. Como evitar viés de variável omitida com dados de painel?

Como em outros desenhos de pesquisa, é preciso estar a par da literatura específica da área para conhecer a especificação teórica dos modelos. Muitas vezes, podemos achar uma variável importante em alguma base de dados já existente. Tipicamente, FE e FD oferecem boas salvaguardas contra variáveis omitidas quando os *unobservables* importantes são fixos às unidades. Mesmo assim, seu uso não substitui conhecimento abalizado da literatura. Lembrando que a inclusão de uma variável irrelevante no modelo explicativo apenas prejudicará a eficiência dos coeficientes. O erro padrão será maior do que deveria e isso reduz a confiabilidade do testes de significância, mas $\hat{\beta}$ ainda fornecerá uma estimativa consistente de β, ou seja, a inclusão de uma variável irrelevante não causa viés. Por outro lado, a exclusão de uma variável teoricamente importante tende a produzir estimativas viesadas, principalmente quando a variável omitida está correlacionada com as demais variáveis incluídas no modelo.

5. O que é e para que serve o teste de Hausman?

O objetivo do teste de Hausman é avaliar se existe correlação entre os erros do modelo e as variáveis independentes. No contexto de dados de painel, o teste é empregado para orientar a decisão entre efeitos fixos e efeitos aleatórios. A hipótese nula sustenta que os efeitos aleatórios são mais adequados. Operacionalmente, devemos estimar os dois modelos (com efeitos fixos e aleatórios) e depois comparar as estimativas. Um resultado significativo (p-valor<0,05) indica que a hipótese nula deve ser rejeitada, ou seja, devemos optar por efeitos fixos. Por outro lado, um resultado não significativo (p-valor>0,05) implica na não-rejeição da hipótese nula, o que significa a adoção do modelo com efeitos aleatórios. O teste de Hausman geralmente vem acompanhado de uma distribuição de chi-quadrado com teste de significância, além das estimativas dos coeficientes. Para mais, ver Hausman (1978).

6. O que é a autocorrelação serial e por que eu devo me preocupar com ela?

Embora a autocorrelação possa ocorrer no espaço, a autocorrelação que nos interessa aqui é aquela que ocorre no tempo, também chamada de correlação serial. O dicionário de Estatística de Cambridge define correlação serial como a dependência de mensurações da mesma entidade em estudos longitudinais. Em geral, quanto maior a proximidade temporal das mensurações, maior a magnitude da correlação. Por outro lado, quanto mais distante as mensurações no tempo, em média, menor o grau de correlação entre os valores da série. Essa estrutura de associação dos valores no tempo viola o pressuposto de independência dos erros no modelo de regressão. O principal problema da autocorrelação serial é a produção de testes de significância inconsistentes, ou seja, não poderemos confiar no p-valor e nos intervalos de confiança. É por esse motivo que a análise de dados longitudinais sempre deve vir acompanhada de estimativas do nível de autocorrelação da série.

7. Como aprender mais sobre análise de dados de painel?

Em Econometria, sugerimos Kennedy (2003), Wooldridge (2002, 2013), Baltagi (2005) e Gujarati (2004). Hsiao (2003) apresenta uma das mais completas introdução ao tema. O artigo "*What to do and (not to do with) with Time-Series Cross-Section Data*" de Beck e Katz (1995) é o trabalho mais amplamente citado em Ciência Política (7.414

citações no *google scholar* em junho de 2021). Fortin-Rittberger (2013) também apresenta um trabalho introdutório. Finkel (1995) enfoca as possibilidades causais a partir de dados longitudinais. Em dúvida sobre efeitos fixos ou aleatórios? Ver Bell e Jones (2015). Mummolo e Peterson (2018) reproduzem vários estudos que empregam dados de painel e demonstram como melhorar a interpretação dos coeficientes estimados utilizados efeitos fixos. Veja a programação do curso de verão da IPSA/USP: <http://summerschool.fflch.usp.br/tracks-and-courses/>.

8. **TSCS x dados de painel?**

Beck (2008) diferencia dados de painel dos dados TSCS. Para ele, os métodos que funcionam bem em painéis não necessariamente funcionam em dados TSCS e vice-versa. Na pesquisa empírica em CPRI, a combinação mais frequente de casos e tempo se materializa em bancos de dados que compilam informações sobre países (N) ao longo de anos (T) (Beck, 2001). Assim, como os dados do tipo TSCS consistem em exatamente os mesmos casos ao longo do tempo, espera-se maior correlação temporal do que para dados de painel. TSCS consiste na coleta de informações sobre unidades observacionais (indivíduos, firmas, unidades federativas, países) dispersas ao longo do tempo (dia, mês, ano). Quando maior for o T em relação ao N, maiores serão as preocupações sobre a estrutura serial dos dados. Por outro lado, quando existem muitos casos (N) para poucos períodos de tempo (T), menos graves serão problemas de autocorrelação.

9. Como recuperar os interceptos de cada unidade quando uso FE ou RE no plm?

Use os comandos fixef() e ranef(). Os valores correspondem ao desvio em comparação ao intercepto geral.

10. Qual a diferença entre utilizar o *plm* ou os pacotes *lme4* e *nlme*?

Dados em painel não são a única forma de lidar com mensurações repetidas. Outra abordagem consiste no uso de modelos lineares mistos, também conhecidos como modelos hierárquicos ou modelos multiníveis, que são executados por pacotes como *lme4* e *nlme.* Enquanto que a perspectiva de dados em painel é principalmente econométrica, em outros campos como saúde, biologia e educação, a abordagem multinível goza de uma tradição mais longa por terem dados tipicamente hierarquizados (e.g.: alunos em salas, salas em escolas). Dados em painel podem ser encarados como uma estrutura aninhada de dois níveis: há indivíduos *i* (nível superior) observados em ocasiões sucessivas *t* (nível inferior). As especificações FE e RE, por admitirem interceptos variáveis por indivíduo, possuem uma lógica próxima aos modelos hierárquicos em que há interceptos aleatórios (*random intercepts*). Além dos interceptos, é possível trabalhar com dados em painel prevendo coeficientes variáveis segundo indivíduo ou período ($\beta_{□□}$) (exemplos em Wooldridge 2013, p. 451-453; Croissant e Millo 2018, seção 8.2), embora essa possibilidade seja mais amplamente explorada nos modelos hierárquicos (*random intercepts* + *random slopes*). Quando temos configurações de dados em que o aninhamento não é no tempo, ou é mais complexo que apenas dois níveis (ex.: municípios em estados, estados em regiões), o uso de modelos hierárquicos é recomendado. Do ponto de vista operacional, o ecossistema de pacotes do R é mais bem desenvolvido em torno do *lme4/nlme* do que do *plm*, de modo que há maior facilidade em integrar os produtos das análises do *lme4/nlme* em outros pacotes gráficos, de testes e de apresentação de resultados. Para mais, ver Skrondal e Rabe-Hesketh (2008) e Croissant e Millo (2008, pp. 33-39).

Anexos

Tabela 1 – Total de artigos coletados por periódico da área de CPRI (2000-2018)

Periódico CPRI	Qualis (2016)	Total de artigos	% Total
Estudos Avançados	B1	896	12
Revista de Administração Pública	A1	726	9
Sociologias	B1	573	7
Revista Brasileira de Ciências Sociais	A2	559	7
Tempo Social	B1	533	7
Revista de Sociologia e Política	A1	518	7
Caderno CRH	A2	503	6
Sociedade e Estado	B1	483	6
Dados - Revista de Ciências Sociais	A1	469	6
Lua Nova: Revista de Cultura e Política	A2	450	6
Novos estudos CEBRAP	A2	420	5
Revista Brasileira de Política Internacional	A1	403	5
Opinião Pública	A1	340	4
Contexto Internacional	A2	323	4
Revista Brasileira de Ciência Política	B1	235	3
Civitas - Revista de Ciências Sociais	B2	183	2
Brazilian Political Science Review	A2	150	2
Total: 17 periódicos		7.764	100

Fonte: elaborado pelos autores

An introduction to panel data regression[1]

Rafael Mesquita - Federal University of Pernambuco
Antônio Alves Tôrres Fernandes - Federal University of Pernambuco
Rafael Mesquita - Federal University of Pernambuco

Resumo

Apesar da crescente oferta de dados em formato de painel, ainda são raros os estudos no Brasil que combinam as dimensões transversal e longitudinal na mesma análise. Por exemplo, em uma amostra de mais de 7 mil artigos publicados entre 2000 e 2018 em periódicos de CPRI, apenas 45 casos citavam técnicas específicas para lidar com observações de unidades espaciais (países, estados, pessoas) repetidas em intervalos regulares do tempo (anos, meses, dias). Diante dos benefícios inferenciais que tal abordagem pode proporcionar e da escassez de pesquisas sobre o tema, este artigo apresenta uma introdução à regressão de painel. Metodologicamente, sintetizamos as principais recomendações da literatura e mostramos a implementação no *R Statistical* com o pacote *plm*, indo desde a seleção de modelos até o tratamento dos dados e apresentação de resultados. Para aumentar o potencial pedagógico do trabalho, disponibilizamos os materiais de replicação, incluindo dados originais e scripts computacionais. Com este artigo esperamos difundir a utilização de análises longitudinais na pesquisa empírica em CPRI no Brasil

Palavras-chave: Dados em painel; TSCS; CPRI no Brasil; Regressão linear; Metodologia política.

1 This article received comments from the Research Methods group from the Political Science Department of the Federal University of Pernambuco (DCP – UFPE). A preliminary version was presented under the title "*Regressão com dados de painel: balanço da utilização na CPRI brasileira e estratégias para maior difusão do método*", in section "*Ensino e Pesquisa em Ciência Política e Relações Internacionais*" of the 12th Conference by the Brazilian Association of Political Science, João Pessoa, Paraíba (2020). We are grateful for the participants' comments and suggestions, in particular to professors Elia Alves and Jakson Aquino. Content partially from Mesquita (2018). Replication materials available at: <https://osf.io/5yx7g/?view_only=ac1691cced8549238d6d6e0a9d2b7f7b>. We are grateful to Ivan Fernandes (UFABC) and Gustavo Fernandes (FGV) for allowing the use of their database.

1. Introdução

O

Abstract

In spite of the growing offer of data organized in panel format, there are still few studies in Brazil combining cross-sectional and longitudinal dimensions. For instance, out of a sample of over 7 thousand articles published between 2000 and 2018 in Brazilian Politicand International Relations journals, only 45 articles cited specific techniques to handle observations comprising spatial units (countries, states, persons) repeated in regular time intervals (years, months, days). Given the inferential gains this approach can afford and the scarcity of research on the topic, this article presents an introduction to panel regression. Methodologically, we summarize the main recommendations of the literature and show how to implement them via *R Statistical* and *plm* package, going from model selection to data treatment and presentation of results. To improve the pedagogical potential of this work, we share replication materials, including the raw data and computational scripts. With this article, we aim to broaden the use of longitudinal analysis in empirical research in Political Science and International Relations in Brazil.

Keywords: Panel Data; TSCS; Political Science and International Relations in Brazil; Linear Regression; Political Methodology.

1. Introduction

A great variety of data in Politics and International Relations (PIR) are measured in panel form, that is, observations of spatial units measured at regular time intervals (Beck, 2001). The notation *Times-Series Cross-Section* (TSCS) is recurrent in the literature and reveals that a database combines spatial (countries, states, municipalities, individuals, etc.) and temporal dimensions (years, months, days, etc.) (Beck, 2008). The expression *time-series* means that the cases are laid out over time, and *cross-section* indicates that a database is formed by multiple transversal observations (Gujarati, 2004).

For certain research questions, that combination of temporal and spatial information can be very important to achieve more reliable answers than what could be produced by using a single approach. For example, if we wanted to verify if economic growth affects election results in Brazilian municipalities (Fernandes and Fernandes 2017), we could initially just conduct a cross-sectional study comparing several municipalities in the last election. That research design, however, would be limited, since we know that in the last few years economic growth in almost the whole national territory was low, meaning we will only know the effects of negative growth rather than positive. In addition, particular circumstances of the recent period (e.g., corruption accusations) with electoral impact could not be controlled for if they are nation-wide. However, overcoming these difficulties is possible if we gather data not only from the last election, but from repeated elections for all municipalities of interest.

In addition, panel data are valuable because they allow for important gains when it comes to measurement, controlling confounders, improvement of causal inferences, and sample size (Baltagi, 2005; Fortin-Rittberger, 2013). Obviously, these advantages come with costs. For example, the structure of the error term requires specific procedures for the results to be consistent and valid (Worrall, 2008). Despite being suitable to the phenomena studied by PIR, the method is still under-utilized: in a sample of over seven thousand published papers between 2000 and 2018 in PIR journals, only 45 cited specific panel regression techniques.

Thus, faced with the analytical benefits of this approach and its limited use in PIR fields, this paper shows didactic solutions on what to do and what not do with panel data (Beck and Katz, 1995). Our target audiences are undergraduate and graduate students in their initial stages of training. This pedagogical contribution becomes more important given the absence of didactic material especially focused on PIR.[2] Most manuals are either dedicated towards other disciplines, such as economics (Wooldridge, 2002; Kennedy, 2003; Stock and Watson, 2004), or have a prohibitive degree of complexity.[3] With this in mind, we have also included a question and answer section with frequently asked questions on the topic, in addition to the computational scripts that enable readers to adapt them to their respective research interests.

The article is organized as follows: the next section presents the results of a bibliometric analysis on the use of panel regression in PIR; section 3 gives an introduction the logic of longitudinal data analysis; section 4 replicates the data from Fernandes and Fernandes (2017), to demonstrate a computational implementation of a panel regression step-by-step; and the final section summarizes our recommendations on how to disseminate longitudinal analysis in PIR empirical research in Brazil.

2 For works in Portuguese with the same didactic goal, see Fávero (2013) and Marques (2000). For a tutorial in English with computational emphasis on the *plm* package, see Henningsen and Henningsen (2019).

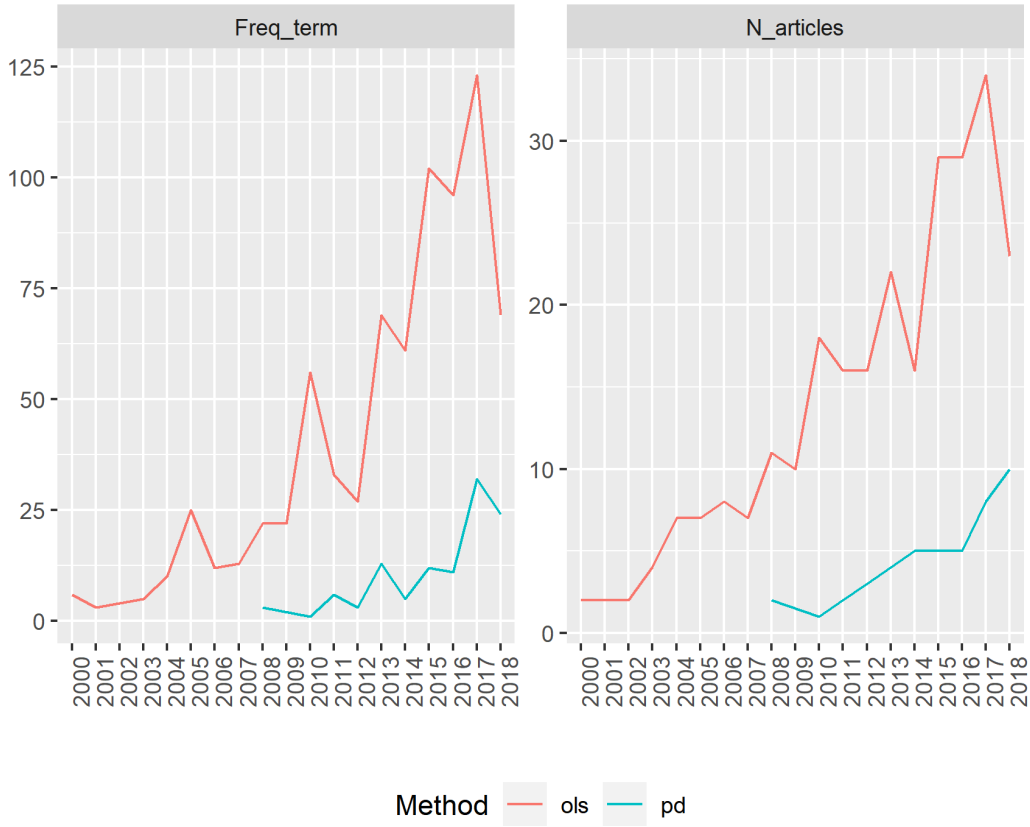
3 See, for example, the degree of difficulty in the final exam for the panel data econometrics course by professor William Greene <http://people.stern.nyu.edu/wgreene/Econometrics/PanelDataEconometricsFinalExam.pdf>

2. An assessment of the use of panel data in Brazilian PIR

How popular are panel data in Brazil PIR? To answer that question, we examined 7,764 articles published between 2000 and 2018 in 17 journals hosted at the Scientific Electronic Library Online Platform (Scielo)[4] (see Table 1 in the Appendix). Similarly to Medeiros et al. (2016), we used automated content analysis to detect keywords in the articles' full text. Using the *rscielo* (Meireles et al., 2019) and *quanteda* (Benoit et al., 2018) packages, we counted the frequency of expressions related to panel data methods. Out of the whole sample, only 45, that is, 0.58%, mentioned expressions related to panel data.[5] Although the interval begins in 2000, only in 2008 did papers with these terms come up, which means the use of this technique in PIR is very recent.

We compared the frequency of panel data with the number of citations for Ordinary Least Squares (OLS), given that this is one of the most disseminated methods in PIR (Krueger and Lewis-Beck, 2008).[6] For each article mentioning panel data (110 citations in 45 articles), there were another five using OLS (758 citations in 263 articles). Figure 1 illustrates that comparison.

Figure 1 – Prevalence of OLS vs. panel data



Source: the authors. The graph to the left counts the occurrence of keywords related to OLS and panel data and the one to the right the total of articles. The blue lines show the total for panel data (PD) and red for OLS.

4 The *rscielo* package automates information gathering from Scielo. As of the writing of this paper, the package identifies a total of 374 journals in all fields of knowledge. Of these, we selected 65 that had been evaluated by Qualis Capes in the PIR field in 2016. Since journals can be evaluated in more than one field, some of them, despite having been evaluated in the field of PIR, had a different concentration area, such as economics, education, or public health. Upon inspecting the list of journals and websites for each, we kept 17 whose focus was unmistakeably PIR. Of this subset, we collected the text of 8,829 papers in total. After removing duplicates and false positives, we arrived at 8,789. Further filtering to publications to only between 2000 and 2018, we reached the final number of 7,764.

5 Terms used for the panel data dictionary: "panel data", "*dados em painel*", "*painel de dados*", "tscs", "time series cross section".

6 Terms used for the OLS dictionary: "ols", "mmq", "mqo", "ordinary least squares", "*mínimos quadrados*", "linear regression", "*regressão linear*", "regression analysis", "*análise de regressão*".

Despite the growth trend, the popularity of panel data methods is today equivalent to OLS' a decade ago. Furthermore, the use of panel data is highly concentrated: 42% of articles were published in the Brazilian Journal of Public Administration (*Revista de Administração Pública*). Of the 17 journals, six did not mention the method at all. These results indicate low dissemination of panel data in national publications in PIR.

Over the remainder of the article, we introduce the essential concepts for panel regression using a bibliographical review of the key manuals and demonstrate a computational step-by-step by replicating Fernandes and Fernandes (2017).

3. Understanding longitudinal data

This section presents the foundations of longitudinal data analysis. First, we describe the different ways to tabulate information over time. We also discuss the main types of panel (short x long; balanced x unbalanced) (section 3.1). Next, we explain the notation of the panel regression model and the four main modalities of estimation (3.2). Lastly, we show the advantages and disadvantages of each approach (3.3 and 3.4).

Longitudinal data can be understood by comparison with cross-sectional data (Menard, 2002). In a cross-sectional design, the information on cases and variables are collected at a specific point in time. For instance, the homicide rate (variable) for all states in the country (cases) in 2013 (time). On the other hand, longitudinal data are collected for at least two different periods of time, which enables the observation of the change in value for each case, for each variable of interest.[7] For example, the homicide rate (variable), for all states in the country (cases), between 2000 and 2010 (time). Figure 2 shows the difference between cross-sectional data, temporal series, and their combination in panel form. As an example, we used the homicide rate per 100 thousand inhabitants in three Brazilian states.

7 Menard (2002) points out four longitudinal research designs: (1) total population designs; (2) repeated cross-sectional designs; (3) revolving panel designs; and (4) longitudinal panel designs. Wooldridge (2013, p. 448) also differentiates between "independently pooled cross section" and "panel data" or "longitudinal data"; the former uses random samples of different individuals in successive points in time (e.g., census sampling), while, in the latter, the subjects remain the same during the time frame.

Figure 2 – Data structure

Cross-Section

UF	Time	Homicide Rate	Overall Mean	Between Case Deviat.
i	t	X_i	$\bar{X}$	$X_i - \bar{X}$
Acre	2011	22,0	24,7	-2,7
Bahia	2011	39,4	24,7	14,7
Santa Catarina	2011	12,8	24,7	-11,9

Time-Series

UF	Time	Homicide Rate	Individual Mean	Intra-case Deviat.
i	t	X_t	$\bar{X}$	$X_t - \bar{X}$
Bahia	2011	39,4	40,2	-0,8
Bahia	2012	43,4	40,2	3,2
Bahia	2013	37,8	40,2	-2,4

Time-Series Cross-Section

UF	Time	Homicide Rate	Overall Mean	Individual Mean	Overall Deviat.	Intra-case Deviat.	Between Case Deviat.
i	t	X_{it}	$\bar{X}$	$\bar{X}_i$	$X_{it} - \bar{X}$	$X_{it} - \bar{X}_i$	$\bar{X}_i - \bar{X}$
Acre	2011	22	26,4	26,5	-4,4	-4,5	0,1
Acre	2012	27,4	26,4	26,5	1,0	0,9	0,1
Acre	2013	30,1	26,4	26,5	3,7	3,6	0,1
Bahia	2011	39,4	26,4	40,2	13,0	-0,8	13,8
Bahia	2012	43,4	26,4	40,2	17,0	3,2	13,8
Bahia	2013	37,8	26,4	40,2	11,4	-2,4	13,8
Santa Catarina	2011	12,8	26,4	12,5	-13,6	0,3	-13,9
Santa Catarina	2012	12,9	26,4	12,5	-13,5	0,4	-13,9
Santa Catarina	2013	11,9	26,4	12,5	-14,5	-0,6	-13,9

Source: the authors, based on Cerqueira (2017).

In our example, the *UF* variable represents the cross-sectional unit of analysis (Brazilian state). The *Time* variable indicates the longitudinal dimension of information, here, the year. X_{it} represents the value of the variable of interest of case i in time t. For example, Acre had a homicide rate of 22 in 2011. The remaining columns show aggregate measures derived from X_{it}. Although these do not usually appear in databases, we reproduced them for teaching purposes.

The overall mean (sometimes called grand mean) is the total sum of cases (237.7) divided by the size of the sample (9), that is, 26.4. The overall deviation is the difference between each observation from the overall mean. For instance, since in 2011 Bahia had a rate of 39.4, it was 13 points above the overall mean (39.4 – 26.4). the deviation between the cases is the difference between the mean of each case and the overall mean. For example, Santa Catarina has a distance of 12.5 – 26.4 = -13.9 points from the overall mean. Lastly, the within-case deviation represents the difference between each observation and the individual mean of that state, meaning that, in 2013, Santa Catarina was 11.9 – 12.5 = -0.6 points below its historical average.

This configuration where each line of the dataset represents a unique combination of case in time is also known as long format and it is the standard arrangement for most statistical software. In several scenarios, there may be variation in the availability of cross-sectional or longitudinal data. Since panel data are the product of the combination of both, this means that different types of panels may be formed, depending on the number of cases or temporal units available, as we explain next.

3.1 Types of panel: short x long, balanced x unbalanced

Panels have two main configurations. There is a short panel, "stacked" or "cross-section dominant", where the number of cases is higher than the amount of time periods, that is, N>T. On the other hand, we say the panel is

long ("temporally dominant") where the number of time periods is higher than the number of cases (T>N). Some authors have standardized referring only to the former as panel data and the latter as Time-Series Cross-Section (TSCS) (Beck, 2001; Fortin-Rittberger, 2013).

For each panel configuration (short or long) different expectations are assumed regarding the distribution of the data. Beck (2001) argues that, for panel data (N>T), the selected units are seen as samples of a population, observed over a short time interval. Thus, inferences are not specific to the units and may be generalizable. For TSCS (T>N), units are seen as fixed and observed for a long period of time. Therefore, inferences are about the units and less generalizable. The traditional examples of the former are census surveys and, for the latter, macroeconomic comparisons between countries.[8]

Another important feature is the availability of information. We have a balanced panel when there is information for all cases in all time periods. On the other hand, we say the panel is unbalanced when some of these cases are absent for certain time periods. An additional complication, which can affect the consistency of estimations, is the nature of the panel's imbalance. The absence of data may be random, which tends to harm the efficiency of the coefficients (standard error, p-value, confidence intervals). However, the imbalance may be related to some variable of interest to the study. For example, it may be easier to find detailed information for richer countries than poorer nations. When the absence of information is systematic for any reason, the estimations tend to be biased. According to Batalgi (2005), unbalanced panels tend to be more frequent than balanced ones, which reinforces the need to describe with precision the process of data collection, treatment, and analysis.

8 Similarly, the asymptotic properties of TSCS presuppose that N is fixed and T can grow to infinity. For panel data, T is fixed, and N can grow to infinity (Fortin-Rittberger, 2013). From this it is deduced that the larger the N, proportionally speaking, the better the data fits to panel data models and, inversely, the larger the T, the better the fit to TSCS models.

We now explain how panel data may be used in regression analysis.

3.2 Notation and types of panel regression

For the purposes of this article, it is important to show the classic notation of the linear regression model, using only cross-sectional data:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad (1)$$

In which Y represents the dependent variable, α indicates the intercept, β represents the variation observed in Y when the independent variable, X, increases in one unit, and ε indicates the stochastic nature of the model. The i in subscript indicates that the observations are indexed by case. Continuing with the example in Figure 2 (homicide rate by state over time), this equation can represent the effect of unemployment (X) on the homicide rate (Y). A panel regression is an extension of the previous model, in which:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \qquad (2)$$

The i and t in subscript inform that the observations are indexed, respectively, by case and time. Our model becomes more powerful, because we accumulate information on the relationship between unemployment and homicide for several states and years. How to move ahead, then, with the regression? We describe below the four traditional approaches to estimate panel models.

Pooled OLS

By gathering all observations in the same linear regression ("pooling"), we have a Pooled OLS (POLS). This approach assumes that all units can be covered by the same regression line. Even if the units of i show differ-

ences between each other, it is considered that the set of independent variables already carries all important information. That is, the main factors that differentiate observations are already included in the model, making it unnecessary to control for other unobserved factors. In our example, it would be as if the homicide rate, in whatever state, was mainly determined by unemployment, and there were no other peculiarities in the states affecting the outcome: if Bahia had the unemployment level of Santa Catarina, it would have the same homicide rate as that state. Since individual heterogeneities are not modelled, we have Equation (2), where there is only one intercept (α) for the whole population and all else that is not observable is allocated as noise in the error term (ε).

However, if there are unobserved factors in ε which are not random but systematic, POLS is no longer appropriate. If we have unobserved elements that are correlated with X, we have a traditional case of omitted variable bias, and the parameter estimate β ($\hat{\beta}$) will be inconsistent.[9] Since observations are repeated over time, the independence of observations is in jeopardy. Suppose that for the three states in the example there is some factor we cannot observe that nonetheless has an impact on the homicide rate, for instance, the influence of organized crime. If the influence of organized crime is correlated with unemployment there will be a bias; if it is specific in each state (e.g., the criminal influence in Acre is the same over time), the errors will not be independent (ε_{it} will be correlated with ε_{it-1}). Because of this, the first motivation to not use the POLS model is the need to work with the heterogeneity hidden in the error term in a more sophisticated way.

The unobserved factors are divided into elements that remain the same for each unit and elements that vary according to unit and time. Therefore, the "composite error" (ε_{it}) of Equation (2) can be separated in two parts: one fixed for each unit (μ_i) and another one variable (ν_{it}).

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it} \qquad (3)$$

The errors persistent to the units (μ_i) are also known as unobserved heterogeneity, unit effect, or fixed effect of the error term, while ν_{it} is usually called variable or idiosyncratic effect. Unlike the other three approaches presented here, POLS does not have strategies to deal with μ_i, because it does not disaggregate ε_{it}. All remaining specifications break down the error factor into fixed and variable.

First Differences

In the First Differences (FD) approach, all variables are differentiated, that is, subtracted from their value in the previous time period. The subtraction is especially efficient in handling dynamic issues with the data, such as serial correlations and non-stationarity, given that, when subtracting an observation at *t* from the observation at *t*-1, everything that has not varied from one moment to the next (including the unobserved heterogeneity) will be erased.

Because it subtracts observations by pairs, FD has costs: it reduces the data's variability and thus increases the standard errors, being, therefore, more suited for large datasets. Furthermore, the differentiation removes all factors constant over time or that vary monotonically (i.e., always increasing at the same rate every year), which hinders the use of dummy variables for countries, geographical distances, and constant growth variables (Wooldridge, 2013, p. 462–463). The use of FD is one of the drastic solutions recommended by Beck (2008) when faced with persistent temporal dependences in the model.

9 If the phenomenon under analysis is better described by a set of variables and one of them and left out of the equation, the effect of the variables preserved in the model will be over- or underestimated. That systematic upward or downward distance in the parameter characterizes a bias.

Fixed Effects

Fixed Effects (FE), as FD, seek to remove the fixed component in the error (μ_i). However, rather than differentiation, they use a "within transformation". The operation is called "within" because it only retains the within-case deviation (see Figure 2). That removal of the value of each observation's mean ("time demeaning" or "group-mean centering") will result in the elimination of α and μ_i, whose values for each unit do not change with the passage of time.[10] Therefore, like FD, the FE estimator omits factors constant in time.

Because it handles the heterogeneity problem and is very robust, the FE specification is seen by Croissant and Millo (2018, p. 5) as a benchmark for panel data. That is the case especially for disciplines that deal with observational data, which was also found in our bibliometric research: fixed effects were the most used method in PIR (26% of articles).

One of the benefits of FE in comparison to POLS is that, while the latter postulates only one intercept for the whole sample, the former takes into account that each unit can have its own intercept. In comparison to FE, POLS will have biased estimators, given that it does not deal with μ_i. Under FE, the model is less biased, although also less efficient.

Computationally, similar results occur when using an FE model and a POLS model with dummies for all cases. In this situation, POLS is usually referred to as Least Square Dummy Variable (LSDV). Figure 3 illustrates the differentiation and mean subtraction procedures.

Figure 3 – Illustration of the differentiation and within transformation procedures

First Differences

ID	Time	Homicide Rate	Unemploym. Rate
i	t	Y_{it}	X_{it}
Acre	2011	22	5,4
Acre	2012	27,4	8
Acre	2013	30,1	9,6
Acre	2014	29,4	9,8
Bahia	2011	39,4	10,5
Bahia	2012	43,4	10,2
Bahia	2013	37,8	9,9
Bahia	2014	40	10
Santa Catarina	2011	12,8	3,6
Santa Catarina	2012	12,9	3,1
Santa Catarina	2013	11,9	3,4
Santa Catarina	2014	13,5	3,1

$$Y_{it-1} = \alpha + \beta X_{it-1} + \mu_i + \nu_{it-1}$$
$$- \left[Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it} \right]$$
$$\Delta Y_{it} = \beta \Delta X_{it} + \Delta \nu_{it}$$

ID	Time	Homicide Rate	Unemploym. Rate
i	t	ΔY_{it}	ΔX_{it}
Acre	d1	5,4	2,6
Acre	d2	-0,7	0,2
Bahia	d1	4	-0,3
Bahia	d2	2,2	0,1
Santa Catarina	d1	0,1	-0,5
Santa Catarina	d2	1,6	-0,3

Fixed Effects

ID	Time	Homicide Rate	Unemploym. Rate	Individual Mean (homic.)	Individual Mean (unempl.)
i	t	Y_{it}	X_{it}	$\bar{Y}_i$	$\bar{X}_i$
Acre	2011	22	5,4	27,2	8,2
Acre	2012	27,4	8	27,2	8,2
Acre	2013	30,1	9,6	27,2	8,2
Acre	2014	29,4	9,8	27,2	8,2
Bahia	2011	39,4	10,5	40,2	10,2
Bahia	2012	43,4	10,2	40,2	10,2
Bahia	2013	37,8	9,9	40,2	10,2
Bahia	2014	40	10	40,2	10,2
Santa Catarina	2011	12,8	3,6	12,8	3,3
Santa Catarina	2012	12,9	3,1	12,8	3,3
Santa Catarina	2013	11,9	3,4	12,8	3,3
Santa Catarina	2014	13,5	3,1	12,8	3,3

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + \nu_{it}$$
$$- \left[\bar{Y}_i = \alpha + \beta \bar{X}_i + \mu_i + \bar{\nu}_i \right]$$
$$\ddot{Y}_i = \beta \ddot{X}_i + \ddot{\nu}_i$$

ID	Time	Intra-case Deviat. (homic.)	Intra-case Deviat. (unempl.)
i	t	$Y_{it} - \bar{Y}_i$	$X_{it} - \bar{X}_i$
Acre	2011	-5,2	-2,8
Acre	2012	0,2	-0,2
Acre	2013	2,9	1,4
Acre	2014	2,2	1,6
Bahia	2011	-0,8	0,4
Bahia	2012	3,3	0,0
Bahia	2013	-2,4	-0,3
Bahia	2014	-0,1	-0,2
Santa Catarina	2011	0,0	0,3
Santa Catarina	2012	0,1	-0,2
Santa Catarina	2013	-0,9	0,1
Santa Catarina	2014	0,7	-0,2

Source: the authors, based on data from Cerqueira et al. (2017) and IBGE (2017).

The figure illustrates the effects of differentiation and within transformation only for the observed variables: homicides and unemployment. Even so, the subtraction also encompasses unobserved factors (indicated in italics in the figure's equations), resulting in the suppression of the overall intercept (α) and the individual effect of the error term (μ_i). In our example, it means that the resulting models, when estimating the effect

10 The α and μ_i components are removed for a consistent estimation of β. Nevertheless, note that some statistical packages report values for individual (μ_i) or population (mean of all μ_i) intercepts. See Wooldridge (2002, section 10.5.3) for details.

of unemployment on the homicide rate, eliminate the fixed influence of organized crime in each state. Not only this, but every unobserved factor that is invariant will be removed and the estimation will only be based on the elements that vary.

Random Effects

As seen previously, a deficiency in FD and FE is the impossibility of testing variables that do not change with time. Even the ones that vary slowly tend to be underestimated (Clark and Linzer 2015). Nevertheless, there are many variables of interest for PIR that are static, such as geography, gender, nationality, among others. The Random Effects (RE) specification bypasses that issue, enabling the inclusion of variables that do not change with the passage of time.

As with FE, it is assumed that the composite error has fixed and idiosyncratic parts, that is, $\varepsilon_{it} = \mu_i + \nu_{it}$. However, instead of eliminating μ_i completely, it is assumed that it has no correlation with X (Bell and Jones 2015, p. 142). If the unit fixed effect is not correlated with the independent variables, it is no longer a source of bias and, consequently, no longer needs to be eliminated (e.g., if we believe that organized crime is independent from unemployment, there is no reason to remove it).

The second distinctive assumption of RE is about unit effects. *Fixed* effects assume that the effect of each unit can be calculated as a unique value. Therefore, FE calculates an intercept for each individual in the form of a parameter. In contrast, RE considers μ_i and ν_{it} as *random* realizations in a distribution. In the example involving unobserved organized crime in the states, FE will estimate that individual impact case-by-case, without assuming that there is a relationship between the level of organized crime in Bahia and Acre. RE gives more structure to the set, because it assumes that the effects of crime tend towards a normal distribution, all states considered (Bell et al., 2019, p. 1061).

Some authors classify RE as a middle ground between POLS and FE (Clark and Linzer 2015, p. 402). POLS does not eliminate μ_i and FE erases it completely. RE partially executes this procedure, since it subtracts a portion from the observations of the within-group mean ("quasi-demeaning"). The proportion of that fraction is given by θ (*theta*, in the *plm* output), which takes on values between 0 and 1. *Theta* can be interpreted as a diagnostic of which component most contributes to the variance of ε: if fixed (μ_i) or idiosyncratic (ν_{it}). When there is little variance in the fixed effects ($\theta \to 0$) the RE estimates approach POLS and, in the opposite scenario ($\theta \to 1$), they approach FE (Wooldridge, 2013, p. 494).

On the other hand, the assumption that there is no relationship between μ_i and X is hard to maintain in most real situations. Another significant difference is that RE estimators will always have some bias and can only be efficient. The choice between FE and RE is, therefore, a trade-off between efficiency (RE) and unbiasedness (FE).

The direct comparison between RE and FE is recurrent in the literature. Disciplines more accustomed with randomized experiments tend to prefer RE (Dieleman and Templin, 2014). Beck (2001, p. 284) argues that observational studies prefer FE to RE since the latter is more recommended when the units are considered to be a sample of a larger population, and the former when the aim is to make inferences restricted to those units. The idea of "interchangeability" captures that notion intuitively: when, for a set of observations, we can replace case A for case B without great loss of information – that is, the *name of the case* does not matter – then we have something close to a random sample (Hausman 1978, p. 1262; King 2001, p. 498). The

RE expectation of a normal distribution of individual effects thus becomes acceptable. In some PIR scenarios, that interchangeability is plausible (e.g., representatives in a legislature), but not in others (e.g., member-states in the UN Security Council).

Lastly, when conducting an empirical analysis, it is always informative to make a side-by-side comparison of the results from the four types of regression: POLS, FD, FE, and RE. The contrast indicates the different results obtained depending on what is done with μ_i: POLS keeps the fixed factors completed inserted in the error component, RE maintains a portion, and FE and FD remove it completely (Wooldridge, 2013, p. 494).

3.3 Benefits of panel data

Box 1 summarizes the main technical benefits and operational issues associated with panel data.

Box 1 – Benefits and issues of panel data

Benefits	Issues
(1) Facilitates the detection of causality; (2) Measures individual variation; (3) Reduces measurement errors; (4) Increases the size of the sample; (5) Controls omitted variables issues.	(1) Serial correlation of the residuals; (2) Non-stationarity; (3) Heterogeneity; (4) Panel heteroskedasticity; (5) Contemporaneous correlation of the errors; (6) Complex dependence structures.

Source: the authors, based on Fortin-Rittberger (2013), Hsiao (2003), and Kennedy (2003).

The first advantage of panel data is to facilitate the detection of causal relationships between X and Y. Pearl (2000) highlights three assumptions to identify a causal relationship: (1) association between the variables (correlation); (2) temporal precedence; and (3) nonspuriousness of the relationship. Unlike cross-sectional data, in which values of X and Y are measured in the same time frame, panel data enable one to observe if the temporal variation of X is correlated with the temporal variation of Y. Thus, in addition to controlling for possible spurious variables, longitudinal data favor the assumption of temporal precedence between the independent and dependent variable.

Furthermore, the longitudinal nature also favors the study of change since it maintains the same unit of analysis over time. Therefore, it is possible to examine the individual variation of different indicators among the cases of interest. For example, the aggregate analysis of the number of homicides in a city is informative. But a database that informs the day of death and neighbourhood brings additional possibilities. It would be possible to observe, for instance, if the majority of occurrences are concentrated in the weekends of a certain neighbourhood. As Fortin-Rittberger (2013) highlights, that possibility to monitor, in space and time, where and when certain phenomena of interest arise brings longitudinal studies, though being observational, closer to the logic of a controlled experiment.

As a third advantage, we have more reliable measures on the change of the variables. According to Blalock (1979), measuring is the main challenge to the development of the social sciences. In particular, in PIR, important concepts, such as democracy, are not directly observable. One of the ways to examine the reliability of a measure is to measure it on several occasions over

time. Consequently, panel data help to overcome possible measuring errors. [11]

In addition to measuring gains, another benefit of panel data is increasing the sample size. One of the recurring problems in research design is indeterminacy, that is, "many variables, small N" (Lijphart, 1971, p. 686). When we have a few observations, it is difficult to argue that the causal relationship proposed is explained by one reason and not any other. Thus, determinacy of a research design relates to the proportion of variables and observations. For this reason, increasing the number of cases is a common recommendation to solve indeterminacy (King, Keohane and Verba, 1994). Furthermore, larger samples solve efficiency issues (smaller standard error), multicollinearity, and make using more robust statistical tests easier (Hsiao, 2003).

Lastly, panel data control the problem of omitted variables (Kennedy, 2003). The omission of an important variable will produce bias in the estimates if it is correlated with other modelled variables. A purely cross-sectional perspective does not offer many solutions to deal with unobserved factors. Frequently, we are simply ignorant on the unobservables that impact Y. However, from the moment that cases are studied over time, tools such as FE, FD, and RE may be used to remove or mitigate these individual heterogeneities (Wooldridge 2002, section 10.1).

Nevertheless, these benefits do not come without a cost, as we show next.

11 Errors can be systematic or random. Systematic error is harmful in descriptive analyses because the estimates produced will be biased. On the other hand, when the systematic error impacts all cases with the same magnitude, it does not affect the consistency of the estimates in regression models. Conversely, random error in the dependent variable tends to reduce efficiency (higher standard error) but does not affect consistency of the estimates. Lastly, when independent variables also present random error, the estimates tend to be inconsistent and, generally, underestimated (King, Keohane and Verba, 1994).

3.4 Challenges with panel data and how to overcome them

For pedagogical purposes, we propose a brief summary of the assumptions of an OLS linear regression as a starting point to address the typical issues of panel regressions.[12] According to Wooldridge (2013, p. 59, 119), the main assumptions for OLS can be summed up into six: (A.1) There is a linear relationship between X and Y; (A.2) Absence of autocorrelation (or random sample); [13] (A.3) Absence of perfect multicollinearity ("full rank"); [14] (A.4) Exogeneity of X (or "zero conditional mean") – the error must not be a function of the independent variables, that is, it is random and none of the independent variables in the model carries information regarding its value;[15] (A.5) Homoskedasticity – there is uniform variance in the regression residuals and these are not correlated to each other – if, for some reason, in a portion of the data, the variance of residuals changes according to the values of the explanatory variables, there is heteroskedasticity;[16] (A.6) Normal distribution of the errors.

If the so-called Gauss-Markov assumptions (A.1-A.5) are respected, it can be said that the parameters estimated are the *Best Linear Unbiased Estimators* (or BLUE) of the population's parameters. "Best" because they have the smallest variance (A.5) since there is homoskedasticity (uniform variance of residuals); "Linear" because of the linear relationship (A.1); and

12 The list of OLS assumptions sometimes varies among authors. For other lists of linear regression assumptions and the Gauss-Markov theorem, see Berry (1993) and Larocca (2012).

13 Random samples and the absence of autocorrelation may seem like distinct expectations, but they are equivalent once we consider that, in both cases, we have $\mathrm{Corr}(\varepsilon i, \varepsilon j = 0)$, for $i \neq j$.

14 The main damage caused by multicollinearity is the increase in the variance of the regression coefficients, which increases the chance of finding non-significant results.

15 The violation of exogeneity equals an omitted variable bias.

16 In other words, the model fits better for certain sections of the data and not others. Heteroskedasticity does not lead to bias or inconsistency of OLS estimators, but because it interferes in the distribution of variances, it hinders the construction of confidence intervals and significance tests (Wooldridge, 2013, pp. 268–269).

"Unbiased" because there is no bias in the parameter, since the regressors are exogenous and do not affect the error (A.4).

However, panel models can easily violate those assumptions. Because they combine temporal series and cross-sectional comparisons, they inherit the issues of both types of data. These problems can be divided into dynamic and spatial. The literature usually recommends attention to dynamic issues first. These especially violate the requirements of absence of autocorrelation (A.2), exogeneity (A.4), and homoskedasticity (A.5). The main ones are:

(1) **Serial correlation of the residuals:** there is correlation between the residuals at moments *t* and *t-1*. This is a persistent characteristic of long panels given that, as Box et al. (2016) caution, an intrinsic aspect of temporal series is that adjacent observations are dependent. The main issue generated by the correlation of errors is an inconsistency of significance tests (p-value and confidence intervals).[17] For this reason, the serial correlation must be detected and solved before moving on to spatial issues. There are different tools for detecting temporal processes in the error structure: the plots of the autocorrelation (ACF) and partial autocorrelation functions (PACF), the Durbin-Watson, Breusch-Godfrey, and Ljung-Box tests, among others. Most of them are directed towards checking first-order autoregressive processes, or AR(1), which are more common.[18]

It is possible to test an AR(1) correlation in the errors of POLS models in a simple and direct way – by running a regression of the residuals at *t* on those at *t-1*, without an intercept, as in $\varepsilon_{it} = \rho\varepsilon_{it-1} + \nu_{it}$. The result for parameter ρ will indicate whether or not the residuals from the past are significantly correlated with the residuals from the future (Wooldridge, 2002, section 7.8.5; 2013, p. 417).

The FE, RE, and FD specifications will require specific tests. To cover the whole arsenal of existing tests goes beyond the scope of this article (for a list, see Croissant and Millo 2008, p. 22-28; 2018, section 4.3). However, we highlight that the selection of tests must be guided by the type of estimation used in the model (POLS, FE, RE, or FD), the N/T proportion, and the number of observations.

Once a serial correlation is found, recommended solutions include adding the dependent variable lagged, inserting robust standard errors, or using FD. Regarding the first alternative: the inclusion of a lagged dependent variable (LDV) is recommended, under the assumption that the process is AR(1) and, afterwards, running a new test to check if the autocorrelation has been eliminated (Beck, 2001). However, this solution has been criticized by Achen (2000, p. 14), who states that the LDV may absorb the whole explanatory power of the remaining regressors and thus produce bias.[19]

(2) **Non-stationarity:** a temporal series is stationary when it is in statistical equilibrium, that is, it has constant mean and variance over time. Although values oscillate, they will return to the mean. It is non-stationary when the mean and variance are not constant, that is, when the series does not tend towards a prior mean after deviations. In other words, non-stationarity may be understood as a persistent effect of prior shocks (Box et al., 2016, p. 7;

17 Technically, autocorrelation affects the efficiency of estimates. In particular, if the autocorrelation is positive, the standard error of the coefficients will be underestimated, which increases the chances of incorrectly rejecting the null hypothesis (type I error).

18 Although it is not common to expect more complex processes than AR(1) for panel data (Fortin-Rittberger, 2013), it is possible that there are higher-order correlations, which will motivate more complex solutions (e.g., the inclusion of variables with multiple lags).

19 When the regressors present some trend, the introduction of an LDV will dominate the regression. The LDV's coefficient will be biased upwards and the remaining regressors, downwards. It is also not recommended when the T is small, because it sacrifices a period of analysis (Achen, 2000).

Fortin-Rittberger, 2013). Deviations of stationarity may come from the type of phenomenon monitored but may also result from the measuring frequency: if information was measured by day or by month, for example, it is possible to find seasonal effects and, for samples with a large T, cyclical behavior. The visual inspection of the time series is one of the simplest ways of identifying trends or seasonality. Different tests in the literature on time series may be employed to verify non-stationarity: ACF, PACF, and unit root tests (e.g., Augmented Dickey-Fuller, ADF). More recently, tests specific to the context of panel data, "Panel Unit Root Tests", have been proposed (see Croissant and Millo 2018, chapter 8). The application of these tests will be hindered if the panel has few time periods. If the series is not stationary, solutions include using FD or, in cases of non-stationarity by trend, including a time index or trend variable (Fortin-Rittberger, 2013). Now we turn to spatial aspects.

(3) **Heterogeneity:** regressions assume that cases are homogenous. All particular characteristics are explained by the independent variables, as previously explained with the concept of "interchangeability" (King, 2001, p. 498). Heterogeneity is present when one or some of the cases observed have distinct characteristics that, having not been modelled, will end up in the error term and thus generate correlations between $\varepsilon_{\square\square}$ and X_{it}. Consequently, heterogeneity violates the assumptions of exogeneity (A.4) and homoskedasticity (A.5). If heterogeneity is present, it will usually be inaccurate to assume that there is only one intercept for the whole population, as in POLS models. If every unit has impactful fixed characteristics, the absence of unit-specific intercepts will lead to a mistaken regression line (Fortin-Rittberger, 2013). On the other hand, violating the homoskedasticity assumption implies producing inconsistent significance tests, which increases the chances of mistaken inferences.

As discussed in section 3.2, abandoning POLS estimates in favor of other specifications is one way of solving heterogeneity. The last three spatial difficulties mainly compromise exogeneity (A.4), homoskedasticity (A.5), and normality of the errors (A.6).

(4) **Panel heteroskedasticity:** the residuals must have constant variance in both directions, that is, between units *i* and time periods *t*. Panel heteroskedasticity is present when residuals present constant variance over time within each unit, but inconstant across units. In other words, each unit will have its own residual variance. It may happen due to the model's poor specification or when one or two cases do not fit well within the specification.

(5) **Contemporaneous correlation of the errors:** a case's error is correlated with the error of others for the same moment in time. For example, if an external shock occurs in a year and affects several units at the same time, the units' residuals for this period will present a different variance from other periods.

(6) **Complex dependence structures:** there are more special cases that, somehow, impose dependence between observations. One of the most explored in observational studies is the spatial or geographical correlation between units. That is, contiguity or proximity between two countries exposes them to more events in common and, because this is not random, may lead to biases.

Generally, tests such as Pesaran's (2004) offer evidence on whether there is dependence between units such as the ones described in items (4) to (6), although there are specialized variations for different types of panel and dependence structure (see Croissant & Millo 2008, p. 28-31; 2018, chapter 10).

Spatial issues in POLS models may require the use of FE, RE, or FD estimators. It is also recommended, for

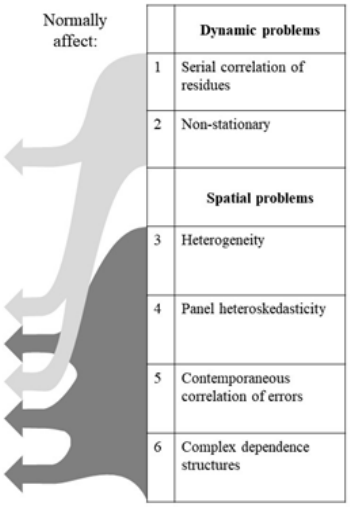
whichever estimator, reviewing the set of variables, in order to consider new factors that may capture the heterogeneity between the units; in addition to using robust standard errors.

In conclusion, it is useful remembering that the likelihood of these issues occurring depends on the panel type. T>N panels tend to have more problems with heteroskedasticity and serial correlation of errors, while N>T panels frequently have issues with non-measured heterogeneity between the cases.

Figure 4 summarizes the list of assumptions and violations discussed in this section.

Figure 4 – Summary of OLS assumptions and dynamic and temporal issues

	Assumption	Formal	Why is it a problem to violate it?
A.1	Linear relation	$Y = \alpha + \beta X + \varepsilon$	Bias
A.2	No autocorrelation, Random sample	$Corr(\varepsilon i, \varepsilon j) = 0$ $i \neq j$	Bias
A.3	No perfect collinearity	$Corr(X1,X2) \neq \pm 1$	Impossible to estimate $\hat{\beta}$; Bias; Inflated SE($\hat{\beta}$), p-value
A.4	Exogeneity	$E(\varepsilon \mid x) = 0$	Bias
A.5	Homoskedasticity	$Var(\varepsilon \mid x) = \sigma^2$	Invalidates SE($\hat{\beta}$), p-value
A.6	Normal distribution of errors	$\varepsilon \sim Normal(0, \sigma^2)$	Invalidates SE($\hat{\beta}$), p-value



Source: the authors.

4. Application

In this section, to illustrate how to conduct a panel analysis, we replicate data by Fernandes and Fernandes (2017). Although it is common to use classic databases from statistical packages for teaching (e.g., "Grunfeld", "EmplUK", etc.), we consider this a valuable opportunity for replication that simultaneously incentivizes learning and a transparency and replicability culture in the social sciences (King, 1995).

The main goal of the article "*A importância do crescimento econômico local na escolha do chefe do Executivo no Brasil*" ("The importance of local economic growth in the choice for president and mayor Brazil") is to analyze the impact of local economic growth on the incumbent's vote share in Brazilian presidential and municipal elections between 2000 and 2010.

To answer the research question, Fernandes and Fernandes (2017) mobilize 17 independent variables to explain the dependent variable: the vote share received by the incumbent. According to the authors: "We estimated the relationship using panel data techniques that enabled us to analyze whether unobservable idiosyncrasies affect the estimation or not. To so do, we used the three traditional models of static panel: pooled OLS, random effects (RE), and fixed effects (FE)" (Fernandes & Fernandes, 2017, p. 660).[20] To replicate the article's data, we used R, RStudio, and the *plm* package (Croissant & Millo, 2008). To make understanding the computational step-by-step easier, we report the scripts throughout the text.[21]

20 In the original: "*Estimamos a relação por meio de técnicas de dados em painel que permitem analisar se as idiossincrasias não observáveis afetam ou não a estimação. Para isso, usamos os três modelos tradicionais de painel estático: painel empilhado, efeitos aleatórios (EA) e efeitos fixos (EF)*" (Fernandes & Fernandes, 2017, p. 660).

21 Note that the example dataset at OSF has its variables named in Portuguese. Translations into English in the scripts reproduced here are merely for clarity.

For the following tests, it is useful to already have in the RStudio environment the four types of specifications that will be compared: POLS, FE, RE, and FD. First, it is necessary to load the database and fit it to execute the models.

```
##### OPENING DATABASE ############

library(haven) # package to read Stata data
library(plm) # package to execute the panel data models

DATA <- read_dta("fernandes_2017.dta") # Database file in the directory

DATA <- pdata.frame(DATA, index = c("codibge", "year")) # here the database is converted to the pdata.frame format
# to execute the models. In 'index', the data's spatial and temporal dimensions are added.
# In this case, 'codibge' represents the spatial dimension (municipal code with the Brazilian Institute of Geography and Statistics) and 'year' the election year.
```

By applying the *pdim()* command, we can see that the database has N= 5,565 municipalities and e T=2-6.[22] In total, there are 28,945 observations. That is, we have a short (N>T) and unbalanced panel. By transforming the database into a *pdata.frame* object, an index containing the spatial and time dimensions is created, enabling the estimation of the models.

```
##### FERNANDES FORMULA 2017 ###################

form <- as.numeric(voteshare) ~ growth+growth_state+growth_country+log_gdp+log_gdp_state+
  log_gdp_country+mayor_base_presid+health_exp+log_pop+leec+
  lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespers

##### POOLED OLS ################################

POLS <- plm(form, data = DATA, model = "pooling")  # pooled model

##### FIXED EFFECTS #############################

mode_fe <- plm(form, data = DATA, model = "within") # The FE model is executed with the model = 'within'

##### RANDOM EFFECTS ########################

mode_re <- plm(form, data = DATA, model = "random") # The RE model is executed with the model = "random".

##### FIRST DIFFERENCES #####################

mode_fd <- plm(form, data = DATA, model = "fd") # The FD model is executed with the model = "fd".
```

22 The period under observation is "electoral year". The observations begin in 2000 and go up to 2010, registering data every two years. This is why the database has T=6 and not T=10.

We executed four models: POLS, FE, RE, and FD. The POLS model is executed through the specification *model = "pooling"* in the *plm* function.[23] Similarly, the FE, RE, and FD models are specified in *model*.

Which of the above models should we use? To answer that, we need to check the assumptions of each one.

4.1 Pooled OLS?

The first and simplest choice is POLS, which applies a single regression line, with the same intercept, for the whole population. Therefore, to judge this specification's adequacy, we need to assess the population's degree of homogeneity and whether or not our set of regressor exhausts that heterogeneity. Is the relationship between incumbent votes and economic growth (and other control variables) similar across the 5,565 municipalities? Or is there significant variation among the units, more than it was possible to explain with only our variables? To decide, we can turn to (a) contextual knowledge, (b) graphical analysis, and (c) testing.

(a) Researcher knowledge on the subject matter is the first plausibility test. What we know about the object studied is often informative enough on the degree of heterogeneity or homogeneity across units. For example, according to the literature, is the electoral dynamics in São Paulo's capital is comparable to what happens in Serra da Saudade (in the state of Minas Gerais), population 781? Consequently, consulting theory on a subject may offer clues on the plausibility of pooling.

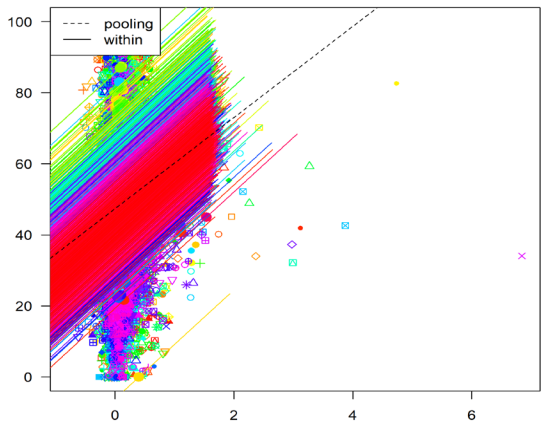
(b) Graphs are useful for exploratory analysis, especially when the N is small. To visualize if we can incorporate all municipalities in a single regression line, we can compare, in the same scatter plot, a single line applied to the whole population vs. separate lines for each municipality. The following example shows a bivariate plot with the main independent variable, "economic growth" (horizontal axis), and the dependent variable "incumbent vote share" (vertical axis).

```
##### GRAPHIC COMPARISON OF POLS VS FE INTERCEPTS #######

mode_fee <- plm(voteshare ~ growth,
                data = DATA, index = c("codibge", "year", group = "codibge"),
                model = "within")
plot(mode_fee)
```

23 The pooled model can also be executed with the *lm* function. However, to execute the tests for assumption check it is necessary that the database is in the format produced by *plm*.

Figure 5 – Pooled OLS and FE comparison for the bivariate model of incumbent vote share x economic growth



Source: the authors, based on data by Fernandes and Fernandes (2017).

The pooling solution (dotted line) proposes a line approximately in the middle of the observations, while the within model (solid colorful lines) better represents the data's variation and indicates that there is a large difference between the intercepts of each municipality.

Another graphic analysis tool is the residuals plot of the final model. Since POLS assumes homogeneity, its residuals must be similarly random for all cases. We can group the regression's residuals by municipality and year as a way to observe if the variance in the error term was more pronounced for certain municipalities or periods, as shown below.

```
##### POOLED OLS RESIDUALS ANALYSIS ###########

library(broom)
library(ggplot2)

POLS_lm <- lm(voteshare ~ growth+growth_state+growth_country+log_gdp+log_gdp_state+
  log_gdp_country+mayor_base_presid+health_exp+log_pop+leec+
  lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespers,
            data = DATA, na.action = "na.exclude")
# running the POLS model with lm is required to add
# the model's residuals to the original data.frame

base <- augment(POLS_lm, data=DATA)
# with the 'augment' command in the broom package, we can join
# the model's residuals with the original database's variables.

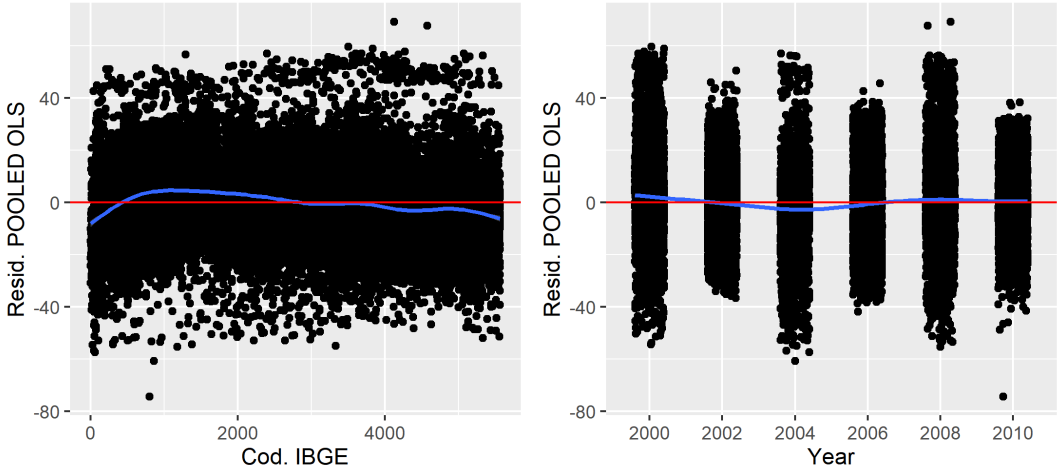
                  #### Municipality ####

ggplot(base, aes(x=as.numeric(as.factor(codibge)), y=.resid))+
  geom_point()+
  geom_smooth()+
  geom_hline(yintercept = 0, color="red") + ylab(" Resid. POOLED OLS") +
  xlab("Cod. IBGE")
                  ####   Year  ####

base$year <- as.numeric(base$year)

ggplot(base, aes(x=jitter(year), y=.resid))+
  geom_point()+
  geom_smooth()+
  geom_hline(yintercept = 0, color="red") + ylab("Resid. POOLED OLS") +
  xlab("Year") + scale_x_discrete(limits=c("2000", "2002", "2004",
                                           "2006", "2008", "2010" ))
```

Figure 6 – POLS model residuals grouped by municipality and year



Source: the authors, based on Fernandes and Fernandes (2017).

The municipalities' residuals are not similar: there is something systematically differentiating the municipalities with IBGE codes lower than 3,000 from those with higher code values. The variance between electoral years is also different. The reason can be seen quickly: the variance is larger for municipal election years (2000, 2004, and 2008) and smaller for presidential ones (2002, 2006, and 2010). This illustrates the issue with pooling and heterogeneity, since we are trying to include mayoral and presidential elections in the same predictive model. These results encourage us to consider the inclusion of dummies for each year or election type.

(c) The *plm* package enables the implementation of different tests with the goal of checking, in a null/alternative hypothesis way, if there are significant individual effects. Both the F-test and the Breusch-Pagan LM test can be applied. The F-test checks if there are significant differences between the POLS and FE models, with the null hypothesis being that POLS is superior. As can be seen by the result of the p-value (<0.05), the FE model is significantly superior.

```
##### POLS F-TEST ##############################

pFtest(mode_fe, POLS) # F-test with the fixed effects and Pooled models
##
##  F test for individual effects
##
## data:  form
## F = 1.4214, df1 = 5548, df2 = 20786, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

The Breusch-Pagan LM test helps us to choose between POLS and RE.[24] The test's null hypothesis predicts that the municipality-specific variance is zero, which would recommend POLS. According to the test's result, H0 is rejected, and POLS can then be ruled out (p-value < 0.05).

```
##### BREUSCH-PAGAN TEST ########################

plmtest(POLS, type="bp", effect = "individual") # Breusch-Pagan test + individual effects
##
##  Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels
##
## data:  form
## chisq = 107.15, df = 1, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

The comments on individual effects also apply to temporal effects. According to the same logic, the goal is to check if each year in the study has specific idiosyncrasies that make observations in that interval have some peculiar heterogeneity. This corresponds to the notion of shocks in time. When assuming both temporal and individual effects, we have the so-called "two-ways" effects. In the tests described above, the presence of these effects can be tested by modifying the "effect" argument, as shown below:

24 The Breusch-Pagan LM test is one of the more traditional ones, but the *plmtest* command can calculate its extensions. These variations are useful depending on the configuration of the data in question (balanced vs. unbalanced, short vs. long). See Croissant and Milo (2018, p. 86).

```
plmtest(POLS, type="bp", effect = "twoways") # Breusch-Pagan test + individual and
time effects
##
##  Lagrange Multiplier Test - two-ways effects (Breusch-Pagan) for
##  unbalanced panels
##
## data:  form
## chisq = 2762.9, df = 2, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Given the results of the graphic analysis and the Breusch-Pagan test, we know that the models tested present spatial and temporal heterogeneity. Thus, we executed them again adding new dummies. In the original study, the authors added to the model dummies for the years and to differentiate municipal and presidential elections (Fernandes and Fernandes 2017, p. 665). Although in most situations annual dummies are the standard approach to capture temporal heterogeneity, the specificity of this database recommends that we use only the election type dummy.[25] It should enhance mainly the POLS and FD models. Meanwhile, for FE and RE, we altered the effect for "two-ways", which will introduce effects by municipality and years, implicating in time dummies such as this and variables such as the national GDP possibly being omitted due to collinearity. With the addition of these variables, it is possible to control the heterogenous effects identified previously.

25 For these data, dummies for each year will very likely have accentuated collinearity with some of the variables, especially variables aggregated at the national level (such as the national GDP, which will have a fixed value every year) and even the new dummy for municipal elections.

```
##### MODELS WITH THE ELECTION VARIABLE ###############

# variable to indicate if the election is pres (0) or munic (1)
DATA$type_ele <- ifelse(DATA$year == 2000 |
                        DATA$year == 2004 |
                        DATA$year == 2008, 1, 0)

# formula with election type dummy
form_d <- as.numeric(voteshare) ~ growth+growth_state+growth_country+log_gdp+log_
gdp_state+
  log_gdp_country+mayor_base_presid+health_exp+log_pop+leec+
  lheu+lses+laseps+ldesporc+ldespcor+linvest+ldespers
  type_ele

##### POOLED OLS W DUMMIES ###############

POLS_dic <- plm(form_d, data = DATA, model = "pooling")


##### FD W DUMMIES ########################

mode_fd_dic <- plm(form_d, data = DATA, model = "fd")


##### FE TWO-WAYS #########################

mode_fe_2w <- plm(form_d, data = DATA, model = "within", effect = "twoways")


##### RE TWO-WAYS #########################

mode_re_2w <- plm(form_d, data = DATA, model = "random", effect = "twoways", random.
method = "walhus")
```

4.2 Fixed or random effects?

With POLS ruled out, we can consider FE or RE. As we have pointed out, for this choice it is important to evaluate both the nature of the data as well as the intensity of the correlation between $_i$ and X. To do so, we can turn to (a) contextual knowledge, (b) relevance or requirement of certain variables of interest, and (c) the Hausman test.

(a) As explained in section 3.2, the researcher must answer the following question on the data's origin: can I consider them a random sample of a broader population? Thinking in exogeneity terms, do we have good reason to expect that some municipalities, just by virtue of being, let's say, Salvador or Sirinhaém, have unobserved peculiarities (e.g., a certain party's historic hegemony) that correlate with the selected variables?

(b) One of FE's limitations (also FD's) is not allowing testing elements that are invariant in time. Where there is substantive interest in testing this type of variable, RE can be considered, always taking into account the trade-off between efficiency and bias. The model by Fernandes and Fernandes (2017) does not use variables fixed in time, but if we wished to go further and test, for instance, control dummies for Brazilian regions

(North, Northeast, Midwest, Southeast, and South), we could not use FE.

(c) The *phtest* command runs the specification test by Hausman (1978), comparing FE and RE models, with the null hypothesis that both are consistent. The null hypothesis will be rejected if the correlation between unobserved individual heterogeneities and independent variables is found. In our case, the null hypothesis that both RE and FE are consistent is rejected. Therefore, we know that the unobserved individual heterogeneity and the regressors are correlated, meaning only FE will have consistent estimators. It is crucial to keep in mind that the Hausman test, albeit the standard one to differentiate between FE and RE, assumes models are well-specified, free of bias, and homoscedastic.[26]

```
##### HAUSMAN TEST ##############################

phtest(mode_fe_2w, mode_re_2w) # Hausman test (fixed and random models)
##
##  Hausman Test
##
## data:  form_d
## chisq = 851.98, df = 15, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
phtest(form_d, data= DATA, method="aux", vcov=vcovHC) # robust version of the test
##
##  Regression-based Hausman test, vcov: vcovHC
##
## data:  form_d
## chisq = 810.83, df = 18, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

4.3 Handling temporal issues

As described in section 3.4, temporal issues are usually (a) non-stationarity and (b) serial correlation of the residuals. The former concerns the raw variables in the database and the latter, the residuals after running the models. Given this sequence, one must first check for stationarity.

(a) Long panels may benefit from tests such as ACF, PACF, ADF, or Panel Unit Root Tests to check stationarity. Since our panel is short (T=6), these tests would not be very informative. If we still want to check if our data are stationary or not, we can look for overall trends towards irreversible growth or retraction in the variables over time. We have done this visual inspection below for the dependent variable "vote share", showing values by municipality and an overall mean for each year. We can have some expectation for a growth trend over time in the incumbent's vote share.

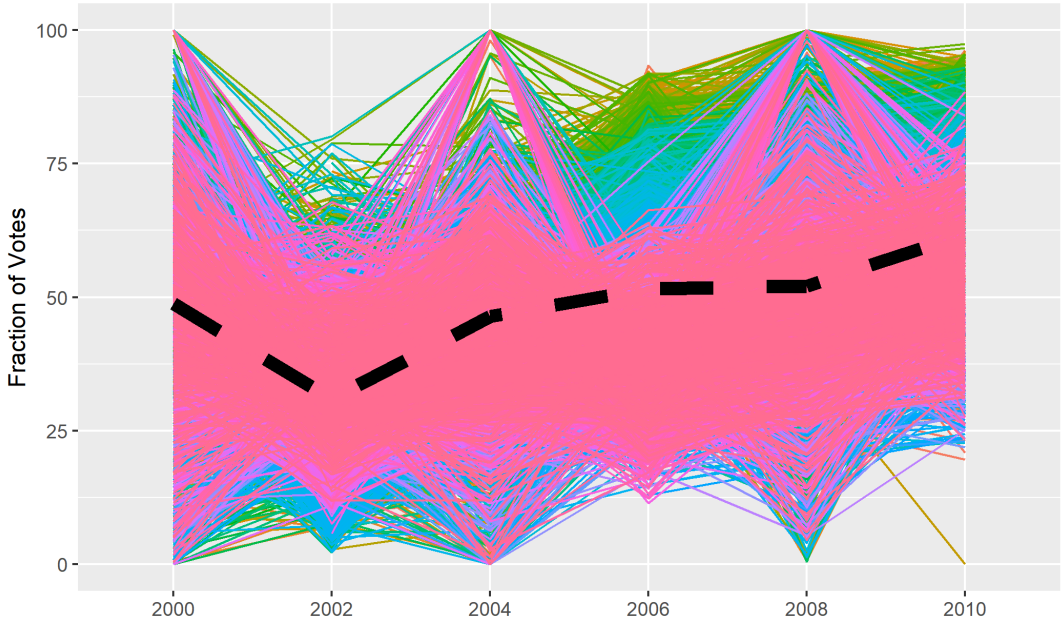
26 If the requirements of absence of misspecification, independence between the error term and independent variables, absence of serial correlation, and heteroskedasticity are not met, it is recommended, in addition to a model revision, the use of robust forms of the test. In *plm*, this is done with the arguments *phtest(..., method="aux")* or yet *phtest(..., method="aux", vcov=vcovHC)* for customized variance-covariance matrices.

```
##### CHECKING STATIONARITY #############

library(dplyr)
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:plm':
##
##     between, lag, lead
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
data_est <- DATA %>%
  group_by(year) %>%
  mutate(meanvotes = mean(voteshare, na.rm=T)) #database with the vote mean

ggplot(data_est, aes(x=year, y=voteshare, group=codibge, color=codibge))+
  geom_line() + geom_line(aes(x=year, y=meanvotes, group=1),
  linetype="dashed", size=3, color="black")+ ylab("Fraction of Votes") +
  xlab("") + guides(color="none")
```

Figure 7 – Vote share over time (municipalities and overall mean)



Source: the authors, based on original data by Fernandes and Fernandes (2017).

If there is a trend, the inclusion of a time trend is recommended, usually a T variable that takes on the values of 1 for the first year, 2 for the second, and so on.

(b) So far, our selection between the POLS, FE, RE, and FD estimators has been based on how well each approach handled unobserved heterogeneity. Now, success in eliminating temporal correlations also becomes important in choosing an estimator. As explained in 3.4, there are several tests for serial correlation that can be run on residuals, depending on which type of panel (short or long) and the estimator employed. The *plm* package cur-

rently offers eight tests for serial correlation.[27] Since the Hausman test has already told us to abandon RE, we will run only two tests, focusing on FE and FD.

To check the presence of serial correlation on fixed effects models, we used the *pwartest* function. We can see in the output that the test's alternative hypothesis is that there is serial correlation. The p-value was lower than 0.05, indicating the rejection of the null hypothesis (there is no serial correlation). That is, according to the test results, the fixed effects model presents serial correlation.

```
##### TESTING SERIAL CORRELATION FOR FE ########################

pwartest(mode_fe_2w) #pwartest for FE model
##
##  Wooldridge's test for serial correlation in FE panels
##
## data:  mode_fe_2w
## F = 71.527, df1 = 1, df2 = 20801, p-value < 2.2e-16
## alternative hypothesis: serial correlation
```

The *pwfdtest* function is used to check the presence of serial correlation in FD models. The null hypothesis is that there is no serial correlation related to the errors. Given the p-value (< 0.05), we reject the null hypothesis. That is, the FD model presents serial correlation issues.

```
##### TESTING SERIAL CORRELATION FOR FD ######################

pwfdtest(mode_fd_dic) #pwfdtest for FD model
##
##  Wooldridge's first-difference test for serial correlation in panels
##
## data:  mode_fd_dic
## F = 6004.9, df1 = 1, df2 = 15307, p-value < 2.2e-16
## alternative hypothesis: serial correlation in differenced errors
```

The tests' results suggest that both FE and FD harbor serial correlation in the error terms. Similarly, the inspection of the raw data indicates an ascending trend, that is, non-stationarity. To remedy the situation, we could consider including a time trend (however, as we saw in 3.2, these cannot be estimated by FE or FD), LDV or employing standard errors robust against serial correlation. As seen previously, temporal issues tend to be more worrisome according to T>N. In our case, given that N is much bigger than T, and to maintain the example's parsimony, we will preserve the current variables and use robust standard errors at the end.

4.4 Handling spatial issues

Having dealt with unobserved heterogeneity, it is expected that remaining spatial issues will be about panel heteroskedasticity and contemporaneous correlation of the errors. The same visual inspection used in section 4.1, grouping residuals by municipality, also informs this decision. Figure 6 has already demon-

27 The six tests not covered in this article are: the generic Wooldridge test for unobserved effects, applicable to POLS (*pwtest*); the Breusch-Godfrey/Wooldridge test for serial correlation, preferably run on RE, but also long FEs (*pbgtest*); the Durbin-Watson test, also preferable for RE (*pdwtest*) and its generalization by Bhargava, Narendranathan, and Franzini (*pbnftest*); the locally-robust test by Bera, Sosa-Escudero, and Yoon for serial correlation or random effects (*pbsytest*); Baltagi and Li test for serial correlation (*pbltest*).

strated that the variance of the residuals between municipalities is systematic.

In addition to the visual inspection, we can turn to specific tests. The *plm* package offers a general command to test cross-sectional dependence (*pcdtest*), wherein the user must choose the type of test according to the panel's characteristics. Since we have a short T and a long N,[28] the most appropriate test is the one proposed by Pesaran (2004).

```
##### TESTING PANEL HETEROSKEDASTICITY ############################

pcdtest(mode_fe_2w, test = "cd") # Fixed Effects model
##
##  Pesaran CD test for cross-sectional dependence in panels
##
## data:  as.numeric(voteshare) ~ cresc + crescuf + crescbr + lpibreal +     lpibuf
+ lpibbrasil + prefeitobasepresidente + persaude +     lpop + leec + lheu + lses +
laseps + ldesporc + ldespcor +     linvest + ldespes + tip_ele
## z = -0.66933, p-value = 0.5033
## alternative hypothesis: cross-sectional dependence
```

The results show that our FE two-ways model is free of contemporaneous dependence, having only serial correlation issues. We must use standard errors robust against serial correlation. We can access different robustness options in *plm*, depending on the type of panel, the way the data are grouped, and the type of problem to be fixed (see Croissant & Millo 2018, chapter 5; Henningsen & Henningsen 2019, pp. 382-384). In our case, because we want to deal with the serial correlation of an FE with two-ways effects, we run the *vcovDC* matrix to estimate robust standard errors:

```
##### HANDLING SERIAL CORRELATION #############

library(lmtest)
mode_fe2w_rbst <- coeftest(mode_fe_2w,
                           vcov = vcovDC)
# adding 'double cluster' robust errors
```

4.5 Results

We present the results of all models in Table 1. To enable the comparison of the original standard errors in the FE model with the robust ones, we put both in separate columns in the table.

Since our main interest is educational and not substantive, we show only a summarized list of variables. Therefore, of the more than 20 variables in the original model, we kept only 11 for the table.

28 Breusch-Pagan LM, activated by the argument test="lm", is recommended for long T and small N; scaled-Breusch-Pagan LM, "sclm", for long T and N; "rho" and "absrho" provide correlation coefficients between the residuals between pairs of observations (see Croissant & Millo, 2018, section 4.4).

Table 1 – Replication results

	Vote Share				
	POLS	RE (2W)	FE (2W)	FE (2W, Robust.)	FD
	(1)	(2)	(3)	(4)	(5)
Municipal growth	1.683***	1.098***	1.148*	1.148	0.190
	(0.608)	(0.040)	(0.652)	(1.517)	(0.669)
State growth	0.095***	0.086***	-0.058*	-0.058	-0.144***
	(0.027)	(0.002)	(0.030)	(0.163)	(0.033)
National growth	3.660***	3.709***			4.231***
	(0.090)	(0.088)			(0.091)
Mayor Base Presid.	1.211***	0.742***	1.289***	1.289***	1.136***
	(0.208)	(0.014)	(0.235)	(0.409)	(0.254)
% Budget Health	0.040***	0.121***	0.044*	0.044	0.058***
	(0.014)	(0.001)	(0.023)	(0.035)	(0.019)
Pop. Munic. (log)	0.302	-0.415***	-6.037***	-6.037**	-3.360*
	(0.258)	(0.018)	(1.449)	(2.609)	(1.824)
Expenditures Educ. Cult. (log)	6.917***	6.461***	1.543***	1.543	0.162
	(0.358)	(0.025)	(0.525)	(1.284)	(0.553)
Expenditures Invest. (log)	2.245***	1.983***	1.430***	1.430	1.644***
	(0.136)	(0.009)	(0.165)	(0.903)	(0.180)
Expenditures Personnel (log)	-1.424***	-0.335***	3.304***	3.304	1.167*
	(0.493)	(0.035)	(0.635)	(2.264)	(0.698)
Municipal Election	1.074***	0.828***			-0.915***
	(0.277)	(0.246)			(0.264)
Intercept	102.026**	152.778***			-4.350***
	(43.701)	(37.792)			(0.373)
N Obs.	26.352	26.352	26.352		20.803
R^2	0.321	0.319	0.039		0.254
Adjust. R^2	0.320	0.318	-0.219		0.253
F-Test	691.129*** (df = 18; 26333)	12.309.190***	55.646*** (df = 15; 20783)		392.979*** (df = 18; 20784)

*** p < 0.001; ** p < 0.01; * p < 0.05

Source: the authors, based on Fernandes and Fernandes (2017). The value of theta for the random effects two-ways model was $\theta = 0.11$.

The specification tests lead us to be more confident in the results obtained via FE. Nevertheless, the comparison of all the models is informative. The POLS and FE results tend to be much closer to each other for many of the regressors, while the FE and FD estimates are more similar. This is expected since, as we saw, the latter two remove individual heterogeneity in a similar way. Thus, reading the POLS/RE and FE/FD results is a way of checking how the estimates behave as we control more and more the municipal idiosyncrasies; the "net effect",

as it were, of the variables, all the peculiarities of the cases being removed.

In some variables the difference is manifested in sign change. For example, the POLS and RE models state that there would be positive effects from the state's economic growth rate on the vote share, or negative ones with personnel expenditures. We have a case of bias given that, controlling for the heterogeneities of each municipality with FE and FD, we see that their impact is actually the inverse.

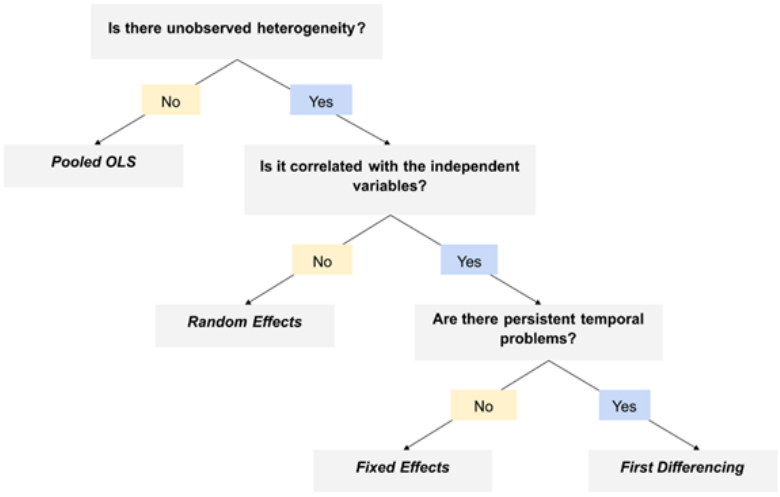
For other regressors, a change in the strength of the relationship is noticed. Sometimes, the effect becomes weaker (e.g., expenditures in education and culture, expenditures in investment), occasionally stronger (dummy for the same mayor-president base, population) and, every so often, stable (percentage of the health budget). Furthermore, while POLS and RE tend to indicate more statistical significance, FD and mainly FE with robust standard errors are more conservative. Notably, the main variable of interest in this study, municipal economic growth, does not retain significant effect on votes on the incumbent.

Lastly, it must be highlighted that the POLS and RE intercepts present unrealistic values due to the fact that the data were not centralized, and that, as we anticipated, FE two-ways omitted time dummies (election type) and those tied to the year (national GDP) due to collinearity.

4.6 Summary of the procedure

To conclude this tutorial, we can represent a model selection process as a decision tree similar to Figure 8.

Figure 8 – How to choose the best estimator?



Source: the authors.[29]

Schematically, first we evaluate to what extent there is unobserved heterogeneity among the database's entities. This can be done with contextual knowledge of the problem, graphic analysis of the intercepts, and by employing specific tests (F and Breusch-Pagan LM). Once POLS is ruled out, we must estimate if there is correlation between the unobserved heterogeneity and the set of explanatory variables, for example, using the Hausman test. If there is no correlation, the use of RE is justified. On the other hand, in the pres-

29 This tree is inspired by a more complex original, which the authors credit to professor Claudia Schwirplies (Uni HH).

ence of correlation between the errors and the independent variables, we can consider using FE or FD. The option between these must take into account the model's temporal issues, as well as the database's size. At the end of the process, tests for time and spatial dependencies are required to check for the need of robust standard errors.

This strategy is a simplification for educational purposes and emphasizes the main decisions that, if ignored, could generate bias or inefficiency. However, as we argue throughout the paper, substantive considerations must also guide that choice.

5. Conclusion

This article presented an introduction to the logic of panel data. We showed the main benefits that come from the analysis of longitudinal information and described the most frequent issues that arise when the dataset gathers, simultaneously, spatial and temporal dimensions. Subsequently, we demonstrated how to implement a regression analysis with panel data in R. Because it is a article with educational purposes, we have also shared all the data and scripts, so that students, professors, and researchers can easily adapt our computational routines to their substantive interests.

At the beginning of this article, we indicated how the use of panel data in Brazilian PIR is still incipient. Thus, we believe it is relevant to explain what can be done to change those circumstances. Among the different options available, we highlight the following: (1) restructure the curricula of undergraduate and graduate courses with the aim of including specific calculus, statistics, and computation classes and offer electives on how to work with longitudinal data, such as professor William Green's class[30]; (2) if item 1 cannot be implemented, ensure the offer of special summer courses such as "Modelling Dynamics" given by professor Lorena Barberia (USP), "Modelling Dynamics in Space and Time" given by professors Guy Whitten and Lorena Barberia (USP), and "Panel Data Analysis" by Igor Viveiros in the MQ/UFMG course;[31] (3) regionally diversify the intermittent offer of data analysis courses. Item 3 being impossible to implement, ensure financial support so that masters and PhD students, specially from the North and Northeast of Brazil, can attend summer schools; (4) encourage the acceptance of online courses as valid credits for graduate curricula; (5) special dossiers on research methods edited by the main PIR journals, with a special emphasis on intuitive and educational articles, such as volume 142 of the Journal of Econometrics.[32] In addition, set space aside in the best PIR journals for the publication of papers with methodological emphasis such as Beck and Katz (1995); and (6) incentivize researchers that are dedicated to the publication of research methods and techniques in PIR through specific financing via funding institutions (CAPES and CNPq).

We are aware of the operational and logistical difficulties that surround each of these initiatives. Nevertheless, considering the potential benefits that may be incorporated in our models of analysis and the limited use of longitudinal data in PIR empirical research, we believe it is about time we take time seriously (Beck, Katz, and Tucker, 1998; Boef and Keele, 2008).

30 http://people.stern.nyu.edu/wgreene/Econometrics/PanelDataNotes.htm

31 http://summerschool.fflch.usp.br/ipsa-usp-summer-school-2017-2/

32 https://www.sciencedirect.com/journal/journal-of-econometrics/vol/142/issue/2

BIBLIOGRAPHICAL REFERENCES

ACHEN, C. H. (2000), "Why lagged dependent variables can suppress the explanatory power of other independent variables". In: *annual meeting of the political methodology section of the American political science association, UCLA*.

BALTAGI, Badi H. (2005), *Econometric Analysis of Panel Data*. England, John Wiley & Sons Ltd.

BERRY, William D. (1993). *Understanding regression assumptions*. Thousand Oaks, Sage.

BECK, Nathaniel; KATZ, Jonathan N. (1995), "What to do (and not to do) with time-series cross-section data", *American political science review*, v. 89, n. 3: 634-647.

BECK, Nathaniel; KATZ, Jonathan N.; TUCKER, Richard. (1998), "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable", *American Journal of Political Science*, v. 42, n. 4: 1260-1288.

BECK, Nathaniel. (2001), "TIME-SERIES–CROSS-SECTION DATA: What Have We Learned in the Past Few Years?", *Annual Review of Political Science*, v. 4, n. 1: 271–293.

BECK, Nathaniel. (2008), "Time-Series Cross-Section Methods", *in* J. Box-Steffensmeier; H. Brady; D. Collier (Eds.). *The Oxford Handbook of Political Methodology*. New York, OUP: 475–493.

BLALOCK, Hubert M. The presidential address: Measurement and conceptualization problems: The major obstacle to integrating theory and research. *American sociological review*, v. 44, n. 6, p. 881-894, 1979.

DE BOEF, Suzanna; KEELE, Luke. (2008), "Taking time seriously", *American Journal of Political Science*, v. 52, n. 1: 184-200.

BELL, Andrew; JONES, Kelvyn. (2015), "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data", *Political Science Research and Methods*, v. 3, n. 1: 133–153.

BELL, Andrew; FAIRBROTHER, Malcolm; JONES, Kelvyn. (2019), "Fixed and random effects models: making an informed choice", *Quality & Quantity*, v. 53, n. 2: 1051-1074.

BENOIT, K. et al. (2018), "quanteda: An R package for the quantitative analysis of textual data", *Journal of Open Source Software*, v. 3, n. 30: 774.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. (2016), *Time series analysis: forecasting and control*. Hoboken/New Jersey, Wiley.

CLARK, Tom S.; LINZER, Drew A. (2015), "Should I Use Fixed or Random Effects?", *Political Science Research and Methods*, v. 3, n. 2: 399–408.

CROISSANT, Yves; MILLO, Giovanni. (2008), "Panel data econometrics in R: The plm package", *Journal of statistical software*, v. 27, n. 2: 1-43.

CROISSANT, Yves; MILLO, Giovanni. (2018), *Panel Data Econometrics with R*, Wiley.

CERQUEIRA, Daniel et al. (2017), *Atlas da violência 2017*.

DIELEMAN, J. L.; TEMPLIN, T. (2014), "Random-Effects, Fixed-Effects and the within-between Specification for Clustered Data in Observational Health Studies: A Simulation Study", *PLoS ONE*, v. 9, n. 10: e110257.

FÁVERO, L. P. L. (2013), "Dados em painel em contabilidade e finanças: teoria e aplicação", *BBR-Brazilian Business Review*, v. 10, n. 1: 131-156.

FERNANDES, Ivan Filipe de Almeida Lopes; FERNANDES, Gustavo Andrey de Almeida Lopes. (2017), "A importância do crescimento econômico local na escolha do chefe do Executivo no Brasil", *Revista de Administração Pública*, v. 51, n. 4: 653-688.

FINKEL, S. E. (1995), *Causal analysis with panel data* (No. 105). Sage.

FORTIN-RITTBERGER, J. (2013), "Time-Series Cross-Section", *in* H. Best; C. Wolf (Eds.), *The SAGE Handbook of Regression Analysis and Causal Inference*. London, SAGE Publications: 387–408.

GUJARATI, Damodar N. *Basic Econometrics*. McGraw-Hill Companies. New York, 2004.

HAUSMAN, Jerry A. (1978), "Specification tests in econometrics", *Econometrica: Journal of the econometric society*, v. 46, n. 6: 1251-1271.

HENNINGSEN, Arne; HENNINGSEN, Géraldine. (2019), "Analysis of Panel Data Using R", *in* M. Tsionas (Ed.) *Panel Data Econometrics: Theory*. London, Elsevier: 345-396.

HSIAO, C. (2003), *Panel data analysis*, Cambridge University Press.

IBGE. (2017), Pesquisa Nacional por Amostra de Domicílios Contínua. Disponível em: https://www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html?=&t=o-que-e. Acesso em 30-06-2020

KING, Gary. (2001). "Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data", *International Organization*, v. 55, n. 2: 497-507.

KING, Gary. (1995), "Replication, replication", *PS: Political Science and Politics*, v. 28, n. 3: 444-452.

KING, Gary; KEOHANE, Robert O.; VERBA, Sidney. (1994), *Designing social inquiry: Scientific inference in qualitative research*. Princeton University Press.

KENNEDY, Peter. (2003), *A guide to econometrics*. MIT Press.

KRUEGER, James S.; LEWIS-BECK, Michael S. (2008), "Is OLS dead?", *The Political Methodologist*, v. 15, n. 2: 2-4.

LAROCCA, R. (2012). "Gauss–Markov Theorem" in SALKIND, N. J. (ed) *Encyclopedia of Research Design*. Thousand Oaks, Sage, 529-533.

LIJPHART, A. (1971), "Comparative Politics and the Comparative Method", *American Political Science Review*, v. 65, n. 3: 682–693.

MARQUES, Luís D. (2000), "Modelos Dinâmicos com Dados em Painel: revisão de literatura". *Working paper*. Available at http://wps.fep.up.pt/wps/wp100.pdf. Accessed 23-06-2020.

MENARD, Scott. (2002), *Longitudinal research*. Thousand Oaks, SAGE Publications.

MEDEIROS, M.; BARNABE, I.; ALBUQUERQUE, R.; LIMA, R. (2016), "What does the field of International Relations look like in South America?", *Revista Brasileira de Política Internacional*, v. 59, n. 1: e004

MEIRELES, Fernando; SILVA, Denisson; BARBOSA, Rogerio. (2019), *rscielo: A Scraper for Scientific Journals Hosted on Scielo*. R package version 1.0.0. https://CRAN.R-project.org/package=rscielo.

MESQUITA, Rafael. (2018) *Liderança regional em perspectiva comparada: Brasil e Turquia*. PhD Thesis, UFPE. Available at https://repositorio.ufpe.br/handle/123456789/32942. Accessed 29-05-2021

MUMMOLO, J.; PETERSON, E. (2018). "Improving the interpretation of fixed effects regression results", *Political Science Research and Methods*, v. 6, n. 4: 829-835.

PESARAN, M. H. (2004), "General diagnostic tests for cross section dependence in panels", *CESifo Working Paper Series*, 1229.

PEARL, J. (2000), *Causality: Models, reasoning, and inference.* New York, Cambridge University Press.

SKRONDAL, A., RABE-HESKETH, S. (2008), "Multilevel and Related Models for Longitudinal Data", *in* Leeuw J.., Meijer E. (Eds.) *Handbook of Multilevel Analysis*, New York, Springer: 275-300.

STOCK, James H.; WATSON, Mark W. (2004), *Econometria*. Pearson.

VERBEEK, Marno; NIJMAN, Theo. (1992), "Testing for selectivity bias in panel data models", *International Economic Review*, v. 33, n. 3: 681-703.

WOOLDRIDGE, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MIT Press.

WOOLDRIDGE, J. M. (2013), *Introductory econometrics: a modern approach*. Australia, South-Western Cengage Learning.

WORRALL, John L. (2008), "An introduction to pooling cross-sectional and time series data", *in* S. Menard (Ed.) *Handbook of longitudinal research: Design, measurement, and analysis*, Elsevier: 233-248.

Questions and Answers

1. **Should I use fixed or random effects?**

Generally, random effects can introduce bias, that is, increase the chance of systematically under- or overestimating the parameter of interest. However, this method tends to reduce the coefficients' variance, which means more efficient estimates (Kennedy, 2003). Therefore, models with random effects tend to produce more consistent estimates when the sample is small and when the correlation between the independent variable and the effects of the observations of relatively small (Clark and Linzer, 2015). On the other hand, fixed effects tend to produce unbiased estimates, especially for large samples (Mummolo and Peterson, 2018). A disadvantage of fixed effects is the reduction of the model's degree of freedom given the inclusion of dummies. However, if the sample is sufficiently large, this will not be a problem in estimating the coefficients. For more information: Wooldridge (2002) and Clark and Linzer (2015).

2. What is the difference between an unbalanced and a balanced panel?

Stock and Watson (2004) define a balanced panel as a dataset in which there are no missing cases, that is, we have information for all observed units (N) for all time periods (T). A panel is unbalanced when, for whatever reason, discontinuities in the data matrix occur, whether due to lack of observations or the absence of time units. Wooldridge (2002) warns that, overall, the statistical treatment will be the same for both panel types. In all cases, the research must inform the reason for the unbalanced panel and detail the criteria for sample selection. For an introduction on the subject, see Mayer (2010).[1] For a detailed discussion on the specificities of the unbalanced panel, Wooldridge (2002) and Baltagi (2005). For a more advanced treatment on selection bias in panel data, see Verbeek and Nijman (1992).

3. Short or long panels?

The panel's amplitude depends on the ratio between entities (N) and time units (T). In a short panel, we have many observed entities over a relatively short period of time. Imagine a dataset with detailed information for all Brazilian municipalities between 2000 and 2010. Undoubtedly, it is a lot of information, but since the number of entities (over five thousand municipalities) is much higher to the number of time frames (ten years), we have a short panel where the cross-sectional dimension is dominant. A long panel is present when the offer of cross-section data is significantly smaller than the temporal range of the information. For example, imagine a dataset of the country's five regions (N, NE, MW, SE, and S) gathered annually between 1950 and 2020. That composition, in which T>N, brings us closer to a time series trend, requiring specific techniques. For more information, see Cameron and Trivedi (2005)[2] and Stock and Watson (2004).

1 See: <https://homepage.univie.ac.at/robert.kunst/pan2010_pres_mayer.pdf>.

2 See: <http://public.econ.duke.edu/~vjh3/e262p_07S/readings/Cameron_&_Trivedi_Microeconometrics_Chapters_2.pdf>.

4. How does one avoid omitted variable bias with panel data?

As with other research designs, it is necessary to be up to date on the field's specific literature to know the theoretical specification of the models. Often we can find an important variable in a database that already exists. Typically, FE and FD offer good safeguards against omitted variables when the important unobservables are fixed to the units. Even then, their use does not substitute authoritative knowledge of the literature. It is important to remember that the inclusion of an irrelevant variable in the explanatory model will only harm the efficiency of the coefficients. The standard error will be higher than it should be and that reduces the reliability of the significance tests, but $\hat{\beta}$ will still provide a consistent estimate of β, that is, the inclusion of an irrelevant variable does not cause bias. On the other hand, the exclusion of a theoretically important variable tends to produce biased estimates, especially when the omitted variable is correlated with the remaining variables included in the model.

5. What is the Hausman test and what is for?

The goal of the Hausman test is to evaluate if there is correlation between the errors of the model and the independent variables. Within the context of panel data, the test is employed to guide the decision between fixed and random effects. The null hypothesis states that random effects are more suitable. Operationally, we must estimate both models (with fixed and random effects) and after comparing the estimates. A significant result (p-value<0.05) indicates that the null hypothesis must be rejected, that is, we must choose fixed effects. On the other hand, a non-significant result (p-value>0.05) implies not rejecting the null hypothesis, which means adopting the model with random effects. The Hausman test usually comes with a chi-square distribution with a significance test, in addition to the estimates of the coefficients. For more, see Hausman (1978).

6. What is serial correlation and why should I be concerned with it?

Although autocorrelation can occur in space, the one that interests us here is the one that happens in time, also known as serial correlation. The Cambridge Dictionary of Statistics defines serial correlation as the dependence between measures of the same entity in longitudinal studies. Usually, the closer the temporal proximity in the measurements, the stronger the magnitude of the correlation. On the other hand, the farther the measurements of time, on average, the weaker the degree of correlation between the values of the series. This association structure of the values in time violates the assumption of independence of the errors in the regression model. The main issue with serial correlation is production of inconsistent significance tests, that is, we will not be able to trust the p-value and the confidence intervals. That is the reason longitudinal data analysis must always be accompanied by estimates of the level of autocorrelation in the series.

7. How do I learn more about panel data analysis?

In Econometrics, we suggest Kennedy (2003), Wooldridge (2002, 2013), Baltagi (2005) and Gujarati (2004). Hsiao (2003) is one of the most complete introductions to the theme. The paper "What to do and (not to do with) with Time-Series Cross-Section Data" by Beck and Katz (1995) is the most widely cited paper in Political

Science (7,414 citations in Google Scholar in June 2021). Fortin-Rittberger (2013) is also an introductory paper. Finkel (1995) gives focus to causal possibilities using longitudinal data. In doubt between fixed or random effects? See Bell and Jones (2015). Mummolo and Peterson (2018) reproduces several studies that employ panel data and demonstrates how to improve the interpretation of the coefficients estimated using fixed effects. See the schedule the IPSA/USP summer school: <http://summerschool.fflch.usp.br/tracks-and-courses/>.

8. **TSCS x panel data?**

Beck (2008) differentiates panel data from TSCS data. According to him, the methods that work well for panels do not necessarily work well for TSCS data and vice-versa. In PIR empirical research, the most frequent combination of cases and time is seen in databases that compile information about countries (N) over years (T) (Beck, 2001). Thus, since data of the TSCS type consist of exactly the same cases over time, a stronger temporal correlation than for panel data is expected. TSCS consists of the collection of information on observational units (individuals, companies, federal units, countries) over time (day, month, year). The bigger the T is in relation to the N, the stronger are concerns over the serial structure of the data. On the other hand, when there are a lot of cases (N) for few time frames (T), autocorrelation problems will be less severe.

9. How do I recover each unit's intercepts when using FE or RE in *plm*?

Use the fixef() and ranef() commands. The values correspond to the deviation in comparison to the general intercept.

10. What is the difference between using the *plm* or the *lme4* and *nlme* packages?

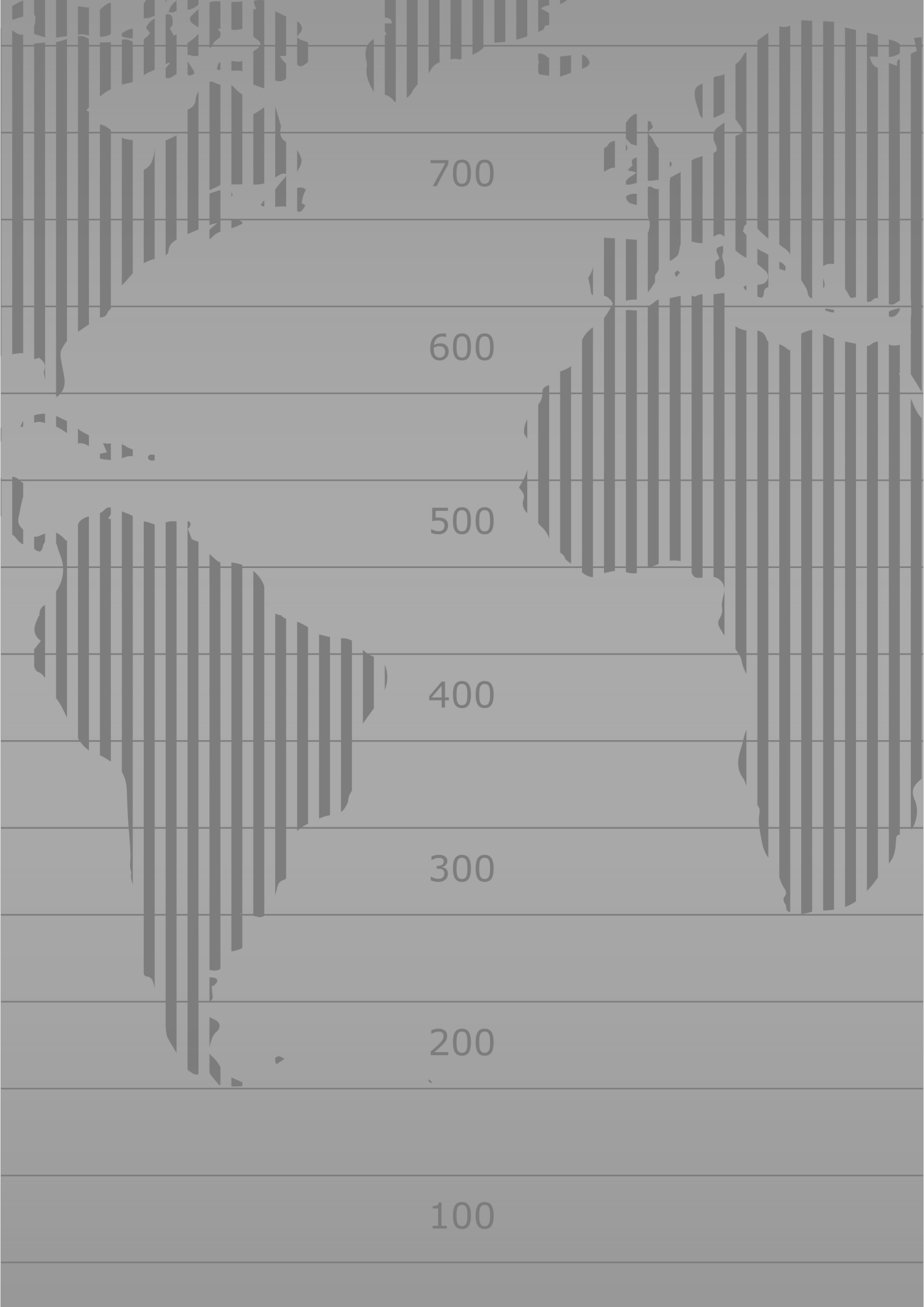
Panel data are not the only way of handling repeated measures. Another approach consists of using mixed linear models, also known as hierarchical or multilevel models, which are executed by packages such as *lme4* and *nlme*. While the panel data perspective is mainly econometric, in other fields such as healthcare, biology, and education, the multilevel approach has a longer tradition as their data are typically hierarchical (e.g., students in classrooms, classrooms in schools). Panel data may be seen as a two-level nested structure: there are individuals *i* (upper level) observed at successive occasions *t* (lower level). Since FE and RE specifications admit intercepts varying by individual, they have a logic close to hierarchical models in which there are random intercepts. In addition to the intercepts, it is possible to work with panel data predicting variable coefficients according to individual or period ($\beta_{\square\square}$) (examples in Wooldridge 2013, p. 451-453; Croissant and Millo 2018, section 8.2), although that possibility is more broadly explored in hierarchical models (random intercepts + random slopes). When we have data configurations in which nesting is not in time or is more complex than in only two levels (e.g., municipalities in states, states in regions), the use of hierarchical models in recommended. From an operational point of view, the ecosystem of R packages is more well-developed around *lme4/nlme* than *plm*, which means it is easier to integrate products from *lme4/nlme* analyses into other graphic, testing, and result presentation packages. For more, see Skrondal and Rabe-Hesketh (2008) and Croissant and Millo (2008, pp. 33-39).

Appendix

Table 1 – Total number of articles collected per journal in the PIR field (2000-2018)

PIR Journal	Qualis (2016)	Total number of articles	% Total
Estudos Avançados	B1	896	12
Revista de Administração Pública	A1	726	9
Sociologias	B1	573	7
Revista Brasileira de Ciências Sociais	A2	559	7
Tempo Social	B1	533	7
Revista de Sociologia e Política	A1	518	7
Caderno CRH	A2	503	6
Sociedade e Estado	B1	483	6
Dados - Revista de Ciências Sociais	A1	469	6
Lua Nova: Revista de Cultura e Política	A2	450	6
Novos estudos CEBRAP	A2	420	5
Revista Brasileira de Política Internacional	A1	403	5
Opinião Pública	A1	340	4
Contexto Internacional	A2	323	4
Revista Brasileira de Ciência Política	B1	235	3
Civitas - Revista de Ciências Sociais	B2	183	2
Brazilian Political Science Review	A2	150	2
Total: 17 journals		7,764	100

Source: the authors


700
600
500
400
300
200
100