The Determinants of Football Match Attendance Revisited

Empirical Evidence From the Spanish Football League

JAUME GARCÍA
Universitat Pompeu Fabra

PLÁCIDO RODRÍGUEZ
Universidad de Oviedo

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. The authors 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. The authors obtain the expected effects on attendance for all the variables. The estimated price elasticities are, in general, smaller than one in absolute value but are sensitive to the specification issues, in particular, the endogeneity of prices.

The analysis of the determinants of attendance at professional team sports events is one of the topics that has received most attention in the empirical literature of the 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 that try to capture the heterogeneity of this type of good (a match). This has been done using different types of data sets, depending on data availability and the objectives of the study.[2]

AUTHORS' NOTE: *We are very grateful to César Rodríguez, Joaquín Lorences, and the participants in a seminar at the University of Oviedo for their comments on earlier stages of this work, and also to an anonymous referee for comments on the previous version of this article. Special thanks are also due to the Liga Nacional de Fútbol Profesional for providing the data used in this article. 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.*

JOURNAL OF SPORTS ECONOMICS, Vol. 3 No. 1, February 2002 18–38

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 articles include prices as an explanatory factor and not many take into account the interpretation of the price elasticities obtained and try to rationalize it. Articles by Heilmann and Wendling (1976); Ferguson, Stewart, Jones, and Le Dressay (1991); Salant (1992); Marburger (1997); and Boyd and Boyd (1998) are among those that try, in the context of a profit maximizing behavior 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 article, we try to bring new evidence from European football[3] 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 economic and sectoral variables for all the matches played in the Spanish First Division League during the seasons 1992-93 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 to estimate price elasticity consistently while also taking into account the possible endogeneity of prices. Our empirical results show that price elasticities are underestimated in absolute value when not taking into account the last issue. We also analyze the incidence of the specification of the functional form in the estimated elasticities. Finally, we also evaluate the importance of the different groups of variables in explaining attendance.

The article is organized as follows. In the next section, we present the specification of the empirical model, followed by a discussion of the estimation results. The article ends with a summary of the main conclusions.

MODEL SPECIFICATION

Model and Variables

We specify and estimate a fairly standard demand equation distinguishing, among the explanatory factors that 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 A1 in the appendix.

The endogenous variable is the (log)number of tickets sold for a match (attendance), not including those for children and season tickets.[4] Among the economic variables we include: prices, measured by the price of the cheapest ticket[5] deflated by the Consumer Price Index; 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, that is, 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 productivity;[6] 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 (Club's 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 include:[7] the number of home team wins in the past 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 past match at home by the home team; a dummy for the away team not having lost a game out of the past four; and two dummies for the home team having 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 uncertainty.[8] 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 because the higher the uncertainty, the smaller the value of the indicator we have defined.

Finally, we include a set of variables that capture the opportunity cost of attending a football match. We model the effect of weather conditions with dummy variables that correspond to the following situations: no rain, high temperature; no rain, low temperature; and rainy days, which is the omitted dummy. We expect that the better the weather is, the higher the attendance will be. The second factor has to do with the game's being televised. Because 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 weekday rather than on the weekend, attendance should decrease.[9] 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 that 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 that 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 that refer to prices,[11] 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's Day match are included in this group. The second group of variables, included in the vector Z_{it}, is made up of those variables that vary depending on the season and 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 past three games, Kuypers's 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 past 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 that 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 that 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:

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

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 that refer to a particular away team (Barcelona, Real Madrid, and Tenerife).[12]

Equation 1 is estimated by ordinary least squares (OLS),[13] which for consistency requires 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 season.[14] This transformation eliminates both the home team effect and the season effect but also eliminates those variables that 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 generalized least squares (Balestra & 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 Equation 1 with that of those models estimated using data (averages) on seasons rather than data on games (e.g., Jones, Schofield, & Giles, 2000). This amounts to transforming our model by averaging variables for each home team in each season.

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. Because, in analyzing Spanish football clubs' optimization behavior, we are interested in estimating price elasticity consistently, we estimate the model by Instrumental Variables (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 that can be sold (capacity). These variables allow us to identify the demand equation.

EMPIRICAL RESULTS

In this section, we present the main results of the specification and estimation of an attendance equation for Spanish football with special attention to the estimates of the price elasticities and their interpretation 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 1,580 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 nonnested 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.

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).[15] All the coefficients in both models have

TABLE 1: Estimates of the Attendance Equation

	1		*2*		*3*		*4*		*5*		*6*	
	Coefficient	t-*Statistic*	*Coefficient*	t-*Statistic*	*Coefficient*	t-*Statistic*	*Coefficient*	t-*Statistic*	*Coefficient*	t-*Statistic*	*Coefficient*	t-*Statistic*
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 1,000)							–0.047	4.64				
Log(income)	0.513	4.93	1,980.4	4.84	–1,652.1	1.20					–1,935.0	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
Club's Day 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												
Number of wins in the past 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 past 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 past 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 past 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	

NOTE: h = home team; v = away team. N = 1,580; endogenous variable is log(attendance). 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 of kilometers.

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; that is, attendance is a normal good. In fact, in the simplest specification of Column 1, the estimated constant income elasticity, which is 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 also has a positive effect on attendance.[16]

Scoring an additional goal either in the past match or in the past match at home and having an extra victory in the past 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 past four games also increases 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 attendance.[17]

Given that there is no information available for betting odds on Spanish football for the period under consideration, we cannot use this variable to proxy the uncertainty of the outcome by means of either the predicted probability of winning a game, as in Peel and Thomas (1988, 1997) and Knowles, Sherony, and Haupert (1992), or the evenness of the contest, as in Forrest and Simmons (2001). Instead, we use a quadratic form in the difference of league positions, which is consistent with measuring maximum uncertainty when the difference in league positions compensates home advantage. We would expect a negative sign for the quadratic term and a positive one for the linear term, but in all cases we estimate a positive sign for the quadratic term. There are some identification problems when using league positions to proxy uncertainty and also quality as mentioned above. We also included a dummy defined in terms of the closeness of the league positions that is not significant in our preferred models (3 and 6), although correctly signed. Consequently, no definite conclusions about the effect of the uncertainty of the outcome can be obtained because there is no information available to measure properly this variable. On the contrary, the variable measuring the uncertainty of the championship for the home team has the expected positive effect, although it has a negative coefficient given the way in which this variable has been defined, as mentioned in the previous section.

Poor weather conditions discourage people from attending football matches because 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 television and those not played on the weekend 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 subscription.[18] Previous empirical evidence in this literature was not very conclusive about the effect of televising a match.[19] Specifically, for a team with an attendance for a nontelevised game equal to the sample mean (3,772), attendance will decrease by 1,386 spectators (36.74%)[20] if the match is televised by a public channel and by 1,042 (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 that correspond to variables that 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 although they are not signed as expected. The explanatory power of the model increases substantially by including these controls.[21]

We also estimated a model (Column 4) that, although similar, has a different functional form for the price and income variables (not in logs). The choice of the specification in Column 3 is confirmed by means of the *J* test (Davidson & 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(attendance) on the variables included in Model 3 plus the predicted (log)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 international variables in the previous specification.

Finally, we estimate 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(prices) using all the variables included in the demand equation plus additional instruments to identify the demand equation. The instruments refer to previous season standings and the fact of being in the second division the past season for both home and away teams. They capture the information the clubs have at the beginning of the season when grouping the games in terms of the potential interest for the spectators. We also include all the variables in the attendance equation because the final decision about prices is taken considering the expected attendance. The results of the estimation of this price equation are presented in Table A2 of the appendix. We must point out the significance of both teams' finishing positions in the previous season in explaining prices. Whereas 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 calculate predicted (log)prices that are used instead of the observed price variables in the demand equation in a kind of nonlinear two-stage least squares estimator as proposed by Amemiya (1983). The results of this estimation are presented in Column 6 of Table 1 and 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 & Blundell, 1984), the estimated parameter has a *t*-statistic of 5.93, rejecting the null hypothesis of exogeneity of the price variable.[22]

Price Elasticities

One of the objectives of this article 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 that 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

TABLE 2: Descriptive Statistics of the Estimated Price Elasticities

	Mean	*Max.*	*Min.*	*10%*	*25%*	*50%*	*75%*	*90%*	*% (η < –1)*
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

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

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 1 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.

When taking into account the possible endogeneity of prices, the estimated price elasticity changes substantially with respect to the previous results. It becomes

$$\eta = -48.781 + 32.190 \log(\text{PRICE}) - 5.349 [\log(\text{PRICE})]^2.$$

As shown in the second row of Table 2, the mean of the estimated elasticities is almost 1 in absolute value (–0.968), with a larger range of variation and a higher percentage of observations with elasticities greater than 1 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 (log)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 (OLS, IV) of our preferred specification (Models 3 and 6). We could distinguish different groups of clubs depending on these values. As pointed out by a referee, variation in price elasticity across clubs is to be expected as clubs have different degrees of local monopoly power, depending on the availability of substitutes in their areas.

The results in the first column of Table 3 show that all clubs have average price elasticities below 1 in absolute value. In fact, out of 27 clubs there are only 2 with a price elasticity greater than 0.5. These elasticities are statistically different from 1. The results in the second column (the model with prices instrumented) show that out of 27 clubs there are 11 with average price elasticities higher than 1 in absolute value. By means of a Wald test we cannot reject for all clubs the null hypothesis of a unit elasticity, but this result very much depends on the high covariances of the esti-

TABLE 3: Average Price Elasticities for Each Team

	Price Elasticity	
Team	*Model 3, Table 1*	*Model 6, Table 1*
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

mated parameters of the cubic profile when using IV. Finally, when comparing both models, the pattern of the distribution of the clubs is similar in both cases but, as was stated above, the elasticities are smaller in the case of the OLS estimation. This underestimation of the price elasticity when not taking into account the endogeneity of prices is consistent with the expected positive correlation of those unobserved factors affecting both prices and attendance.

Consequently, results from Model 3 agree with the empirical evidence on estimated price elasticities for professional team sports events, whereas those from Model 6 seem to give some support for a nearly elastic or elastic demand for a significant proportion of clubs. 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 profits.[23] On the other hand, price elasticities can be underesti-

TABLE 4: Significance of Each Set of Explanatory Variables

	Model Without Home Team and Season Effects				*Model With Home Team and Season Effects*			
	SSR	*K*	*r*	*F Test*	*SSR*	*K*	*r*	*F Test*
Basic model	604.07	33			469.68	61		
Set of excluded variables								
Economic variables	731.62	26	7	46.63	513.83	54	7	20.39
Ex ante 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

NOTE: SSR = residual sum of squares; K = number of parameters; r = number of restrictions. The basic models are those in Columns 2 and 3 of Table 1 for the models without and with effects, respectively.

mated as a consequence of the omission of other components of total cost of attendance, like transport costs, for which we do not have information in our data set.

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 limited companies seems to support an explanation for our results based on maximizing an objective function more general than a profit function.[24]

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 a 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 effect on attendance. On the other hand, when controlling for the unobserved effects the effect of the economic variables has substantially reduced, in particular, that of income. 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 when the absolute value of a game affects attendance (Késenne, 2000). Our results do not show that home team quality (budget, number of wins in the past three games, current league position, number of goals scored in the past match at home, and result of the past game) has a larger effect on attendance than away team quality (budget, number of internationals, and no defeat in the past four games). In fact, when including in the away team quality variables the dummies corresponding to either Barcelona or Real Madrid as visitors, the effect 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.

CONCLUSIONS

In this article 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 is usual in this literature, we estimated price elasticities that, 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. In fact, when this last issue is taken into account, results seem to support a nearly elastic demand for a significant proportion of clubs.

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 limited companies, future research needs to extend the sample period in attempting to characterize their economic behavior 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.

APPENDIX
TABLE A1: Descriptive Statistics and Sources

Variable	*Mean*	SD	*Source*
Attendance	3,772.59	5,101.24	LNFP
Economic variables			
Price[a]	2,047.53	662.16	LNFP
Income[a]	1,292.86	277.44	BBVA
Population[b]	1,089.36	1,058.43	BBVA
Ex ante quality			
Budget (h)[a]	1,790.54	1,736.59	LNFP
Budget (v)[a]	1,776.39	1,727.63	LNFP
Number of internationals (v)	11.85	4.90	*Dinámico*[c]
Away team Barcelona	0.0487		
Away team Real Madrid	0.0487		
Rivalry	0.0468		
Club's Day match	0.0563		LNFP
Current quality			
Number of wins in the past 3 games (h)	0.9127	0.8255	
Score past game (h)	–0.4025	1.7253	
Goals past game at home (h)	1.0887	1.1477	
Standings (h)	10.7006	6.0825	
No defeat in past 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)[d]	180.719	143.969	
Opportunity cost			
No rain, hot	0.5361		*Dinámico*
No rain, cold	0.3627		*Dinámico*
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	Road map
Home team Tenerife	0.0494		
Away team Tenerife	0.0487		

NOTE: LNFP = Liga Nacional de Fútbol Profesional; BBVA = Fundación BBVA, *Renta Nacional de España y su Distribución Provincial*; and *Dinámico* = football yearbook; h = home team; v = away team.
a. These variables are expressed in real terms (1991 *pesetas*). Income in thousands of pesetas and budgets in millions of pesetas.
b. Population is in thousands.
c. We also used information from Sarmiento (1994).
d. This is based on Kuypers (1996) measure.

TABLE A2: Estimation of a Price Equation (Reduced Form)

Variable	*Coefficient*	t-*Statistic*
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
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's Day match	0.138	4.42
Current quality		
Number of wins in the past 3 games (h)	0.010	0.97
Score past game (h)	–0.001	0.23
Goals past game at home (h)	0.004	0.79
Standings (h)	–0.012	4.34
No defeat in past 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	

NOTE: h = home team; v = away team. Endogenous variable is log(price). The distance variable is measured in thousands of kilometers. The uncertainty for the title variable is measured in thousands.

NOTES

1. See Schofield (1983); Cairns, Jennett, and Sloane (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. Evidence for the English Football League from Simmons (1996), using season data for individual clubs, suggests that "casual" spectators are more price-sensitive than season ticket holders.

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. On the other hand, this is almost equivalent to the use of a constructed average ticket price using each class of seat's share of stadium capacity as proposed by Coffin (1996). This is because there is almost proportionality among the prices of the different types of seats, and their share of stadium capacity has not changed substantially in the period under consideration.

6. As far as we know, Falter and Pérignon (2000) is the only article 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. Most of the clubs have more than three different prices during the season. In fact, the proportion of the total price variation in each season due to within-club variation represented 51.4% in the 1992-93 season, 53.9% in 1993-94, 56.5% in 1994-95, and 53.0% in 1995-96 when considering the cheapest price.

12. Some studies have shown a tendency to control for the unobserved component corresponding to each fixture in a particular season, as in Baimbridge, Cameron, and Dawson (1996) and Carmichael, Millington, and Simmons (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.

13. These estimates would be inconsistent if clubs were capacity constrained. In our sample, only 3.4% of the observations correspond to this situation. Results obtained estimating a Tobit model do not differ significantly from those presented here.

14. This is the usual approach when transforming a dynamic model for panel data previous to its estimation by Instrumental Variables or Generalized Method of Moments.

15. 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 ($F_{4,\,1{,}550} = 2.37$). When fitting a polynomial of fourth order all the coeffcients of the (log)price variables become not significant because of collinearity problems.

16. 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.

17. 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 final 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.

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

19. For U.S. professional football, Welki and Zlatoper (1994) found that games that are blacked out for local television are more poorly attended; for English football, Kuypers (1996) did not find a significant effect for this variable; 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, Cameron, and Dawson (1996) argue that the net effect of television on attendance is indeterminate.

20. Note that we cannot 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.

21. The *F* statistic for testing the null hypothesis of not including these controls is 13.16, rejecting this null at a 5% significance level ($F_{33,\,1,518} = 1.46$).

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

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

24. This transformation took place in 1992 and all the clubs were involved with the exception of Athletic of Bilbao, Barcelona, Osasuna, and Real Madrid. Before 1992, all football clubs were nonprofit associations.

REFERENCES

Amemiya, T. (1983). Nonlinear regression models. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics, volume 1* (pp. 333-389). Amsterdam: North-Holland.

Baimbridge, M., Cameron, S., & 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., & 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., & Lye, J. (1992). Attendance at Australian rules football: A panel study. *Applied Economics*, *24*, 1053-1058.

Boyd, D. W., & 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., & Eaton, J. W. (1996). Rebuilding attendance in Major League Baseball: The demand for individual games. In J. L. Fizel, E. Gustafson, & L. Hadley (Eds.), *Baseball economics: Current research* (pp. 9-31). Westport, CT: Praeger.

Cairns, J. A. (1988). *Uncertainty of outcome and the demand for football* (Discussion Paper No. 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., & 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., & Simmons, R. (1999). Elasticity of demand for rugby league attendance and the impact of BskyB, *Applied Economics Letters*, *6*, 797-800.

Coffin, D. A. (1996). If you build it, will they come? Attendance and new stadium construction. In J. L. Fizel, E. Gustafson, & L. Hadley (Eds.), *Baseball economics: Current research* (pp. 33-46). Westport, CT: Praeger.

Davidson, R., & 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, MA: Lexington Books.

Downward, P., & Dawson, A. (2000). *The economics of professional team sports*. London: Routledge.

El Hodiri, M., & Quirk, J. (1971). An economic model of a professional sports league. *Journal of Political Economy*, *79*, 1302-1319.

Falter, J. M., & 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., & Le Dressay, A. (1991). The pricing of sports events: Do teams maximize profit? *Journal of Industrial Economics*, *39*, 297-310.

Forrest, D., & Simmons, R. (2001). *Outcome uncertainty and attendance demand in sport: The case of English soccer.* Retrieved from www.aems.salford.ac.uk/salgroup

Fort, R. (2000). European and North American sports differences (?). *Scottish Journal of Political Economy*, *47*, 431-455.

Heilmann, L., & Wendling, W. R. (1976). A note on optimum pricing strategies for sport events. In R. E. Machol & S. P. Ladany (Eds.), *Management science in sports* (pp. 91-99). Amsterdam: North-Holland.

Hoehn, T., & 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., & 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., & Haupert, 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 No. 96-01). University College London.

Marburger, D. R. (1997). Optimal ticket pricing for performance goods. *Managerial and Decision Economics*, *18*, 375-381.

Peel, D., & 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., & 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* [Economic analysis of attendance in Spanish football]. Doctoral dissertation, Universidad de Oviedo, Spain.

Salant, D. J. (1992). Price setting in professional team sports. In P. M. Sommers (Ed.), *Diamonds are forever: The business of baseball* (pp. 77-90). Washington, DC: The Brookings Institution.

Sarmiento, M. (1994). *Selección española de fútbol: Sus hombres, uno a uno* [Spanish national football team: Their men one by one]. Madrid, Spain: Fundación del Fútbol Profesional.

Schofield, J. A. (1983). Performance and attendance at professional team sports. *Journal of Sports Behaviour*, *6*, 196-206.

Simmons, R. (1996). The demand for English League Football: A club-level analysis. *Applied Economics*, *28*, 139-155.

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., & Blundell, R. (1984). An exogeneity test for a simultaneous equation Tobit model with an application to labour supply. *Econometrica*, *54*, 679-685.
Szymanski, S., & Kuypers, T. (1999). *Winners and losers*. London: Viking.
Szymanski, S., & Smith, R. (1997). The English football industry: Profit, performance and industrial structure. *International Review of Applied Economics*, *11*, 135-153.
Welki, A. M., & 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., & Sim, B. (1995). The demand for semi-pro league football in Malaysia 1989-91: A panel data approach. *Applied Economics*, *27*, 131-138.

Jaume García is a professor of economics at the Universitat Pompeu Fabra in Barcelona, Spain. He received his undergraduate degree from the University of Barcelona and his M. Sc. (Econometrics and Mathematical Economics) and Ph. D. (Economics) degrees from the London School of Economics and Political Science. His fields of specialization are microeconometrics, labour economics, health economics, and sports economics.

Plácido Rodríguez is a lecturer in economics at the Universidad de Oviedo in Salamanca, Spain. He is a doctor of economics and licentiate on laws. He has served as President of Real Sporting de Grijón, a First Division soccer club, and is a member of AIES.