The Demand for Soccer[1]

Jacob Goeree[2] and Mark Minkman[3]

January 1996

Abstract: In this paper we investigate the determinants of the number of attendants of a soccer match. We propose a censored regression model to account for the fact that the grounds of the clubs have limited capacity. This enables us to determine the real, or potential, demand for soccer. The independent variables of our model consist of two types: club specific and match specific variables. The former include current form, current league position and club dummies (both for the home and away team). The match specific variables include geographical distance, a derby and top-match dummy and a dummy to account for the day the match is played. Our results stress the importance of club specific variables that do not change within a season. We find that the top-clubs, Ajax, Feyenoord and PSV, generate in their away matches a gross total of four and a half million guilders per season for the other clubs. We also find that FC Groningen looses almost a quarter of a million guilders per season due to its isolated geographical position.

---

[1]We would like to thank the participants of the "Bijeenkomst voor Sport en Statistiek" and an anonymous referee for their comments. We also would like to thank Joost Schraag helping us collect (part of) the data.

[2]Department of Actuarial Science & Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

[3]Foundation for Economic Research (SEO), University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

1. Introduction

The financial significance of soccer in the Netherlands is growing. Due to the present outstanding success of the Dutch, (former) European and World Champion Ajax, the last year has shown a tremendous increase in attention for Dutch soccer. The value of the product 'Ajax' is estimated to be one hundred million guilders per year (see Wijman (1995)). Furthermore, nine of the ten best viewed television programmes were live broadcasted soccer matches, indicating the importance of soccer for the (public) broadcasters. Therefore, it is quite surprising that there has been little research that concerns the economic and financial implications of Dutch soccer. In this note we investigate what determines the number of attendants of a soccer match in the Dutch Premier League.

2. Data and model specification

Our data consisted of 701 matches played in the Dutch Premier League (also known as the PTT-telecompetition) in the seasons 92/93, 93/94 and 94/95. They were obtained from various newspapers and soccer magazines. The independent variables that determine the number of attendants at a match can roughly be divided into two categories: club specific and match specific variables. The variables in the former category come in pairs, one for the home and one for the away team. The reason for taking both home and away variables is intuitively clear: first, there are the fans of the away team that attend, and secondly, there are clubs that attract extra 'local,' or home, spectators in away matches. Both reasons indicate that the characteristics of the away team influences the number of attendants of the match. The club specific variables are:

- *Current form* (f^h, f^a)

  These variables are included to capture the effect of a team that is in a 'winning mood.' The variables f_i^h, f_i^a are based on the number of points obtained in the five matches prior to match i. They are defined as follows: $f^h = 1$ if the home team has gained 0, 1 or 2 points; $f^h = 2$ if it has gained 3 or 4 points; $f^h = 3$ if it has gained 5 or 6 points, etc. A similar definition holds for f^a.

- *Current league position* (p^h, p^a)

  Being at the top (bottom) of the league table may have a postive (negative) effect on the number of attendants, which is what this variable seeks to capture. The variables p_i^h, p_i^a are determined immediately prior to match i, and take values between 1 and 18, which is the number of teams in the Premier League.

- *UEFA/Championsleague significance* (u^h, u^a)

  Following the approach of Jennet (1984), we include variables that measure the (theoretical) changes of reaching a league position at the end of the season that entitels a team to play in the European competition (either Championsleague or UEFA competition). They are based on an index that runs from 1 to 2; for teams that remain in contention[1] for an European ticket, the index increases to 2 as the season reaches its end. For teams that drop out the index is set to 1.

- *Club dummies*

  These dummies are included to summarize all the effects that do not change within a season. For instance, a team that is temporarily out of form may still attract a large number of spectators due to its record of past successes. So one facet that the club dummies represent is the long term form, or status, of a club. But the dummies also reflect other club specific aspects that are constant during the season such as the image of teams, the attractiveness of the home team's stadium, the loyalty of the fans, etc.

The club specific variables take into account the characteristics of the home and away team seperately. However, they are incapable of capturing the effects that depend on both clubs, such as

- *Geographical distance* (Δ)

  One would assume that a great geographical distance between the home cities of the home and away team has a negative effect on attendance. This may be due either to long travel times or to the fact that away teams that come from far are less appealing to the home fans.

- *Derby and top-match* (d, t)

  The derby dummy is 1 for the following matches: NAC – RKC, NEC – Vitesse, NAC – Willem II, Willem II – RKC, Feyenoord – Sparta, Roda JC – MVV, Heerenveen – Groningen, and Go Ahead – Twente (and of course for the same matches with the home and away team interchanged), and zero otherwise. It measures the positive effect of local interest and rivalry. The top-match dummy is 1 when both the home and away club are from the set {Ajax, Feyenoord, PSV} (which are widely accepted to be the three top-teams in the Netherlands), and zero otherwise.

- *Day of match* (m)

  Although the majority of the matches are played on sunday, some clubs play their home games on a different day (*e.g.* PSV, whose home matches are scheduled on saturday). In addition there are matches that are (re)scheduled to working days. The dummy variable m, which is 1 if the day of play is sunday and 0 otherwise, is included to measure the effect of this.[2]

[1] This is the case when it is (mathematically) still possible for a team to reach the number of points which is ultimately needed to obtain an European ticket.

[2] Actually, we started with 4 dummies, one for working days and one for fridays, saturdays and sundays. However, it turns out that only on sundays the number of attendants differs significantly from other days.

Finally, we included a season dummy s to capture the empiral fact that attendance grows each year.[3]

The above variables constitute the following model:

$$y_{ij} = \beta_j' x_{ij} + \gamma' z_{ij} + \delta s + \epsilon_{ij}, \tag{2.1}$$

where y_{ij} denotes the number of attendants at match i of home team j, and the x_{ij} (z_{ij}) are club (match) specific independent variables discussed above, whose effects are measured by the β's and γ's respectively. Furthermore, the ϵ_{ij} are assumed to be normally distributed with zero mean and variance σ_j^2.

We cannot employ the ordinary least squares (OLS) technique to estimate the above model, since we have to take into account the fact that the attendance of some clubs varies more than that of others.[4] This is why the above model allows for an heteroscedastic disturbance, which depends on the club label j. This heteroscedastic model can in principle be estimated using the generalized least squares (GLS) technique. However, due to the limited capacity of the grounds the GLS technique would underestimate the parameters. This problem is most prominent in home games of Ajax, Heerenveen, and PSV, who play to a full house in 85%, 55%, and 30%, of their home matches respectively. The solution of this problem[5] is well known: one should use a censored regression model (*e.g.* , Maddala (1983)). This solves the finite capacity problem for all but one club, namely Ajax. Their stadium "De Meer" is (almost) always filled to capacity so that the dependent variable shows virtually no variation. Fortunately, Ajax has played some of its matches in the Olympic stadium, which has a much larger capacity. These matches have not always been fully booked[6], yielding the necessary variation of the dependent variable y_{ij} for the model to be estimable.

Since we use a censored regression model the β's, γ's and δ reflect the effect of the independent variable on the *true* demand for soccer, rather than on the number of spectators which may be (much) smaller due to the capacity constraints of the grounds.

3. Results and discussion

The model was estimated using a maximum likelihood procedure. Table 1 gives the parameter estimates. The estimates of the club specific standard deviations σ_j indicate the presence of

[3]This set of variables partially overlaps with that of Dobson and Goddard (1992) who analyzed the demand for standing and seated viewing accommodation in the English Premier League. Their approach differs from that of ours in that they do not use a censored model to correct for the finite capacity of the grounds, see below.

[4]In section 3 this is shown more rigorously using the likelihood ratio test.

[5]An additional problem is that for reasons of safety the authorities may demand a maximum attendance that is less than the capacity of the grounds. We have accounted for this by adjusting the capacity in such cases.

[6]Although in matches against Feyenoord and PSV it has, which means that for Ajax we are in fact dealing with a doubly censored model.

heteroscedasticity. To verify this we also estimated a (censored) homoscedastic model. This results in a standard deviation of $\sigma = 2068$. Comparing the likelihoods of the two models yields a likelihood ratio test statistic (χ^2 distributed with 17 degrees of freedom) that equals 329, which is highly significant.

We also compared the estimates with the GLS-estimates. The error one thus makes is best illustrated by the estimate of the Ajax home dummy: the GLS-estimate equals 20235, whereas the correct estimate is 30213 (see Table 1). The hypothesis that the limited capacity of the grounds has no effect on the estimates, may also be rejected by comparing the likelihoods of the GLS and the censored model. The likelihood ratio test statistic is again highly significant.

Note that all the estimates have the sign that is to be expected, except maybe for the RKC away dummy and perhaps the sign of the day of match dummy. The latter turns out to be negative, probably due to the fact that on sundays a particular match has to compete with many other matches, and/or other (sports)events. In Table 1 we have only listed the parameters that were significantly different from zero. In particular there were no significant effects of the club specific variables of the away team that change within the season, in fact only the club dummy (which remains constant during the season) of some away teams mattered. The main difference between our results and those of Dobson and Goddard (1992) is that in our case short term factors, such as current form and current league position, do not play an important role. The fact that they are important in the English Premier League means that the number of attendants varies more in England: an English team that is temporarily out of form (in form) is punished (rewarded) more than a Dutch team.

It is clear from the above estimates that the demand for soccer is largely determined by the status of the home and away team. (This is nicely illustrated by the present competition (95/96) in which Sparta is exceeding its long term form, but hardly attracts any extra spectators.) One has to bear in mind that these numbers do not simply reflect the mean attendance of the home team; in fact they represent the mean of the *true* demand for soccer, which can differ due to the limited capacities of the grounds. Moreover, although the influence of the other variables is small for one match, their cumulative effect (over a season) can become quite substantial. This point will be elaborated upon below.

One of the more striking results is the lack of capacity that Ajax has to deal with. This season (95/96) Ajax still performs in "De Meer" that has a capacity of 19500, occasionally escaping to the larger Olympic stadium with a capacity of 45000. Our analysis implies that Ajax would have earned an extra four million guilders per season (taking the price of a ticket to be twenty guilders) if they had played their matches in the Olympic stadium. To a lesser degree a similar result holds for PSV and Heerenveen.

There is a clear significant effect of the away dummies of the top-clubs Ajax, Feyenoord and PSV. These clubs attract seven, five and three thousand extra spectators per game respectively, when they play an away match, thus generating a gross total of four and a half million for the other clubs. The top-clubs thus form a strong financial driving force for the Dutch soccer competition. To a lesser degree Heerenveen (500) and NAC (1000) also attract extra spectators in their away matches. Surprisingly, RKC *lowers* the attendance by one thousand when it plays an away match.

The other estimated parameters such as derby, top-match, and season, are all significant and have the expected sign. Of particular interest is the influence of the geographical distance Δ. Using the estimated value for the parameter corresponding to Δ one may conduct the following Gedanken experiment: what would be the benefit for FC Groningen if it would play in Utrecht instead of Groningen ? This benefit is readily computed to be almost a quarter of a million (again taking the ticket price to be twenty guilders) per season. This indicates that FC Groningen might be better off in a more regional competition, for instance in the German Bundesliga Nord.

Our analysis may be extended in the following directions. One may want to discriminate between card holders and buyers of single tickets (see Dobson and Goddard (1992)). It would be of great interest to determine the consequences of price changes, *i.e.* determine the price elasticity of the demand for soccer. A totally different point is the influence of wheather conditions, a variable we left out due to lack of data. Finally, it would be very interesting to use the above results (or extensions thereof) to optimize financial benefits for all clubs involved, for instance by accounting for the effects of the variables that change within a season when drawing up the fixture, or by regionalizing the competition.

References

J. Borland and J. Lye, *"Attendance at Australian rules football: a panel study,"* Applied Economics **24** (1992), 1053–1058.

B. Burkitt and S. Cameron, *"Impact of league restructuring on team sport attendances: the case of rugby league,"* Applied Economics **24** (1992), 265–271.

Dobson, S. M. and J. A. Goddard, *"The demand for standing and seated viewing accommodation in the English Football League,"* Applied Economics **24** (1992), 1155–1163.

Jennet, N., *Attendances, uncertainty of outcome and policy in the Scottish Football League,* Scottish Journal of Political Economy **31** (1984), 176 – 198.

Maddala, G. S., *Limited-dependent and qualitative variables in Econometrics,* Econometric Society Monographs, Cambridge University Press, New York (1983).

Wijman, E., *"Ajax als supermerk: de merchandising,"* Adformatie **51/52** (1995), 70 – 71.

club specific parameters	*estimate*	*standard error*	*t-value*
current form home f^h	205	63.5	3.2
current position home p^h	−77.5	22.0	−3.5
UEFA significance home u^h	2207	551	4.0
distance Δ	−3.24	1.02	−3.2
home dummies			
Ajax	30213	2394	12.6
Feyenoord	21658	1156	18.7
Go Ahead	4528	759	6.0
Groningen	10599	789	13.4
Heerenveen	12027	807	14.9
MVV	4305	757	5.7
NAC	8473	794	10.7
NEC	3646	1000	3.7
PSV	23224	935	24.8
Roda JC	4322	824	5.2
Sparta	2857	776	3.7
FC Twente	4778	780	6.1
FC Utrecht	4803	780	6.2
Vitesse	4063	743	5.5
Volendam	2088	765	2.7
Willem II	5050	732	6.9
away dummies			
Ajax	6866	305	22.5
Feyenoord	4882	265	18.4
Heerenveen	597	270	2.2
NAC	1036	264	3.9
PSV	3018	263	11.5
RKC	−798	240	−3.3
match specific parameters			
derby d	1452	250	5.8
top-match t	5657	1952	2.9
day of match m	−310	145	−2.1
season 92/93	−483	166	−2.9
season 93/94	−461	142	−3.2
standard deviations σ_j			
Ajax	4597	1315	3.5
Feyenoord	5229	579	9.0
Go Ahead	1104	125	8.8
Groningen	1230	144	8.6
Heerenveen	1523	300	5.1
MVV	1285	146	8.8
NAC	1519	242	5.3
NEC	2786	485	5.7
PSV	2947	463	6.4
RKC	951	123	7.7
Roda JC	1619	210	7.7
Sparta	1732	197	8.8
SVV/Dordrecht'90	780	152	5.1
FC Twente	1789	195	8.8
FC Utrecht	1966	239	8.2
Vitesse	996	118	8.4
Volendam	1115	140	8.0
Willem II	1259	148	8.5