SOCIOECONOMIC ANALYTICS

An extensive analysis of Brazil and the Netherlands determinants of football attendance

Matthijs Edel [1] - https://orcid.org/0009-0008-1868-729X

[1] Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, Netherlands
Email: matthijs@edelmail.nl

How to cite this paper: Edel, M. (2024). An extensive analysis of Brazil and the Netherlands determinants of football attendance. *Socioeconomic Analytics,* 2(1) 6-18. https://doi.org/10.51359/2965-4661.2024.260880

RESEARCH ARTICLE

Socioeconomic Analytics
https://periodicos.ufpe.br/revistas/SECAN/
ISSN Online: 2965-4661

Submitted on: 28.12.2023.
Accepted on: 24.01.2024.
Published on: 02.02.2024.

Abstract

Understanding attendance at football stadiums holds great significance for sports economists and football clubs. Consequently, extensive research has been conducted to analyze the factors influencing football stadium attendance. However, much of this research has been confined to short-term analyses or focused solely on European countries. This study seeks to broaden the scope by examining long-term trends in the Netherlands and exploring the dynamics in Brazil. In the Netherlands, factors such as unemployment and overall interest in football emerge as significant determinants of stadium attendance. Surprisingly, hooliganism does not appear to have a notable impact, and the influence of leisure time is unclear. In the Brazilian context, stadium capacity and goal difference do not show significant effects on attendance, and the impact of the club's division is ambiguous.

Keywords

Football stadium attendance, sports economics, ordinary least squares, OLS.

1. Introduction

In many countries, football is the biggest sport, both to play by the people themselves, and to watch professionals play, either at home on the television, or at the stadium. The latter has a whole culture around it. In the eyes of many fans, football is more than just a sport, it is a substantial part of their lives, with strong emotions related to their favorite clubs. A lot of fans buy season tickets and attend every game, whatever the cost may be, home games or away games, even if it is played on the other side of the country, or even in a different country. However, the largest part of a filled stadium is not filled with these people. Most people love watching the game at the stadium, but do not go every time.

This depends on a plethora of factors, investigated by plenty of studies (e.g. García and Rodríguez (2002) and Pawlowski and Nalbantis (2015)). According to Borland and MacDonald (2003), five main categories are consumer preferences, economic factors such as price, quality of viewing, contest characteristics like quality or uncertainty, and capacity. However, most of these studies used match-level for only a few years, like Garc´ıa and Rodr´ıguez (2002) and Pawlowski and Nalbantis (2015), both studying four seasons. Other examples are Pawlowski and Anders (2012), studying only one season, and Martins and Cró (2018), studying five seasons.

Only a few attempts have been made to study the long run developments in stadium attendance. Also, little attention is paid to macroeconomic factors when analysing stadium attendance. Furthermore, international developments in stadium attendance are rarely examined. All of these things are extensively investigated in Van Ours (2021), on which the current paper builds upon. However, in that paper, the assumptions corresponding to the estimation method and the standard errors were not checked. Also, we believe an omitted variable bias may have occurred in that paper. Hence, it might include unreliable results.

Moreover, former studies have mainly focused on stadium attendance in Europe. Hence, not much can be said about other cultures or economies. As explained in Schreyer and Ansari (2021), who performed a scoping review on stadium attendance, examining South America will contribute to the robustness of former findings. They postulate that too little is investigated in South America and that doing so makes findings more generalizable.

Football in South America is extremely popular. It is the biggest sport among amateur players, and professional football also affects the lives in South America a lot, not only on a club level, but also on a national level. South America was the first to hold a regular continental championship (Copa Am´erica), and out of 21 World Championships played, 9 are won by a South American team, and the most World Championships are won by Brazil (5). South America has a plethora of great football players, of which great names like Ronaldinho, Pelé, Ronaldo, Messi, and Súarez are only a few.

Madalozzo (2008) tried to explain stadium attendance in Brazil and found that structure variables, quality variables, performance variables, and an uncertainty variable have the largest effect. Buraimo et al. (2018) did this for Peru, and found that football attendance behaves differently compared to Europe. For example, one of the findings was that price does not affect football attendance significantly. Also, Ferreira and Bravo (2007) did this for Chile, and again found that price does not affect football attendance significantly, whereas some performance indicators and stadium capacity do.

However, as mentioned before, not much research has been conducted in this continent and doing so will contribute to the robustness of former findings. Determining football stadium attendance is of crucial importance for football clubs. Fans supporting the team are sometimes called ‘the 12th man’, and may contribute to the team’s performance. Also, it is quite a substantial part of a club’s revenue. It would be beneficial to explain the attendance for both clubs and sports economists. Sports economist could come up with ideas for policies and clubs could implement this in order to maximise attendance, and to act on it if it is expected to be low due to unchangeable circumstances.

This research is the result from the Bachelor Thesis of Edel (2021) that attempts to examine whether stadium capacity and performance indicators influence football stadium attendance in the Netherlands and Brazil and how these countries compare to each other. Also, for the

Netherlands, this paper investigates whether the unemployment rate, hooliganism, general interest in football, and time spent on leisure influences football stadium attendance. Unfortunately, this analysis is not possible for Brazil due to a shortage of data.

To do so, we use several data sets. For the analysis of the Netherlands, we use a panel data set with club-specific determinants, and a data set with season-specific determinants. For the analysis of Brazil we use a panel data set with club-specific determinants only. For the analysis of the Netherlands, we perform regressions in two stages. In the first stage regression, we use the panel data set, and we use either Fixed Effects, First Differences, or Random Effects to estimate the coefficients, including seasonal dummies. In the second stage, we regress the coefficients corresponding to the seasonal dummies on the season-specific determinants.

For the analysis of Brazil we perform only the first stage. Also, in each of the aforementioned regressions, we test whether the assumptions hold and we use different standard errors if they do not hold.

We found that stadium capacity and performance indicators affect stadium attendance in the Netherlands significantly positively. Furthermore, unemployment affects stadium attendance in the Netherlands significantly negatively, and hooliganism does not seem to have a significant effect. General interest in football has a significantly positive effect, and the effect of time spent on leisure is not that clear. For Brazil, we found that neither stadium capacity nor goal difference affects stadium attendance significantly, and that the effect of the division in which a club plays is not that clear. However, due to a limited number of observations, these results must be carefully interpreted and no real conclusions can be drawn from them.

This paper is one of the first to properly analyse the long run developments in stadium attendance with the inclusion of a macroeconomic factor. Furthermore, this paper adds to the robustness of findings in South America and carefully compares stadium attendance determinants in countries with different cultures. Moreover, this paper is one of the first to properly analyse international football developments. The structure of this paper is as follows. In Section 2, the data is described. In Section 3, the model is specified. Then, in Section 4, all tests needed to test the assumptions are provided. In Section 5, the estimation procedure is explained, and in Section 6, all results are shown. Ultimately, in Section 7, conclusions are drawn and suggestions for further research are given elsewhere.

2. Data Description

We use several data sets, two data sets for an analysis of the Netherlands, one data set for an analysis of Brazil. Due to the scarcity of data of South America, we do not have data to perform an international spillover analysis of South America.

2.1 The Netherlands and Brazil

For the analysis of the Netherlands, we have two data sets. A panel data set with club-specific determinants, and a data set with season-specific determinants. For the analysis of Brazil we have, due to a deficiency of data availability, only one data set, being one with club-specific determinants.

2.1.1 The Netherlands

For the analysis of the Netherlands, we use the same data as Van Ours (2021), which can be found in the Data Repository of Erasmus University Rotterdam (DOI: 10.25397/eur.13634066). The panel data set, including club-specific determinants, contains information of 30 Dutch clubs in seasons 1956/1957 until 2018/2019. These clubs are all the clubs that played professional football in all 63 seasons and played at least one season in the Eredivisie, which is the top league in the Netherlands. The names of these clubs, the number of seasons they played in the Eredivisie, First Division, and Second Division, the average stadium attendance in thousands, and the average stadium capacity in thousands, as proxied by the highest attendance during a season, is shown in Table 1.

Table 1: Club statistics in seasons 1956/1957 until 2018/2019. Adapted from Edel (2021)

Club	Number of Seasons			Attendance	Capacity
	Eredivisie	First Division	Second Division		
ADO Den Haag	45	18	0	9.0	15.7
AFC Ajax	63	0	0	28.6	46.8
AZ Alkmaar	42	19	2	8.9	13.6
De Graafschap	21	34	8	6.5	9.5
FC Den Bosch	15	45	3	4.4	9.3
FC Dordrecht	6	53	4	2.7	5.4
FC Eindhoven	3	58	2	3.3	6.1
FC Groningen	52	11	0	12.6	17.4
FC Twente	61	2	0	13.1	19.8
FC Utrecht	63	0	0	12.3	20.2
FC Volendam	25	38	0	4.5	9.2
Feyenoord	63	0	0	34.0	53.3
Fortuna Sittard	32	31	0	5.9	12.9
Go Ahead Eagles	31	29	3	6.7	12.0
Helmond Sport	2	54	7	2.8	5.2
Heracles Almelo	19	42	2	5.3	8.3
MVV Maastricht	36	27	0	6.2	12.5
NAC Breda	50	13	0	10.5	15.4
NEC Nijmegen	40	15	8	8.2	15.5
PEC Zwolle	19	29	15	5.2	8.5
PSV	63	0	0	23.0	27.4
Roda JC Kerkrade	50	5	8	8.7	15.1
SBV Excelsior	22	37	4	3.2	7.0
SBV Vitesse	34	25	4	10.5	15.8
SC Cambuur	7	52	4	5.6	9.3
sc Heerenveen	27	24	12	10.3	13.4
Sparta Rotterdam	53	10	0	8.9	17.9
Telstar	14	48	1	3.4	7.5
VVV-Venlo	22	37	4	5.1	9.0
Willem II	43	20	0	8.4	13.4

Four clubs played in the Eredivisie throughout the whole period: Ajax, FC Utrecht, Feyenoord and PSV. Helmond Sport has only two seasons in the Eredivisie. PEC Zwolle has been present the most in the Second Division (15 seasons). Average stadium attendance is the highest for Feyenoord (34,000), and the lowest for FC Dordrecht (2,700). Feyenoord also has the highest capacity (53,300), whereas Helmond Sport has the lowest (5,200).

For each of these clubs, we have the average number of attendants in each season, the capacity of each season, and for each season the division they played, the number of points they obtained, their ranking, and the goal difference. Since the capacity might not be a fixed number,

as this could easily be expanded by having additional places to stand or playing in a different stadium, we proxy the capacity of each season by the highest number of attendants during a match in that season. Descriptive statistics of these variables are shown in Table 2.

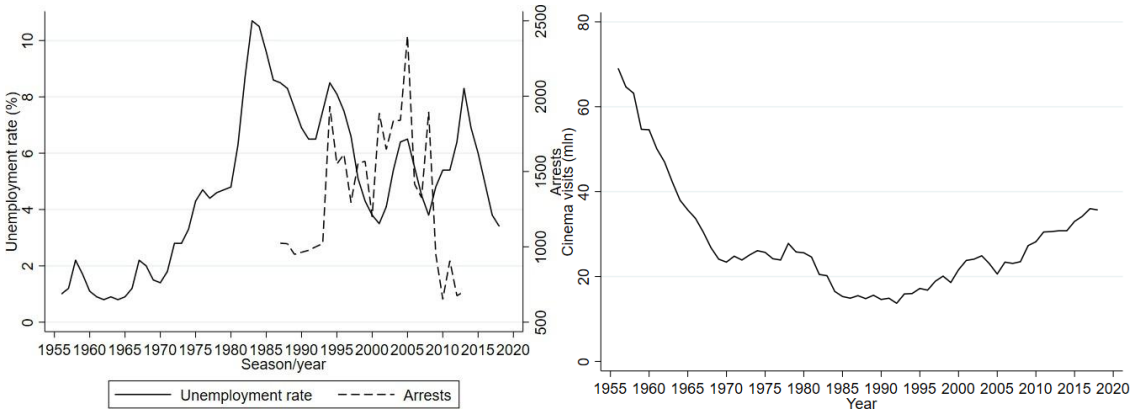
Table 2: Descriptive statistics of club determinants in seasons 1956/1957 until 2018/2019. Adapted from Edel (2021)

Variable	Mean	Minimum	Maximum	Number of observations
Attendance	9,266	661	52,987	1890
Capacity	15,075	1,200	68,000	1890
Dummy for First Division	0.41	0	1	1890
Dummy for Second Division	0.05	0	1	1890
Points/100	0.49	0.13	1.01	1890
Ranking/100	0.09	0.01	0.21	1890
Goal difference/100	0.03	-0.73	0.90	1890

We see that attendance has a mean of 9,266 and ranges from 661 to 52,987 with stadium capacity having a mean of 15,075 and ranging from 1,200 to 68,000. Also, we notice that 41% of the observations is from the First Division and 5% from the Second Division. Furthermore, we observe that the least number of points at the end of the season is 13 and that the highest number of points is 101, with an average of 49. The ranking ranges from 1 to 21 (first place to last place in the First Division in 1971/1972) with an average of 9. The worst goal difference is 73, whereas the best is 90. The average is 3.

The second data set, including season-specific determinants, includes the unemployment rate in 1956 until 2018 and the number of arrests related to hooliganism in seasons 1987/1988 until 2013/2014. Their developments throughout time are shown in the left panel of Figure 1, where we see that the unemployment rate has mainly been rising after 1970, reaching its peak in 1983. Thereafter, it has been fluctuating a lot, but with a downward trend. The number of hooliganism related arrests has also been fluctuating a lot, reaching its peak in 2005. It also contains the average stadium attendance in the Premier League (England), and the number of cinema visits, of which the development throughout time is shown in the right panel of Figure 1. We see that it has been decreasing since the beginning of our sample, reaching its low in 1992, whereafter it has mainly been rising.

Figure 1: Unemployment rate from 1956 to 2018 and the number of hooliganism related arrests from 1987 until 2013 (Left), and the number of cinema visits in millions from 1956 until 2018 (Right). See Edel et al. (2021).



Stadium attendance is likely to be affected by hooliganism and macroeconomic factors such as unemployment. However, including stadium attendance in the Premier League and the cinema visits might not be as straightforward as including hooliganism and unemployment. The number of cinema visits is used as a proxy for leisure. If people go out to the cinema more often, it might be because they have more time to spend on leisure, which then also means that they have more time to go to a football match. Ultimately, the Premier League is the top league in Europe, so it could be that the stadium attendance in this league is an indicator for interest in football in general.

Moreover, we add, for each season, the following variables to this data set. The number of points won by a country for the UEFA coefficient list in the corresponding season, and a dummy equalling 1 in the corresponding season if the Netherlands' national team played a final at either a World Championship (WC) or a European Championship (EC) the summer before, 2 if they won the WC or EC, and 0 otherwise. Let us now elaborate on the coefficient list. The UEFA coefficient list ranks countries by their number of points. Clubs playing in the Champions League or Europa League, the top leagues throughout Europe (formerly known as Europa Cup I and Europa Cup II respectively), can earn points when they win matches or get to a certain round in these tournaments.

At the end of each season, the points obtained by clubs in the same country are added together, divided by the number of clubs, and added to the current country's number of points. The points of five years before are then deducted as it is a moving window of five years. UEFA qualifies the teams playing in these leagues based on this list, where the top teams of the top countries are qualified. The descriptive statistics of all the variables are reported in Table 3.

Table 3: Descriptive statistics of season determinants in seasons 1956/1957 until 2018/2019. Adapted from Edel (2021).

Variable	Mean	Minimum	Maximum	Number of observations
Unemployment rate	4.8	0.8	10.7	63
Hooliganism related arrests	1326	652	2401	27
Cinema visits in millions	27.9	13.7	69.1	63
Premier League attendance in thousands	29.9	18.8	38.3	63
Dummy for EC and WC	0.1	0	2	63
UEFA coefficients	6.2	0	13.6	63

We see that the unemployment rate has a mean of 4.8% and ranges from 0.8% to 10.7%. The hooliganism related arrests range from 652 to 2401, with an average of 1326. The cinema visits are 27.9 million on average and range from 13.7 million to 69.1 million, and the attendance in the Premier League range from 18.8 thousand until 38.3 thousand, with an average of 29.9 thousand. Furthermore, the dummy for the EC and WC has a mean of 0.1, and ranges from 0 to 2. Ultimately, the points obtained for the UEFA coefficient list range from 0 to 13.6, with an average of 6.2.

2.1.2 Brazil

For the analysis of Brazil, we have a panel data set with club-specific determinants. This data set contains information of home games of three clubs in the seasons 2006 until 2014. These clubs, Náutico, Santa Cruz, and Sport Recife (Sport), are the top clubs in Pernambuco, a state

in Brazil. Pernambuco has nearly 10 million inhabitants, almost 5% of all inhabitants in Brazil (a bit over 200 million). These clubs are all located in the capital of Pernambuco, Recife, which has over 1.5 million inhabitants.

These three clubs are the only clubs in Pernambuco that play in the *Série A* or *Série B* quite regularly (the top leagues in Brazil), out of the 40 clubs that play in one of these leagues each year. The data specifying which clubs played in the *Série A* or *Série B* is gathered from https://www.flashscore.com/. For the clubs that we analyse, the number of seasons they played in the *Série A*, *Série B*, and a lower division, the average stadium attendance in thousands, and the stadium capacity in thousands, is shown in Table 4.

Table 4: Club statistics in seasons 2006 until 2014. Adapted from Edel (2021). Data from Nepomuceno et al. (2017, 2022).

Club	Number of Série A	Série B	Seasons lower division	Attendance	Capacity
N´autico	5	4	0	12.4	15.6
Sport	6	3	0	20.8	33.3
Santa Cruz	1	2	6	21.9	43.6

Both Náutico and Sport did not play in a lower division than the *Série A* or *Série B* throughout this period, whereas Santa Cruz played most of its seasons in a lower division. Nonetheless, Santa Cruz has the highest average stadium attendance (21,900) and the highest capacity (43,600). Náutico has both the lowest average attendance (12,400) and capacity (15,600).

For each of these clubs, we have the average number of attendants in each season, and the capacity in these seasons. Furthermore, we have, for each season, information on whether they played in the *Série A*, *Série B*, or a lower division, and the average goal difference at home games. We use the average since we do not have the same number of observations for each season and across clubs. Hence, a big (or small) goal difference might simply mean that we have many (or a few) observations for that particular club during that year, rather than good (or bad) performance, which might lead to unreliable results. For this analysis, we do not use the number of points obtained or the ranking, since this data is not available for the *Série C* during this period. Descriptive statistics of these variables are shown in Table 5.

Table 5: Descriptive statistics of club determinants in seasons 2006 until 2014. Adapted from Edel (2021). Data from Nepomuceno et al. (2017, 2022).

Variable	Mean	Minimum	Maximum	Number of observations
Attendance	18,343	8,490	36,470	27
Capacity	30,822	15,611	43,573	27
Dummy for Série A	0.44	0	1	27
Dummy for Série B	0.33	0	1	27
Average goal difference per home game	0.01	-0.02	0.02	27

We see that the average attendance is 18,343 and ranges from 8,490 to 36,470 with stadium capacity having a mean of 30,822 and ranging from 15,611 (N´autico) to 43,573 (Santa Cruz). Furthermore, we observe that 44% of the observations is from the S´erie A and 33% from the S´erie B. Ultimately, the mean of the average goal difference per home game is 0.01, with the

worst goal difference being -0.02, and the best being 0.02.

3. Model specification

For the analysis of the Netherlands, the same approach as in Van Ours (2021) is used, a regression in two stages. The first stage is a dummy regression, defined as:

$$y_{it} = \alpha_i + \beta x_{it} + \gamma_t D_t + e_{it}, \quad \textbf{(1)}$$

Where:
y_{it} is the log of the stadium attendance;
x_{it} consists of ranking divided by hundred, goal difference divided by hundred, points divided by hundred, the log of the stadium capacity, and two dummy variables for the First Division and Second Division;
D_t are seasonal dummies, with γ_t the seasonal fixed effects;
β is a vector of parameters;
α_i are the club-specific effects;
e_{it} are the error terms.

We perform four regressions. First, we estimate this equation using all variables mentioned above using the full sample (1956/1957 - 2018/2019). Then, we will use the same sample, but omit the points divided by 100. As explained in Van Ours (2021), this variable is highly correlated with ranking and goal difference, and it is thus hard to estimate the effect of each variable separately. Ultimately, we estimate these two regressions over the period where we have data on hooliganism related arrests available (1987/1988 - 2013/2014). Note that the Second Division did not exist anymore in this last sample. Hence, this variable is omitted in the last two regressions.

After estimating this first stage, the second stage regression is performed, where the estimated season parameters γ'_t are regressed on the seasonal variables as follows:

$$\gamma'_t = \delta z_t + u_t \quad \textbf{(2)}$$

Where z_t consists of the log of the unemployment rate, a time trend, the log of the Premier League attendance, the log of the cinema visits, and the log of the number of arrests. In this stage, we add, as opposed to Van Ours (2021), our variable for the WC and EC, and the points earned for the coefficient list. Moreover, δ is a vector of parameters, and ut are the error terms. We perform six regressions. First, we omit the Premier League attendance and the cinema visits to observe the effect of the time trend without it being picked up by other variables. Then, we add the Premier League attendance, and thereafter we add the cinema visits as well. These three regressions will be estimated both for the entire period, with the hooliganism related arrests omitted, and the period for which the hooliganism related arrests are available, with the hooliganism related arrests included.

For the analysis of Brazil, we perform a similar analysis. However, due to the absence of data required for the second stage, we only perform the first stage. For this stage, xit in Equation 1 consists of the average goal difference at home divided by hundred, the log of the stadium capacity, and two dummy variables for the S´erie A and S´erie B. Besides, for Brazil, we perform only one regression, since we do not have different samples and we only have one performance indicator. Moreover, we compare the first stage regressions of the Netherlands

and Brazil, by means of the explanatory power of both the model as a whole and the explanatory variables separately.

4. Results and Discussion

Results for the first regression reports the presence of serial correlation in both FE and FD. Hence, FE is used for this regression, and, in order to obtain reliable results, for the other first stage regressions as well. The presence of serial correlation was also found for the other three regressions, so Driscoll and Kraay (1998) standard errors were used in each regression. The parameter estimates and the corresponding standard errors can be found in Table 6. Columns (1) and (2) represent the full sample, where the number of points are included in the former and omitted in the latter. Columns (3) and (4) represent the sample where data on hooliganism is available, where again the number of points are included in the former and omitted in the latter.

Table 6: Parameter estimates of the stadium attendance in the Netherlands: first stage.

	1956/1957-2018/2019		1987/1988-2013/2014	
	(1)	**(2)**	**(3)**	(4)
Dummy First Division	-0.31 (0.03)***	-0.31 (0.03)***	-0.44 (0.03)***	-0.44 (0.02)***
Dummy Second Division	-0.42 (0.04)***	-0.42 (0.04)***	-	-
Ranking/100	-0.80 (0.37)**	-0.65 (0.23)***	-0.63 (0.54)	-0.64 (0.34)*
Goal difference/100	0.29 (0.07)***	0.24 (0.07)***	0.17 (0.11)	0.18 (0.11)
Points/100	-0.13 (0.18)	-	0.00 (0.26)	-
Log stadium capacity	0.73 (0.04)***	0.73 (0.04)***	0.57 (0.03)***	0.57 (0.03)***
Within R^2	0.866	0.866	0.869	0.869
Observations	1890	1890	810	810

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

When looking at the full sample, the attendance in the First and Second Division is significantly lower than in the Eredivisie, namely 27% and 34% respectively. As mentioned before, it is hard to estimate the effects of the three performance indicators separately and we indeed see that the number of points is not significant. If we omit this variable, ranking becomes highly significant. Going up in ranking then yields 0.65% extra attendance, and one additional goal scored or one goal less conceded yields 0.24% extra attendance. Furthermore, a 1% expansion in capacity, yields 0.73% extra attendance.

When looking at the period with data on hooliganism available, we do not observe many big changes compared to the full sample. Nonetheless, when the number of points are included, none of the performance indicators are significant, and when the number of points are omitted, only ranking is significant, and only at the 10% significance level. We thus conclude that performance indicators are not that determining in this period, whereas Van Ours (2021), inferred that they are.

When performing the Breusch-Godfrey tests for the second stage, we found the presence of serial correlation in each regression in the full sample and the regression with all variables in the sample where the arrests are available. Hence, Newey-West standard errors were used for each of those regressions. For the remaining two regressions, no significant heteroskedasticity was found and we thus use ordinary OLS standard errors. The parameter estimates and the corresponding standard errors can be found in Table 7. Columns (1), (2) and (3) represent the

full sample, where both the attendance in the Premier League and the cinema visits are omitted in column (1), the attendance in the Premier League is included in column (2), and both the attendance in the Premier League and the cinema visits are included in column (3). Columns (4), (5), and (6) represent the sample where data on hooliganism is available, where their mutual differences are the same as columns (1), (2), and (3).

Table 7: Parameter estimates of the stadium attendance in the Netherlands: second stage.

	1956/1957-2018/2019			1987/1988-2013/2014		
	(1)	(2)	(3)	(4)	(5)	(6)
Log unemployment rate	-0.26 (0.03)***	-0.11 (0.03)***	-0.08 (0.03)***	-0.15 (0.03)***	-0.10 (0.03)***	-0.10 (0.03)***
Time/10	0.14 (0.01)***	0.08 (0.02)***	0.09 (0.02)***	0.28 (0.01)***	0.19 (0.03)***	0.19 (0.03)***
Log Premier League	-	0.57 (0.10)***	0.42 (0.09)***	-	0.36 (0.10)***	0.36 (0.11)***
Log cinema visits	-	-	0.15 (0.04)***	-	-	-0.00 (0.09)
UEFA coefficients	-0.00 (0.00)	-0.01 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Dummy for EC and WC	-0.04 (0.03)	-0.03 (0.02)	-0.03 (0.01)**	0.00 (0.02)	-0.00 (0.01)	-0.00 (0.01)
Log arrests R^2	- 0.821	- 0.921	- 0.944	0.08 (0.02)*** 0.986	0.03 (0.02) 0.991	0.03 (0.03) 0.991
Observations	63	63	63	27	27	27

*** $p < 0.01$, ** $p < 0.05$

For both periods, the unemployment rate has a negative significant effect on attendance, the attendance in the Premier League has a significant positive effect, and the attendance is significantly increasing over time. The magnitude does not differ a lot between the periods. Furthermore, where the cinema visits are significant in the first sample, with an increase in attendance of 0.15% when cinema visits increase with 1%, it lost this significance during the second sample. The UEFA coefficients are not significant in any of the regressions, and the dummy for the EC and WC is only significant once. Also, compared to Van Ours (2021), the other estimates are hardly affected and the R^2 barely increases, even in the regression where the dummy for the EC and WC is significant.

Hence, these variables do not have any explanatory power when explaining stadium attendance. A possible explanation for the UEFA coefficients might be that the teams earning points are thus playing well enough in the Champions League or Europa League to obtain points. Hence, they will also play well in the national league, which thus leads to better performance indicators, and so there is no separate effect of the UEFA coefficients. For the EC and WC, a possible explanation could be that club competitions are quite separated from national competitions. Hence, a national team playing well does not get people excited to watch club competitions. Ultimately, hooliganism does not seem to have a significant effect, and where Van Ours (2021) finds that the effect is significant in column five's regression, we find that it is not.

After performing the Wooldridge tests, we found did not find the presence of serial correlation or heteroskedasticity in FE. Hence, FE is used, with OLS standard errors. The parameter estimates and the corresponding standard errors are the following. The *Série A* has an estimate of -0.26, with a standard error of 0.24. The *Série B* has an estimate of -0.29, with a standard error of 0.23. Both are not significant. Furthermore, the average goal difference per home game divided by hundred has an estimate of 971.0311, with a standard error equal to 996.09, which is also not significant. The within R^2 equals 0.667. Since the log of the stadium capacity is a fixed number, this variable was omitted since this becomes 0 for each observation when

subtracting the club average because of the use of FE. However, to still get an estimate for the stadium capacity, we estimated the club-specific effects αi in Equation 1, and regressed them on the log of the stadium capacity. This gave an estimate of 0.39, with a standard error equal to 0.26, which is not significant. However, no conclusions can be drawn from this since this regression only had three observations.

Moreover, it is extremely hard to properly estimate regression coefficients with 12 regressors (three explanatory variables and nine-year dummies) with only 27 observations. Hence, these results must be carefully interpreted. As an attempt to better estimate the coefficients, we ran six additional regressions, three with each possible combination of two variables, and three with regressor separately, each without year dummies. Here we found that the goal difference was insignificant in each regression, but that the *Série A* and *Série B* were significant on a 10% significant level in one of the regressions. This might indicate that goal difference indeed does not explain stadium attendance, and that division needs to be further investigated as this might have explanatory power. However, no real conclusions can be drawn due to the limited number of observations.

When comparing the Netherlands and Brazil based on the R^2, we see that the model of the Netherlands has a larger explanatory power. However, this might be because of the performance indicators points and ranking, which we do not have for Brazil. When comparing the significance of the variables, we see that division is highly significant in the Netherlands, but the significance in Brazil is not so clear. Goal difference is significant in one sample in the Netherlands, but not in the other, and in Brazil it is not significant. Stadium capacity is highly significant in the Netherlands, but not in Brazil. This might suggest that attendance is highly performance-dependent in the Netherlands, but that people in the Brazilian football culture support their club no matter how they play. Nonetheless, we again cannot draw real conclusions due to the limited number of observations.

5. Conclusion

This paper has scrutinised the effects of stadium capacity and performance indicators on football stadium attendance in both the Netherlands and in Brazil. Besides, it has shown the similarities and differences between these country with respect to this issue. Furthermore, this paper has examined the effects of the unemployment rate, hooliganism, general interest in football, and time spent on leisure on football stadium attendance in the Netherlands.

For the Netherlands, we found that stadium capacity affects stadium attendance significantly positively. Also, the lower the league a club plays in, the less people attend the stadium. Besides, other performance indicators, such as ranking, goal difference, and number of points obtained during a season, seem to positively affect stadium attendance. However, these effects are extremely hard to measure due to their high correlation.

Furthermore, the unemployment rate has a significantly negative effect, and hooliganism does not seem to affect stadium attendance. The attendance in the Premier League has a significant positive effect on stadium attendance in the Netherlands, which might indicate that stadium attendance rises as general interest in football rises. Furthermore, it is unclear whether the time spent on leisure significantly affects stadium attendance.

For Brazil, we found, as opposed to the Netherlands, that stadium capacity does not have a significant effect on stadium attendance. Furthermore, we found that the goal difference does

not affect attendance significantly. Moreover, we did not find any unambiguous results regarding the division. Ultimately, the explanatory power of the model seems to be lower in Brazil than in the Netherlands. However, these results should be carefully interpreted, and no real conclusions can be drawn because we have a limited number of observations. Also, we were forced to omit certain variables due to missing data, and only a specific set of clubs was analyzed. Further research could focus more on South America once more data is available.

Conflict of Interest Declaration

The author has no conflicts of interest to declare.

References

1. Borland, J. and MacDonald, R. (2003). Demand for sport. Oxford Review of Economic Policy, 19(4):478–502.
2. Buraimo, B., Tena, J., and de la Piedra, J. D. (2018). Attendance demand in a developing foot- ball market: The case of the Peruvian first division. European Sport Management Quarterly, 18(5):671–686.
3. Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. Review of Economics and Statistics, 80(4):549–560.
4. Edel, M. (2021). What determines football stadium attendance? An extensive analysis of Brazil and Europe, with the Netherlands in particular. Bachelor Thesis Econometrics and Operations Research. Erasmus University Rotterdam.
5. Ferreira, M. and Bravo, G. (2007). A multilevel model analysis of professional soccer attendance in Chile 1990-2002. International Journal of Sports Marketing and Sponsorship, 8(3):49–66.
6. García, J. and Rodríguez, P. (2002). The determinants of football match attendance revisited: Empirical evidence from the Spanish football league. Journal of Sports Economics, 3(1):18–38.
7. Madalozzo, R. (2008). A model of attendance demand at the Brazilian football league. Insper Working Paper, WPE: 113/2008.
8. Martins, A. M. and Cr´o, S. (2018). The demand for football in Portugal: New insights on outcome uncertainty. Journal of Sports Economics, 19(4):473–497.
9. Nepomuceno, T. C. C., de Moura, J. A., e Silva, L. C., & Costa, A. P. C. S. (2017). Alcohol and violent behavior among football spectators: An empirical assessment of Brazilian's criminalization. International journal of law, crime and justice, 51, 34-44.
10. Nepomuceno, T. C. C., de Carvalho, V. D. H., Silva, L. C. E., de Moura, J. A., & Costa, A. P. C. S. (2022). Exploring the Bedouin syndrome in the football fan culture: Addressing the hooliganism phenomena through networks of violent behavior. International journal of environmental research and public health, 19(15), 9711.
11. Pawlowski, T. and Anders, C. (2012). Stadium attendance in German professional football– The (un)importance of uncertainty of outcome reconsidered. Applied Economics Letters, 19(16):1553–1556.
12. Pawlowski, T. and Nalbantis, G. (2015). Competition format, championship uncertainty and stadium attendance in European football–A small league perspective. Applied Economics, 47(38):4128–4139.

13. Schreyer, D. and Ansari, P. (2021). Stadium attendance demand research: A scoping review. Journal of Sports Economics, forthcoming.
14. Van Ours, J. C. (2021). Common international trends in football stadium attendance. PLOS ONE, 16(3):e0247761.