

The determinants of football match attendance revisited: Empirical evidence from the Spanish football league.

Jaume García* and Plácido Rodríguez**

An attendance equation is estimated using data on individual games played in the Spanish First Division Football League. The specification includes as explanatory factors: economic variables, quality, uncertainty and opportunity costs. We concentrate the analysis on some specification issues such as controlling the effect of unobservables given the panel data structure of the data set, the type of functional form and the potential endogeneity of prices. We obtain the expected effects on attendance for all the variables. The estimated price elasticities are smaller than one in absolute value as usually occurs in this literature but are sensitive to the specification issues.

Keywords: Attendance, price elasticity, panel data

JEL classification: C23, L83

* Corresponding author:
Department of Economics and Business
Universitat Pompeu Fabra
Ramon Trias Fargas 24-26
08005 Barcelona (Spain)
e-mail: jaume.garcia@econ.upf.es

** Department of Economics
Universidad de Oviedo
Avenida del Cristo, s/n
33071 Oviedo (Spain)
e-mail: placido@econo.uniovi.es

ACKNOWLEDGEMENTS

We are very grateful to César Rodríguez and Joaquín Lorences and the participants in a seminar at the University of Oviedo for their comments on earlier stages of this work. Special thanks are also due to the *Liga Nacional de Fútbol Profesional* for providing the data used in this paper. The first author wishes to express his gratitude for financial support from DGESIC grant No. PB98-1058-C03-01. The second author wishes to thank FUNCAS for its financial support. The usual disclaimer applies.

1. Introduction

The analysis of the determinants of attendance at professional team sports events is one of the topics which has received most attention in the empirical literature of the socio-economics of sports.¹ The usual approach is the estimation of a demand equation, which is either linear or can be linearized, including as explanatory factors the usual economic variables (prices and income) and the sectoral variables which try to capture the heterogeneity of this type of good (a football match). This has been done using different types of data sets depending on data availability and the objectives of the study².

Most of the studies do not pay too much attention either to econometric specification issues or to the economic implications of the results. With respect to the latter point, not all the papers include prices as an explanatory factor and not many take into account the interpretation of the price elasticities obtained and try to rationalize it. Papers by Heilmann and Wendling (1976), Ferguson *et al.* (1991), Salant (1992), Marburger (1997) and Boyd and Boyd (1998) are among those which try, in the context of a profit maximizing behaviour for the professional teams, as assumed in El Hodiri and Quirk (1971), to give a theoretical explanation for the usual empirical finding of a price elasticity of less than one. This finding can also be explained in a context where teams have objective functions other than profits, as proposed by Sloane (1971).

In this paper we try to bring new evidence from European football³ to bear on this empirical issue making use, for the first time in this literature, of a data set corresponding to the Spanish Football League, one of the most important and highly

regarded in Europe. We use a complete data set with observations of both economic and sectoral variables for all the matches played in the Spanish First Division League during the seasons 1992-95 to 1995-96, which allows us to specify an attendance equation that is more detailed in terms of the explanatory variables than in previous studies. We use the panel data structure to control for some unobservables in order to estimate price elasticity consistently, while also taking into account the possible endogeneity of this variable. We analyze the incidence of some specification issues, in particular the functional form, in the estimated elasticities. We also evaluate the importance of the different groups of variables in explaining attendance.

The paper is organized as follows. In Section 2 we present the specification of the empirical model. In Section 3, we discuss the estimation results. The paper ends with a summary of the main conclusions.

2. Model specification

Model and variables

We specify and estimate a fairly standard demand equation distinguishing, among the explanatory factors which have an effect on attendance, the following groups of variables: economic variables, variables proxying the expected quality of the match, those measuring the uncertainty of the result, and those capturing the opportunity cost of attending a match.

The basic data set comes from the information received by the *Liga Nacional de Fútbol Profesional* from each club about the number of people attending each match and the prices charged. The sources and the descriptive statistics of the variables are presented in Table A.1. in the Appendix.

The endogenous variable is the (log)number of tickets sold for a match (attendance), not including those for children and season tickets⁴. Among the economic variables we include: prices, measured by the price of the cheapest ticket⁵ deflated by the CPI; real income per capita in the province of the team playing at home; and the population in the province of the home team, which is distributed, when there are two or more teams in a province, according to the number of season ticket holders corresponding to each team. We expect a negative effect for prices and a positive one for both income (i.e. a normal good) and population.

The expected quality of the match can be measured by what we call *ex ante* quality, i.e. the quality of both teams at the beginning of the season, independent of performance previous to the match; and by those variables proxying the most recent performance of both teams (current quality). In the first group we include the budgets (in real terms) of both teams, because they depend, among other things, on the salaries of the players, which should proxy their productivity⁶; the number of players who have played for their national team (internationals); two dummies for those matches where the away team is either Barcelona or Real Madrid, historically the two most important teams in Spain; a dummy for those games of special interest because of historical or regional rivalry; and, finally, a dummy indicating whether season ticket holders have to pay to attend the

match (the Club Day match), a usual practice of most Spanish clubs. This is an indicator of the club's expectation about the quality (or interest) of the game.

Among the variables capturing the recent performance of both teams we include⁷: the number of home team wins in the last three games; the result (as the difference between goals scored for and against) of the most recent game played by the home team; the home team's current position in the league; the number of goals scored in the last match at home by the home team; a dummy for the away team not having lost a game out of the last four; and two dummies for the home team using the value one if the latter has no chance of winning the championship or of leaving the relegation zone. We expect all variables increasing quality to have a positive effect on attendance.

With respect to uncertainty we distinguish, as it is done in Kuypers (1996), between match and seasonal uncertainty⁸. We use as measures of match uncertainty a quadratic form of the difference between the league positions of the home and the away teams previous to the match and a dummy equal to one if the home team is between three positions ahead and five positions behind the away team. To measure seasonal uncertainty with respect to winning the championship, we have chosen the second indicator proposed by Kuypers (1996), which is the product of the number of games left before the championship is decided and the number of points the team trails behind the leader, being equal to zero when there is no possibility of the team's winning the championship. Uncertainty increases attendance, and in the particular case of the measure of seasonal uncertainty we expect a negative sign since the higher the uncertainty, the smaller the value of the indicator we have defined.

Finally, we include a set of variables which capture the opportunity cost of attending a football match. We model the effect of weather conditions with dummy variables which correspond to the following situations: no rain, high temperature; no rain, low temperature; and rainy days, which is the omitted dummy. We expect the better the weather is, the higher the attendance will be. The second factor has to do with the game's being televised. Since in Spain football games are televised by both public and private channels, in the latter case only for subscribers, we define two different dummies depending on which channel is broadcasting the match. The omitted group corresponds to matches not televised. We expect that televising games will reduce attendance, especially, if the match is televised by a public channel.

The day of the match also has to play a role in determining attendance. Specifically, if a match is played on a week-day, rather than on the weekend, attendance should decrease⁹. We model this variable by means of a dummy.

Finally, we include the distance between the towns of both teams as a way of capturing the demand which comes from away team supporters. We give special consideration to the case of Tenerife, located on the Canary Islands, by including two dummies: one for Tenerife playing at home and the second one for Tenerife playing away. We expect distance to have a negative effect on attendance.

Econometric specification

Given the panel data structure of our data set, the variables included in the demand equation can have different sources of variability. There is time variation because the

observations correspond to games played during four seasons (1992-93 through 1995-96)¹⁰ from September through May/June. On the other hand, there is variation depending on the teams which correspond to each observation (home and away teams). Consequently, we use three subindices to identify each observation. They refer to the season (t), the home team (i) and the away team (j), with four types of explanatory variables in terms of the sources of variation. We have a set of variables, included in the vector X_{ijt} , which vary in all three dimensions as also happens with the endogenous variable Y_{ijt} . Variables which refer to prices, the difference between the two teams' current league positions, weather conditions, a match's being televised, a match's not being played on the weekend, and a Club Day match are included in this group. The second group of variables, included in the vector Z_{it} , is made up of those variables which vary depending on the season and on the home team. Variables such as income, population, the home team budget, the home team's result in the most recent game played, the number of wins by the home team in the last three games, Kuypers' measure of uncertainty for the home team with respect to the championship, and those variables referring to the home team's not having any chance of winning the championship or leaving the relegation zone, belong to this second group. The third one includes variables, such as the away team's budget, the number of internationals and the away team's having not lost a match out of the last four, which show variation depending on the away team and the season. They are included in the vector V_{jt} . Finally, in the fourth group there are the variables, included in vector W_{ij} , which simultaneously depend on both teams playing the match. Variables capturing historical and local rivalries and the distance between the cities of the two teams are in this group. Note that there is, in fact a fifth dimension of variation because some of the variables in Z_{it} and V_{jt} show variation

across different games of a particular season whereas others are constant through the season.

On the other hand, we can control for the presence of unobservables which can also have different sources of variation and whose omission in the specification and estimation of the model can cause inconsistency of estimates, to the extent they are correlated with the regressors. We consider home team effects (α_i), away team effects (η_j) and season effects (τ_t) apart from the usual disturbance term (u_{ijt}).

Consequently, the general specification of the model has the following form:

$$Y_{ijt} = X_{ijt}'\beta + Z_{it}'\gamma + V_{jt}'\delta + W_{ij}'\theta + \alpha_i + \eta_j + \tau_t + u_{ijt} \quad (1)$$

where β , γ , δ and θ are the vectors of parameters. In the empirical model we do not take into account the unobserved away team effects because we consider that the explanatory variables related to the visitor capture the basic effects, specifically, those dummies which refer to a particular away team (Barcelona, Real Madrid and Tenerife)¹¹.

Equation (1) is estimated by OLS which for consistency require the unobserved effects to be uncorrelated with the explanatory variables. We also take the usual approach to controlling for these effects by including dummy variables for the home team and the season effects. Finally, we also transform the model by taking a special type of “differences”. Specifically, we subtract from each variable the value of the observation corresponding to the previous match played at home by the home team in that season¹². This transformation eliminates both the home team effect and the season effect but also

eliminates those variables which show only this kind of variation (Z_{it}). The latter elimination does not happen when using the within group transformation.

On the other hand, we do not estimate the model by GLS (Balestra and Nerlove, 1966) because, contrary to what happens with other panel data sets, the individual dimensions (the number of teams) are, in some sense, small. Consequently, when we specify a model with more explanatory variables than teams in the data set, as happens in our empirical model, we cannot obtain an estimate of the variance of the home team effects. Nor can we use Hausman's test for the null of home team effects' being uncorrelated with the explanatory variables. Note too that we can compare the specification in (1) with that of those models estimated using data (averages) on seasons rather than data on games (e.g. Jones *et al.*, 2000). This amounts to transforming our model by averaging variables for each home team in each season, which will not allow us to identify the coefficients γ in equation (1).

Finally, the price variable has potential problems of endogeneity as is usual in demand analysis. This could explain, apart from data availability, why most of the empirical studies do not include it, but, in fact, estimate a kind of reduced form model. Since, in analyzing Spanish football clubs' optimization behavior, we are interested in estimating price elasticity consistently, we estimate the model by IV instrumenting the price variable using the value predicted from a reduced form equation. We include in this (log) price equation as explanatory variables all the variables, apart from prices, which appear in the attendance equation. We also include variables that refer to performance in the previous season (the final league positions of both home and away teams and dummies for either team's being in the second division in the previous season) and the

number of tickets which can be sold (capacity). These variables allow us to identify the demand equation.

3. Empirical results

In this section we present the main results of the specification and estimation of an attendance equation for the Spanish football (*General results*) with special attention to the estimates of the price elasticities and their interpretation (*Price elasticities*) and to the *contribution of each group of variables to explaining attendance*.

General results

We estimated different versions of equation (1), in which the endogenous variable (attendance) is always in logs, using 1580 observations. The results are reported in Table 1. First, we consider the equation linear in price and income variables, both in logs, (column 1) and we compare it with a more general specification based on a cubic polynomial for these two variables (column 2). We also control for the potential correlation between home team and season effects and the regressors by including the corresponding dummies (column 3), our preferred specification. We compare the last specification against a non-nested one where the price and income variables are not in logs and have a cubic profile (column 4). We also estimate the preferred specification by applying OLS to the model transformed by taking differences as defined above (column 5). Finally, we estimate the preferred specification by instrumenting the price variable (column 6). The results of these estimations are reported in Table 1.

(TABLE 1)

Results presented in column (2) show that the model with cubic polynomials both in (log) prices and (log) income better fits the attendance equation than does the standard model linear in logs (column 1)¹³. All the coefficients in both models have the expected sign and almost all of them are significant at a 5% level. Additionally, the estimates do not differ very much between the two specifications except for the price and income variables.

The income elasticity (η_Y) obtained from the model in column (2) is

$$\eta_Y = 1980.4 - 425 * \log(\text{INCOME}) + 22.8 * [\log(\text{INCOME})]^2$$

which is in the interval [-0.137, 5.373] evaluated at the values of the observations of our sample. This means that although elasticity has a quadratic form, the relevant values for our sample are basically positive, i.e., attendance is a normal good. In fact, in the simplest specification of column (1), the estimated constant income elasticity (the coefficient of (log) income) is positive and significant.

All variables proxying *ex ante* quality have the expected positive sign. The coefficients of the budget variables are very similar (statistically the same), with the effect on attendance of a match's being played by "rival" teams found to be more important than the fact that either Barcelona or Real Madrid are the away team. This could be explained by the fact that its effect is captured through the budget variables. In fact, when defining the budget variables in logs, the coefficients of the dummies corresponding to Real Madrid and Barcelona become significant but the fit of the model is worse than that of the corresponding model reported in Table 1. Additionally, the

number of internationals playing for the away team has also a positive effect on attendance¹⁴.

Scoring an additional goal either in the last match or in the last match at home and having an extra victory in the last three games have a similar effect on attendance as shown in column (2). On the other hand, the fact that the away team is unbeaten in the last four games increases also attendance. The other three variables included in this group proxying current quality (current interest) of the match produce some problems when we interpret their effects because they may also proxy the uncertainty of the match and of the season. In particular, those games in which the home team has no chance of either winning the championship or leaving the relegation zone have *ceteris paribus* smaller attendance¹⁵.

Uncertainty variables also have the expected positive effect on attendance; the sign of the variable defined for the uncertainty of who will win the championship is negative given the way in which the variable is defined. On the other hand, the variables defined in terms of the league positions of the teams also produce problems in interpretation. The results indicate that the closer the positions in the league, the greater the uncertainty and, consequently, the greater the attendance, but at the same time the better the league position of the away team with respect to the home team, the greater the attendance¹⁶.

Poor weather conditions discourage people from attending football matches since they are played outdoors: the better the weather conditions, the higher the attendance. This negative effect is also obtained for the distance variable. On the other hand, games shown live on TV and those not played on the week-end, show significantly lower

attendance. This effect is more important when matches are televised on a public channel, to which everybody has free access, rather than on private channels to which access is by subscription¹⁷. Previous empirical evidence in this literature was not very conclusive about the effect of televising a match¹⁸. Specifically, for a team with an attendance for a non televised game equal to the sample mean (3772), attendance will decrease by 1386 spectators (36.74%)¹⁹ if the match is televised by a public channel and by 1042 (27.62%) if it is televised by a private channel.

When we include dummies to control for the home team and season effects (column 3) the pattern of the effects we mentioned above does not change in either sign or significance except for a few cases which correspond to variables which show variability only in the home team and season dimensions. This is the case with the budget variable and the income variable whose parameter estimates are not significant while they are not signed as expected. The explanatory power of the model increases substantially by including these controls²⁰.

We also estimated a model that, while similar, has a different functional form for the price and income variables (not in logs). The results are presented in column (4) and we can observe that the results do not change with respect to those in column (3), except for the income variables which are significant in this specification. In fact, when looking at the R^2 for both models they are very similar but higher for the specification in which prices and income variables are in logs. Given that we are estimating the same number of parameters in both models, this implies a preference for the model in column (3). This is also confirmed by means of the J test (Davidson and MacKinnon, 1981) for the null that the model in column (3) rather than the model in column (4) is the true one.

When running a regression of $\log(\text{attendance})$ on the variables included in model (3) plus the predicted $(\log)\text{attendance}$ from model (4), the t-statistic for the coefficient of that predicted (\log) attendance is 1.22, whereas when considering model (4) as the null and using the same type of test, the t-statistic is 4.20. Consequently, we choose the specification in model (3) as our preferred specification in terms of the functional form and the set of explanatory variables to be included.

Results seem to be quite robust to different transformations of the model to control for the home team and seasonal effects. In column (5) we present the results corresponding to the OLS estimates of the model transformed using the special type of differences we mentioned above. This implies that those variables with no variation within a season will cancel out as happens with the home team and season effects. The most relevant change is the higher significance of the effect of Barcelona or Real Madrid being the away team. The explanation may be that these variables are, in some sense, capturing the *ex ante* quality of the away team measured by the budget and the number of internationals variables in the previous specification.

Finally, we estimated the preferred version (column 3) of the attendance equation by correcting the possible endogeneity of prices. As mentioned above, we estimate a reduced form equation for $\log(\text{prices})$ using all the variables included in the demand equation plus four additional instruments to identify the demand equation. The results of the estimation of this price equation are presented in Table A.2 of the Appendix. We must point out the significance of two teams' finishing positions in the previous season in explaining prices. While the higher the away team's position, the higher the price charged, we find the opposite effect for the home team's position.

From these equations we calculated predicted (log) prices which are used instead of the observed price variables in the demand equation in a kind of non-linear two-stage least squares estimator as proposed by Amemiya (1983). The results of this estimation are presented in column (6) of Table 1 and they are very similar to those obtained in column (3), except for the magnitude of the coefficients of the price variables. Nevertheless, when testing for the endogeneity of prices by means of introducing the residual of the price equation as an additional regressor in the demand equation (Smith and Blundell, 1984), the estimated parameter has a t-statistic of 5.93, rejecting the null hypothesis of exogeneity of the price variable²¹.

Price elasticities

One of the objectives of this paper has to do with analyzing the sensitivity of the estimated price elasticities to different assumptions of our model, in particular, the functional form and the exogeneity of prices. As we stated before, the model with a cubic profile for (log) prices (column 3) was preferred to the linear version (column 1). This has important implications in terms of the price elasticities because the linear model implies a constant elasticity whereas for the cubic version the elasticity will vary with prices.

The estimated price elasticity for the linear model is -0.63 , statistically different from a unit elasticity which would be the value in a context of clubs acting as profit maximizers and costs not depending on attendance in a standard monopolistic model.

When estimating the more general model with a cubic profile, the price elasticity (η) becomes

$$\eta = -14.017 + 9.394 \log(\text{PRICE}) - 1.587 [\log(\text{PRICE})]^2$$

In the first row of Table 2 we present the descriptive statistics of the estimated elasticities for the sample used in the estimation. The mean is even smaller in absolute value (-0.295) than the estimated elasticity for the linear model and only 5.19% of the observations have elasticity greater than one in absolute value. In fact, the value for the first decile (-0.571) is also smaller in absolute value than the estimated elasticity for the first model.

(TABLE 2)

Consequently, these results agree with the empirical evidence on estimated price elasticities for professional team sports events. As mentioned in the introduction, this evidence can be rationalized either in a context of profit maximization under different modifications of the standard model or in a context where a club's objective function has arguments other than profits²².

Although the estimated model does not allow us to identify which theoretical framework applies to the Spanish case, the recent transformation of the Spanish clubs into private firms²³ seems to support an explanation for our results based on maximizing an objective function more general than a profit function.

When taking into account the possible endogeneity of prices, the estimated price elasticities change substantially with respect to the previous results. As shown in the second row of Table 2, the mean of the estimated elasticities is almost one in absolute

value (-0.968), with a larger range of variation and a higher percentage of observations with elasticities greater than one in absolute value (27.91%) as compared with the previous model. Consequently, although in both cases most of the observations correspond to clubs working in the inelastic part of the demand curve, the pattern is significantly different after taking into account price endogeneity. It is also relevant to point out that when estimating a linear version of our specification by IV, price elasticity becomes insignificant.

In Table 3 we present the average of the estimated price elasticities for each team for the two versions of our preferred model (the one in which price variables are instrumented and the other in which they are not). We could distinguish different groups of clubs depending on these values. The results in the second column (the model with price instruments) show that out of 27 clubs there are 11 with elasticities higher than one in absolute value and 8 with elasticities smaller than 0.5 in absolute value. The pattern of the distribution of the clubs is similar if we consider the OLS results but, as was stated above, the elasticities are smaller than in the second case.

(TABLE 3)

Contribution of each group of variables to explaining attendance

A final aspect we wish to evaluate is the contribution of each group of variables we included in our model to explaining attendance. We do this by performing F tests for the null hypothesis of the coefficients of each group of variables' separately being equal to zero. In fact, in using the F test, we are comparing the average reduction on the residual sum of squares by each additional estimated parameter included in a particular group of

variables against the average reduction when including all the variables. This gives us a measure of what group of variables most reduces the residual sum of squares when the number of extra parameters to be estimated is taken into account.

In Table 4 we present the results of this exercise for two models: that without home team and season effects (model 2 in Table 1) and that with those effects (model 3 in Table 1). Clearly, in both cases the group of variables capturing *ex ante* quality of the two teams is the group with the highest impact on attendance. On the other hand, when controlling for the unobserved effects the impact of the economic variables is substantially reduced. The group of variables proxying the opportunity costs of attending a match is the second most important group in explaining attendance ahead of home team effects.

Finally, we wish to comment on the importance of home and team quality variables, because of the implications on the effect of revenue sharing on competitive balance, as stated in Késenne (2000) when the absolute value of a game affects attendance. Our results do not show that home team quality (budget, number of wins in the last three games, current league position, number of goals scored in the last match at home and result of the last game) has a larger effect on attendance than away team quality (budget, number of internationals and no defeat in the last four games). In fact, when including in the away team quality variables the dummies corresponding to either Barcelona or Real Madrid as visitors, the impact of the away team quality is clearly higher than that of the home team. So, the necessary conditions for revenue sharing having an effect on competitive balance do not seem to be satisfied.

(TABLE 4)

4. Conclusions

In this paper we have estimated an attendance equation for the Spanish Football League using data on the individual games played during the seasons 1992-93 to 1995-96. We concentrated our attention on specification issues. We have included all the types of variables (economic and sectoral) proposed in the literature as explanatory factors in this kind of demand equations. Additionally, we have given attention to the functional form of the equation and the potential endogeneity of prices, specifically, with respect to their implications for estimated price elasticities. We also have employed the panel data structure of our data set to control for the effect of unobservables potentially correlated with the regressors.

As it is usual in this literature, we estimated price elasticities which, in general, are less than one in absolute value, but these estimates show substantial differences depending on the functional form and consideration of the potential endogeneity of prices.

At the same time, we have measured the contribution of each group of explanatory factors on explaining attendance, concluding that those variables related to *ex ante* quality of the two teams are those with the highest explanatory power.

As the sample period corresponds precisely to the initial stages of most Spanish football clubs' roles as private firms, future research needs to extend the sample period in attempting to characterize their economic behaviour more accurately. This would permit a more detailed analysis of the effect of televising football matches on attendance, given that the *pay per view* option could be included in the analysis.

References

- Amemiya, T. (1983) Non Linear Regression Models, in *Handbook of Econometrics, Volume 1*, (ed.) Z. Griliches and M.D. Intriligator, North-Holland, Amsterdam, pp. 333-389
- Baimbridge, M., Cameron, S. and Dawson, P. (1996) Satellite Television and the Demand for Football: A Whole New Ball Game?, *Scottish Journal of Political Economy*, **43**, 317-333
- Balestra, P. and Nerlove P, M. (1966) Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas, *Econometrica*, **34**, 586-612
- Borland, J. (1987) The Demand for Australian Rules Football, *Economic Record*, **63**, 220-230
- Borland, J. and Lye, J. (1992) Attendance at Australian Rules Football: A Panel Study, *Applied Economics*, **24**, 1053-1058
- Boyd, D.W. and Boyd, L.A. (1998) The Home Field Advantage: Implications for the Pricing of Tickets to Professional Team Sporting Events, *Journal of Economics and Finance*, **22**, 169-179
- Bruggink, T.H. and Eaton, J.W. (1994) Rebuilding Attendance in Major League Baseball: The Demand for Individual Games, in *Baseball Economics. Current Research* (eds.) J.L. FizeL, E. Gustafson and L. Hadley, Praeger Publishers, Westport, Connecticut, pp. 9-31
- Cairns, J.A. (1988) Uncertainty of Outcome and the Demand for Football, Discussion Paper 88-02. Department of Economics. University of Aberdeen
- Cairns, J.A. (1990) The Demand for Professional Team Sports, *British Review of Economic Issues*, **12**, 1-20
- Cairns, J.A., Jennett, N. and Sloane, P.J. (1986) The Economics of Professional Team Sport: a Survey of Theory and Evidence, *Journal of Economic Studies*, **13**, 3-80
- Carmichael, F., Millington, J. and Simmons, R. (1999) Elasticity of Demand for Rugby League Attendance and the Impact of BskyB, *Applied Economics Letters*, **6**, 797-800
- Davidson, R. and MacKinnon, J. (1981) Several Tests of Model Specification in the Presence of Alternative Hypotheses, *Econometrica*, **49**, 781-793
- Demmert, H.G. (1973) *The Economics of Professional Team Sports*, Lexington Books, Lexington, Massachusetts

- Downward, P. and Dawson, A. (2000) *The Economics of Professional Team Sports*, Routledge, London
- El Hodiri, M. and Quirk, J. (1971) An Economic Model of a Professional Sports League, *Journal of Political Economy*, **79**, 1302-1319
- Falter, J.M. and Pérignon, C. (2000) Demand for Football and Intramatch Winning Probability: An Essay on the Glorious Uncertainty of Sports, *Applied Economics*, **32**, 1757-1765
- Ferguson, D.G., Stewart, K.G., Jones, J.C.H. and Le Dressay, A. (1991) The Pricing of Sports Events: Do Teams Maximize Profit?, *Journal of Industrial Economics*, **39**, 297-310
- Fort, R. (2000) European and North American Sports Differences (?), *Scottish Journal of Political Economy*, **47**, 431-455
- Heilmann, L. and Wendling, W.R. (1976) A Note on Optimum Pricing Strategies for Sport Events, in *Management Science in Sports*, (eds.) R.E. Machol and S.P. Ladany, North-Holland Publishing Company, Amsterdam, pp. 91-99
- Hoehn, T. And Szymanski, S. (1999) The Americanization of European Football, *Economic Policy*, **28**, 203-240
- Jennett, N. (1984) Attendance, Uncertainty of Outcome and Policy in the Scottish League Football, *Scottish Journal of Political Economy*, **31**, 176-198
- Jones, J.C.H., Schofield, J.A. and Giles, D.E.A. (2000) Our Fans in the North: the Demand for British Rugby League, *Applied Economics*, **32**, 1877-1887
- Késenne, S. (2000) Revenue Sharing and Competitive Balance in Professional Team Sports, *Journal of Sports Economics*, **1**, 56-65
- Knowles, G., Sherony, K. and Hauptert, M. (1992) The Demand for Major League Baseball: A Test of the Uncertainty of Outcome Hypothesis, *American Economist*, **36**, 72-80
- Kuypers, T. (1996) *The Beautiful Game?. An Econometric Study of Why People Watch English Football*, Discussion Paper in Economics 96-01, University College London
- Marburger, D.R. (1997) Optimal Ticket Pricing for Performance Goods, *Managerial and Decision Economics*, **18**, 375-381
- Peel, D. and Thomas, D. (1988) Outcome Uncertainty and the Demand for Football: An Analysis of Match Attendances in the English Football League, *Scottish Journal of Political Economy*, **35**, 242-249
- Peel, D. and Thomas, D. (1997) Handicaps, Outcome Uncertainty and Attendance Demand, *Applied Economics Letters*, **4**, 567-570

- Rodríguez, P. (2001) *Análisis económico de la demanda de entradas en el fútbol español*, Ph.D. thesis, Universidad de Oviedo
- Salant, D.J. (1992) Price Setting in Professional Team Sports, in *Diamonds Are Forever. The Business of Baseball* (ed.) P.M. Sommers, The Brookings Institution, Washington, pp. 77-90
- Sarmiento, M. (1994) *Selección española de fútbol: sus hombres, uno a uno*, Fundación del Fútbol Profesional, Madrid.
- Schofield, J.A. (1983) Performance and Attendance at Professional Team Sports, *Journal of Sports Behaviour*, **6**, 196-206
- Sloane, P.J. (1971) The Economics of Professional Football: The Football Club as a Utility Maximizer, *Scottish Journal of Political Economy*, **17**, 121-146
- Smith, R. and Blundell, R. (1984) An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labour Supply, *Econometrica*, **54**, 679-685
- Szymanski, S. (1997) The English Football Industry: Profit, Performance and Industrial Structure, *International Review of Applied Economics*, **11**, 135-153
- Szymanski, S. and Kuypers, T. (1999) *Winners and Losers*, Viking, London
- Welki, A.M. and Zlatoper, T.J. (1994) US Professional Football: The Demand for Game-Day Attendance in 1991, *Managerial and Decision Economics*, **15**, 489-495
- Wilson, P.R.D. and Sim, B. (1995) The Demand for Semi-Pro League Football in Malaysia 1989-91: A Panel Data Approach, *Applied Economics*, **27**, 131-138

Table 1: Estimates of the attendance equation
(Endogenous variable: log(attendance))
(N = 1580)

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Coef.</i> <i>t-stat.</i>	<i>Coef.</i> <i>t-stat.</i>	<i>Coef.</i> <i>t-stat.</i>	<i>Coef.</i> <i>t-stat.</i>	<i>Coef.</i> <i>t-stat.</i>	<i>Coef.</i> <i>t-stat.</i>
Socio-economic variables						
Log(Price)	-0.630 (10.04)	-8.984 (3.21)	-14.017 (5.30)		-12.879 (4.88)	-48.781 (3.08)
Log(Price)**2		2.831 (2.90)	4.697 (4.98)		4.289 (4.53)	16.095 (3.00)
Log(Price)**3		-0.310 (2.76)	-0.529 (4.76)		-0.478 (4.27)	-1.783 (2.94)
Price				-0.136 (4.79)		
Price**2				0.004 (4.43)		
Price**3 (divided by 1000)				-0.047 (4.64)		
Log(Income)	0.513 (4.93)	1980.4 (4.84)	-1652.1 (1.20)			-1935.0 (1.32)
Log(Income)**2		-212.5 (4.90)	171.46 (1.17)			201.17 (1.29)
Log(Income)**3		7.600 (4.96)	-5.945 (1.14)			-6.984 (1.27)
Income				-0.004 (3.36)		
Income**2				0.267 (2.75)		
Income**3				-5.920 (2.38)		
Log(Population)	0.247 (5.05)	0.342 (6.81)	0.026 (0.28)	-0.019 (0.21)	-0.033 (0.59)	-0.034 (0.34)
Ex ante quality						
Budget (h)	0.017 (8.12)	0.013 (6.09)	-0.010 (1.14)	-0.012 (1.35)		-0.016 (1.75)
Budget (v)	0.014 (3.03)	0.014 (3.11)	0.014 (3.41)	0.013 (3.17)		0.014 (3.25)
Number of internationals (v)	0.015 (2.41)	0.016 (2.51)	0.015 (2.81)	0.015 (2.85)		0.016 (2.89)
Away team Barcelona	0.455 (2.07)	0.428 (1.99)	0.407 (2.06)	0.445 (2.24)	1.328 (20.01)	0.462 (2.19)
Away team Real Madrid	0.275 (1.27)	0.271 (1.28)	0.264 (1.37)	0.302 (1.56)	1.172 (17.40)	0.300 (1.49)
Rivalry	0.491 (5.79)	0.450 (5.34)	0.453 (5.50)	0.438 (5.22)	0.420 (5.53)	0.496 (2.27)
“Day of the club” match	0.217 (2.85)	0.190 (2.44)	0.197 (2.53)	0.193 (2.44)	0.166 (2.49)	0.246 (2.27)
Current quality						
No. of wins in the last 3 games (h)	0.047 (1.95)	0.044 (1.93)	0.028 (1.38)	0.031 (1.53)	0.017 (0.75)	0.036 (1.59)
Score last game (h)	0.046 (4.54)	0.043 (4.32)	0.040 (4.49)	0.041 (4.53)	0.034 (4.08)	0.040 (4.41)
Goals last game at home (h)	0.045 (3.20)	0.046 (3.32)	0.038 (3.11)	0.037 (3.08)	0.048 (4.40)	0.038 (3.12)
Standings (h)	-0.008 (1.39)	-0.005 (0.91)	-0.016 (2.88)	-0.015 (2.62)	-0.036 (4.90)	-0.018 (2.15)
No defeat in last 4 games (v)	0.119 (2.66)	0.099 (2.27)	0.103 (2.58)	0.115 (2.86)	0.027 (0.56)	0.125 (3.04)
No chance to win the championship (h)	-0.215 (3.91)	-0.218 (4.01)	-0.160 (3.17)	-0.163 (3.20)	-0.052 (0.55)	-0.159 (2.79)
No chance of leaving relegation zone (h)	-1.162 (4.25)	-1.074 (3.92)	-1.011 (4.40)	-1.035 (4.47)	-0.577 (0.77)	-1.101 (4.31)

Uncertainty							
Difference in league positions (h-v)	0.022 (6.12)	0.022 (6.19)	0.021 (6.79)	0.021 (6.67)	0.034 (12.17)	0.022 (5.14)	
Difference in league positions**2 (h-v)	0.001 (3.61)	0.001 (3.78)	0.001 (3.35)	0.001 (3.25)	0.000 (1.19)	0.001 (2.74)	
Closeness of league positions	0.098 (2.30)	0.088 (2.15)	0.047 (1.31)	0.050 (1.37)	0.055 (1.59)	0.046 (1.25)	
Uncertainty of championship (h)	-0.001 (5.85)	-0.001 (6.08)	-0.001 (3.54)	-0.001 (3.81)	0.000 (0.27)	-0.001 (3.04)	
Opportunity cost							
No rain, hot	0.374 (6.22)	0.355 (5.94)	0.303 (5.49)	0.305 (5.50)	0.269 (5.49)	0.307 (5.48)	
No rain, cold	0.334 (5.61)	0.325 (5.48)	0.270 (4.90)	0.273 (4.95)	0.246 (5.08)	0.271 (4.88)	
Televised by public channels	-0.427 (5.78)	-0.458 (6.57)	-0.464 (7.26)	-0.459 (6.93)	-0.425 (8.18)	-0.454 (6.59)	
Televised by a private channel	-0.321 (4.79)	-0.323 (5.22)	-0.318 (5.69)	-0.318 (5.62)	-0.344 (7.28)	-0.330 (5.48)	
Not played on the weekend	-0.235 (4.01)	-0.233 (3.99)	-0.216 (4.00)	-0.220 (4.10)	-0.245 (4.99)	-0.239 (4.28)	
Distance	-0.525 (6.86)	-0.503 (6.69)	-0.497 (6.74)	-0.496 (6.63)	-0.521 (7.60)	-0.501 (6.37)	
Home team Tenerife	0.327 (4.06)	0.451 (5.59)	0.955 (6.48)	0.939 (6.27)		0.878 (5.18)	
Away team Tenerife	-0.507 (5.34)	-0.494 (5.24)	-0.492 (5.79)	-0.491 (5.76)	-0.546 (7.42)	-0.498 (5.71)	
Constant	0.667 (0.46)	-6137.7 (4.77)	5337.2 (1.24)	30.78 (5.26)		6270.3 (1.36)	
Home team effects	NO	NO	YES	YES	NO	YES	
Season effects	NO	NO	YES	YES	NO	YES	
R ²	0.6252	0.6489	0.7270	0.7229	0.4776	0.7121	

Notes: (h) and (v) refer to the home team and the away team respectively.

In model (4), the quadratic and cubic terms of income are divided by 10^6 and 10^{12} , respectively, and the cubic term of the price variable by 10^3 .

The distance variable is measured in thousands kilometres.

Table 2: Descriptive statistics of the estimated price elasticities

	<i>Mean</i>	<i>Max.</i>	<i>Min.</i>	<i>10%</i>	<i>25%</i>	<i>50%</i>	<i>75%</i>	<i>90%</i>	<i>% (h<-1)</i>
Model 3	-0.295	-3.947	-0.116	-0.571	-0.321	-0.178	-0.124	-0.120	5.19
Model 6	-0.968	-14.100	-0.352	-1.713	-1.049	-0.542	-0.407	-0.368	27.91

Table 3: Average price elasticities for each team

<i>Team</i>	<i>Price elasticity Model 3 Table 1</i>	<i>Price elasticity Model 6 Table 1</i>
Albacete	-0.2679	-0.7778
Athletic de Bilbao	-0.1729	-0.5067
Atlético de Madrid	-0.1598	-0.5097
Barcelona	-0.2875	-0.9236
Betis	-0.4142	-1.1625
Burgos	-0.2154	-0.6844
Cádiz	-0.2898	-1.1038
Celta	-0.2179	-0.7875
Compostela	-0.3934	-1.1198
Deportivo de La Coruña	-0.1978	-0.6277
Español	-0.2208	-0.6863
Logroñés	-0.3250	-0.9137
Lleida	-0.5441	-1.6093
Mérida	-0.3696	-1.3574
Osasuna	-0.2774	-0.9976
Oviedo	-0.3377	-0.9498
Racing de Santander	-0.1577	-0.5391
Rayo Vallecano	-0.3320	-1.0726
Real Madrid	-0.7358	-2.7388
Real Sociedad	-0.1891	-0.5503
Salamanca	-0.2464	-0.6676
Sevilla	-0.3290	-1.1294
Sporting de Gijón	-0.2407	-0.7758
Tenerife	-0.3002	-1.1486
Valencia	-0.3041	-1.0355
Valladolid	-0.2964	-0.9421
Zaragoza	-0.3748	-1.3269

Table 4: Significance of each set of explanatory variables

	Model without home team and season effects				Model with home team and season effects			
	<i>SSR</i>	<i>K</i>	<i>r</i>	<i>F-test</i>	<i>SSR</i>	<i>K</i>	<i>r</i>	<i>F-test</i>
Basic model	604.07	33			469.68	61		
<u>Set of excluded variables</u>								
Economic variables	731.62	26	7	46.63	513.83	54	7	20.39
<i>Ex ante</i> quality	774.20	26	7	62.20	621.20	54	7	69.96
Current quality	658.22	26	7	19.80	514.87	54	7	20.86
Home team quality	641.53	28	5	19.17	488.63	56	5	12.25
Away team quality	626.50	30	3	19.14	490.73	58	3	22.66
Uncertainty	646.07	29	4	26.87	493.39	57	4	19.16
Opportunity cost	715.08	25	8	35.51	529.55	54	7	27.64
Home team effects					596.98	36	25	16.46
Season effects					482.58	58	3	13.89

Notes: The basic models are those in columns (2) and (3) of Table 1 for the models without and with effects, respectively.

SSR: Residual sum of squares.

K = Number of parameters.

r = Number of restrictions

APPENDIX

Table A.1 Descriptive statistics and sources

<i>Variable</i>	<i>Mean</i>	<i>St. Dev.</i>	<i>Source</i> ⁴
Attendance	3772.59	5101.24	LNFP
Socio-economic variables			
Price ¹	2047.53	662.16	LNFP
Income ¹	1292.86	277.44	BBVA
Population ²	1089.36	1058.43	BBVA
Ex ante quality			
Budget (h) ¹	1790.54	1736.59	LNFP
Budget (v) ¹	1776.39	1727.63	LNFP
Number of internationals (v)	11.85	4.90	<i>Dinámico</i> ⁵
Away team Barcelona	0.0487		
Away team Real Madrid	0.0487		
Rivalry	0.0468		
“Club Day” match	0.0563		LNFP
Current quality			
No. of wins in the last 3 games (h)	0.9127	0.8255	
Score last game (h)	-0.4025	1.7253	
Goals last game at home (h)	1.0887	1.1477	
Standings (h)	10.7006	6.0825	
No defeat in last 4 games (v)	0.1587		
No chance of winning the championship (h)	0.1791		
No chance of leaving relegation zone (h)	0.0089		
Uncertainty			
Difference in league positions (h-v)	0.3329	8.2663	
Closeness in league positions	0.3006		
Uncertainty of championship (h) ³	180.719	143.969	
Opportunity cost			
No rain, hot	0.5361		<i>Dinámico</i>
No rain, cold	0.3627		<i>Dinámico</i>
Televised by public channels	0.1006		LNFP
Televised by a private channel	0.0987		LNFP
Not played on the weekend	0.0715		LNFP
Distance	544.447	268.072	<i>Road map</i>
Home team Tenerife	0.0494		
Away team Tenerife	0.0487		

- Notes:
- 1 These variables are expressed in real terms (1991 pesetas). Income in thousands of pesetas and budgets in millions of pesetas.
 - 2 Population is in thousands.
 - 3 This is based on Kuypers (1996) measure.
 - 4 LNFP: Liga Nacional de Fútbol Profesional
BBVA: Fundación BBVA, *Renta nacional de España y su distribución provincial*
Dinámico: Football yearbook
 - 5 We also used information from Sarmiento (1994)

Table A.2 Estimation of a price equation (reduced form)
(Endogenous variable: log(price))

<i>Variable</i>	<i>Coef.</i>	<i>t-stat.</i>
Previous season standings (h)	0.007	3.37
Previous season standings (v)	-0.004	2.32
Previous season in Second Division (h)	0.017	0.48
Previous season in Second Division (v)	-0.036	1.31
Log(Capacity)	-0.035	1.34
Socio-economic variables		
Log(Income)	-228.83	0.36
Log(Income)**2	22.221	0.33
Log(Income)**3	-0.711	0.30
Log(Population)	0.036	0.67
Ex ante quality		
Budget (h)	0.100	0.22
Budget (v)	0.009	0.06
Number of internationals (v)	0.003	1.20
Away team Barcelona	0.131	1.55
Away team Real Madrid	0.107	1.29
Rivalry	0.142	4.14
“Club Day” match	0.138	4.42
Current quality		
No. of wins in the last 3 games (h)	0.010	0.97
Score last game (h)	-0.001	0.23
Goals last game at home (h)	0.004	0.79
Standings (h)	-0.012	4.34
No defeat in last 4 games (v)	0.012	0.62
No chance of winning the championship (h)	-0.038	1.85
No chance of leaving relegation zone (h)	-0.150	2.17
Uncertainty		
Difference in league positions (h-v)	0.005	3.93
Difference in league positions**2 (h-v)	-0.000	1.15
Closeness in league positions	0.002	0.15
Uncertainty of championship (h)	0.270	2.62
Opportunity cost		
No rain, hot	0.022	1.04
No rain, cold	0.003	0.13
Televised by public channels	-0.023	0.79
Televised by a private channel	-0.036	1.53
Not played on the weekend	-0.032	1.52
Distance	-0.042	1.33
Home team Tenerife	-0.225	3.29
Away team Tenerife	-0.015	0.48
Constant	781.12	0.39
Home team effects	YES	
Season effects	YES	
R ²	0.5052	

Notes: (h) and (v) refer to the home team and the away team respectively

The distance variable is measured in thousands of kilometres

The uncertainty for the title variable is measured in thousands

¹ See Schofield (1983), Cairns *et al.* (1986), Cairns (1990) and Downward and Dawson (2000) for surveys of this literature.

² See Rodríguez (2001) for a recent survey of the empirical specification issues related to the estimation of attendance equations.

³ See Szymanski and Smith (1997), Szymanski and Kuypers (1999) and Hoehn and Szymanski (1999) for complete analysis of the British football industry.

⁴ The attendance of season ticket holders will be explained by a different model where some variables, in particular the economic variables, will not play any explanatory role. See Rodríguez (2001) for some preliminary results for this type of attendance in the Spanish Football League.

⁵ This form of measuring the price variable has been used previously in the literature. See Jennett (1984), Borland (1987), Borland and Lye (1992) and Falter and Pérignon (2000) among others. We prefer this to the usual average ticket price as in this manner we avoid the inclusion of the endogenous variable (attendance) in the definition of the price variable. In any case, as we shall see later on, we control for the possible endogeneity of the price variable we use.

⁶ As far as we know Falter and Pérignon (2000) is the only paper in the literature on attendance at professional team sporting events that includes this type of variable in a demand equation.

⁷ We report the variables included in the final specification. Other variables proxying the same effects have been included in previous estimations, not reported here but available on request.

⁸ See Cairns (1988) for a complete discussion of how to model uncertainty in these demand equations.

⁹ We have also considered the possibility that the scheduling of a match might have an influence on attendance, but the estimated effect was not significant.

¹⁰ Each team played 38 matches each season, except in 1995-96 when they played 42 matches.

¹¹ Some studies have shown a tendency to control for the unobserved component corresponding to each fixture in a particular season, as in Baimbridge *et al.* (1996) and Carmichael *et al.* (1999), whereas other studies include dummies for the initial and final games of the season, as in Peel and Thomas (1988) and Wilson and Sim (1995). In our study when attempting to control for this effect, we did not obtain significant estimates.

¹² This is the usual approach when transforming a dynamic model for panel data previous to its estimation by IV or GMM.

¹³ The F-statistic for testing the null hypothesis of a linear specification against the alternative of a cubic one is 26.92, rejecting this null at a 5% significance level ($F_{4,1550}=2.37$).

¹⁴ We have not included the number of internationals of the home team in the final specification because it was wrongly signed when included, unless we eliminated the budget of the home team as a explanatory variable.

¹⁵ We considered the possibility of including a dummy variable for the home team having won the championship, but in three out four of the seasons included in the sample the champion was not known until the last game was played and in the remaining season the championship was won one game before the end. For this reason, we decided not to include it in the model as this dummy would have the value one for only a single observation.

¹⁶ Given that there is no information available for betting odds on Spanish football, we can not proxy the predicted probability of winning a game as a measure of uncertainty of the outcome by means of this variable as in Peel and Thomas (1988, 1997) and Knowles *et al.* (1992).

¹⁷ In these seasons, the “pay-per-view” system was not still available in Spain.

¹⁸ For US professional football, Welki and Zlatoper (1994) found that games which are blacked out for local TV are more poorly attended; for English football Kuypers (1996) did not found a significant effect for this variables; and for major league baseball Bruggink and Eaton (1996) obtained different effects for games televised on a local free channel and on premium cable. Negative effects of television on baseball attendance were found in Demmert (1973). Baimbridge *et al.* (1996) argue that the net effect of television on attendance is indeterminate.

¹⁹ Note that we can not interpret the coefficients of the dummies for a match being televised as a rate of increase of the endogenous variable because they are not small rates. The figures calculated above are not based on this approximation.

²⁰ The F-statistic for testing the null hypothesis of not including these controls is 13.16, rejecting this null at a 5% significance level ($F_{33, 1518}=1.46$)

²¹ Given the cubic profile for prices, we also introduced the square and cubic residuals of the price equation in the demand equation, with the coefficients of the linear and quadratic terms being significant.

²² See Fort (2000) for a comparison of European and North American sports in terms of team objectives.

²³ This transformation took place in 1992 and all the clubs were involved with the exception of Athletic of Bilbao, Barcelona, Osasuna and Real Madrid.