ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 01/2007

Identification of Segments of Soccer Clubs in the Spanish League First Division with a Latent Class Model

Carlos Barros, Julio del Corral y Pedro Garcia-del-Barrio

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE ECONOMÍA

PERMANENT SEMINAR ON EFFICIENCY AND PRODUCTIVITY

IDENTIFICATION OF SEGMENTS OF SOCCER CLUBS IN THE SPANISH LEAGUE FIRST DIVISION WITH A LATENT CLASS MODEL

Carlos P. Barros[♣], Julio del Corral[♦] and Pedro Garcia-del-Barrio[♠]

Efficiency Series Paper 03/2007

Abstract: This paper identifies different groups in a cost function framework of soccer clubs in the Spanish Football League First Division. In particular, we have clustered the sample - comprising data for seasons 1994/95 to 2004/05 - into three groups. To do so, we have implemented a stochastic frontier latent class model, a procedure that also permits us to analyze also the efficiency of the clubs with respect to their own frontiers. The results reveal that some of the clubs could improve their efficiency levels substantially.

Key words: cost efficiency, latent class model, soccer, stochastic frontier.

[♣] ISEG – School of Economics and Management, Technical University of Lisbon, Portugal.
e-mail: cbarros@iseg.utl.pt

[♦] Departamento de Economía, Universidad de Oviedo, Spain.

[♠] Universitat Internacional de Catalunya, Spain.

1. Introduction

Efficiency in sports is a subject that has attracted a lot research in the past. The literature have followed two main approaches, namely, Data Envelopment Analysis (DEA)[1], which is a non-parametric technique (Fizel and D'Itry, 1997; Barros and Santos 2003; Haas, 2003A, 2003B; Espitia-Escuer and García-Cebrian, 2004; Barros and Leach, 2006A), and econometric models that are either stochastic frontier or deterministic[2] (Scully, 1994; Hofler and Payne, 1997; Dawson, Dobson and Gerrard, 2000A, 2000B; Carmichael, Thomas and Ward, 2001; Kahane, 2005; Barros and Leach, 2006B, 2007). All these papers assume that the technology is the same for all units. Hence, if a unit uses a different technology, the efficiency level calculation is biased, since the efficiency level is sensitive to the estimated technology.

The motivation for the present research is based on the hypothesis that some heterogeneity could arise in the production process of translating inputs into output (i.e., winning performance) among clubs, due to different unobserved endowments in the Spanish soccer league. By unobserved endowments, we mean unobserved variables such as the attitude of fans and the media towards the team, and the club's historical sporting achievements, traditions, etc. It is important to be aware that only unobserved endowments can give rise to different parametric differences. This hypothesis is tested using a latent class model (Orea and Kumbhakar, 2004; Greene, 2005). These models assume that there is a finite number of structures (classes) underlying the data and each unit can be assigned to a particular group, using the estimated probabilities of class membership. Thus, they are able to estimate different technologies within a sample. Moreover, the number of groups can be tested using statistical criteria.

Thus, the aim of this research is twofold. First, we cluster Spanish soccer clubs competing in the first division of the league by segments[3]; second, we study the

[1] See Charnes et al. (1995) for details.

[2] See Kumbhakar and Lovell (2000) or Greene (2007) for good overviews.

[3] An alternative procedure to split the teams into several groups is cluster analysis. Cluster analysis aims both to maximize the variance between groups and minimize the variance within groups according to some variables. In contrast, the latent class models not only can use the differences in some variables but it also uses the goodness of fit to the estimated technologies. Hence, the latent class model benefits from richer information than traditional clustering methods.

efficiency of these clubs in a cost function framework. In this regard, it is important to examine the bias produced when the assumption that all clubs use the same technology is incorrect. To this end, we use the stochastic frontier latent class model in line with Greene (2005) and we compare this with the standard stochastic frontier model. The scope of the present research involves combining both sports and financial variables in order to evaluate efficiency in the soccer industry. While sporting achievement is easily observable on the field, financial outcomes (as reported in the financial accounts of the clubs) have no such transparency. Moreover, not all sports clubs make their financial statements public. This is indeed the case in the Spanish Football League, for which Deloitte and Touche no longer publishes its annual Spanish football financial report, owing to the lack of available information.

The rest of the paper is organized as follows. In Section 1, we review the related literature, followed in Section 2 by a description of the contextual setting. Sections 3 and 4 describe the sporting cost function and the model respectively. Section 5 presents the data and the results. Finally, Section 6 provides our conclusions and some policy implications.

2. Literature review

The analysis of efficiency is a major issue in the domain of sports management (Slack, T., 1997). There are two contemporary approaches to measure efficiency: firstly, the econometric or parametric approach and secondly, the non-parametric approach. We find several papers that have used the econometric approach to efficiency analysis in soccer. For instance, Dawson, Dobson and Gerrard (2000A) analyzed the managerial efficiency of English soccer, estimating several production frontiers. They used as output the winning percentage and as input several measures of player quality. Dawson, Dobson and Gerrard (2000B), using a similar approach, provided comprehensive empirical evidence on the robustness of estimates of coaching efficiency in English soccer to the full range of robustness available estimation methods, as well as alternative input-output[4] specifications. Carmichael, Thomas and

[4] They used the winning percentage ratio as output. This percentage is computed by the number of points obtained as a proportion of the total potential attainable points. Alternatively, they resorted to the adjusted winning percentage ratio, in which drawn matches are considered as ‘half-wins’.

Ward (2001) analyzed the efficiency of the English Football Association Premiership clubs, using as output the number of points gained during the season. Barros and Leach (2006B, 2007) estimated cost stochastic frontiers for the English Premiership, using as outputs both the points achieved in the season and spectator attendance. Ascari and Gagnepain (2007) estimated average wage equations in Spanish soccer, using a stochastic frontier model to evaluate empirically the consequences of the rent-seeking behavior of clubs on their costs.

Looking beyond soccer, several papers have used the econometric approach to efficiency analysis in sports economics. For example, Zak, Huang and Siegfried (1979) explored the production efficiency in the National Basketball Association, using a deterministic frontier. Porter and Scully (1982) studied the managerial efficiency of baseball managers with a deterministic approach. Scully (1994) showed that coaching tenure was related to managerial efficiency in basketball, baseball and football, using survival analysis. He had previously calculated the managerial efficiency, using deterministic as well as stochastic frontier models. Hofler and Payne (1997) applied a stochastic frontier model to the NBA, using the number of wins as output. Finally, Kahane (2005) investigated the relationship between inefficiency and discriminatory hiring practices in the National Hockey League (NHL), using a stochastic frontier model. He used as output the proportion of possible points gained in the regular season.

Among the papers adopting a non-parametric approach, we first mention Espitia-Escuer and García-Cebrian (2004), in which the efficiency of Spanish First Division soccer clubs is decomposed, by means of DEA, into technical efficiency and scale efficiency. They use as output the number of points achieved in the league season. Fizel and D'Itri (1997) applied the DEA technique to measure the managerial efficiency in college basketball. They use the winning percentage as output and *ex-ante* quality measures of players as inputs. Barros (2003) analyzes the incentive regulation on sports organisational training activities, disentangling technical and allocative efficiency with DEA techniques. Finally, Haas (2003A) examines the efficiency of the US Major Soccer League with DEA, while Barros and Santos (2003) estimate Malmquist indexes for Portuguese sports organizational training activities.

The above survey of the previous research indicates that, to our knowledge, there is no published paper diverging from the assumption that the technology is the same for all individual clubs. Furthermore, it seems that the employment of the number of points obtained by the team during the season as output is a common procedure in the related literature.

3. Contextual setting

The Spanish Football League First Division has become one of the most important soccer leagues in the world. Many of the world's best players participate in it and, as a result, Spanish clubs have an outstanding record of success in the European club competitions.[5] Given the importance of this league, it is only natural that several papers have used data from the Spanish Professional Football League (LFP). For instance, García and Rodríguez (2002) explored the determinants of attendance in the Spanish league; Espitia-Escuer and García-Cebrian (2004) studied the efficiency of the First Division clubs, using DEA. Other studies on Spanish soccer include Ascari and Gagnepain (2006), who analyzed the sport's financial crisis, and Garcia-del-Barrio and Pujol (2007), who explored the relationship between pay and performance and the role of superstar players in such winner-takes-all markets as the Spanish soccer league.

The Spanish First Division is characterized by the fact that many of the top players, in addition to the largest attendances, are concentrated in only two teams (*Real Madrid* and *F.C. Barcelona*), which regularly occupy the first two positions in the league table. Table 1 displays some descriptive statistics on financial and sport issues of the Spanish soccer clubs in our sample.

Table 1. Team statistics (averages for seasons 1994/95 to 2004/05)

Teams	Number of available Seasons	Position	Wages	Revenues
		(1)	(2)	(3)
Real Madrid	9	3	89,627,500	138,318,000

[5] *F.C.Barcelona* won the European Champions League in season 2005/06 while *Sevilla C.F.* won consecutively the UEFA Cup in 2005/06 and 2006/07. Furthermore, in 2006/07, the UEFA Cup final was played by two Spanish teams: *Sevilla C.F* and *R.C.D. Español*.

F.C. Barcelona	8	2.8	62,858,200	113,313,000
Valencia C.F.	8	4.9	31,809,900	52,683,200
Atlético de Madrid	6	10.3	27,922,400	43,897,200
Deportivo de la Coruña	6	6	21,920,400	32,206,700
Athletic de Bilbao	8	7.6	22,791,400	32,102,000
Real Sociedad	8	9.6	14,657,100	26,139,100
Sevilla F.C.	5	13	13,835,600	22,273,200
Real Betis Balompié	8	9	7,385,100	22,079,200
R.C.D. Mallorca	6	7.7	14,455,700	20,515,000
R.C.D. Español	9	11.4	12,386,200	20,399,500
Celta de Vigo	8	7.5	7,562,840	19,761,400
Villarreal C.F.	4	13.8	8,472,160	19,276,300
C.D. Tenerife	5	13.6	10,666,900	19,137,400
Club At. de Osasuna	2	14	14,487,700	19,047,500
Real Zaragoza C.D.	9	12.1	13,220,600	18,467,100
Málaga C.F.	5	10.4	15,160,700	16,262,900
C.D. Alavés	5	11.6	12,307,700	15,704,100
Real Valladolid	4	9.5	9,992,470	12,097,100
Real Oviedo	7	15.3	5,823,570	9,915,570
Racing de Santander	7	15	8,475,120	9,849,230
U.D. Salamanca	2	17.5	5,665,860	9,751,820
Rayo Vallecano	3	13	5,198,230	8,825,070
S.D. Compostela	3	12.7	4,937,870	7,654,700
C.D. Numancia	2	18.5	4,527,860	6,920,900
Albacete Balompié	2	19	3,747,040	5,105,310

SOURCE: Authors' elaboration.

The averages are computed for the N seasons in which data was available for teams in the First Division League.

(1): Average final league position.

(2) and (3): measured in euros, year 2000.

From the information presented in Table 1, we can conclude that clubs displaying the highest expenditures are normally ranked in the leading positions. Similarly, clubs in the middle of the table have an average expenditure, while the lower clubs are those with the lowest levels of expenditure.

4. Sporting cost functions

In production economics, the production process is usually analysed by using a dual approach (i.e., cost functions or profit functions). The assumption underlying cost functions is that the units have a minimizing cost behavior.

Thus, a cost function is specified as:

$$C = C^{*}(w, y, t) \tag{1}$$

where C is cost, w are input prices, y is output and t represents the state of the technology. As expressed by Carmichael, Thomas and Ward (2001), sports team output is conventionally measured in terms of team success, reflected by the winning performance in individual matches or during a season's competition. As documented in the literature review, the number of points achieved in the league is the most common form of measuring the level of output in the production process. Concerning input prices, researchers clearly agree in the literature that the main input of team sports is the quality of the players. Nevertheless, there is much discussion as to how such quality should be measured (e.g., Gerrard, 2001; Dawson, Dobson and Gerrard, 2000B). In this regard, using a cost function framework is simpler and more appropriate than the production function framework, since it must employ the input price.

A simple procedure to construct the player's price is to divide players' total wages by the number of players. Here arises another issue, which is how to determine the number of players. In Spain, the General Rules of the Royal Spanish Soccer Federation establishes a maximum number of 25 players allowed to be registered by the clubs, a rule affecting all the teams in our sample.[6] Accordingly, this maximum can serve well to represent the number of players, since it is unlikely that clubs employ fewer players than allowed. On one hand, clubs may pick young, home-grown players (from junior teams) during the course of the season. If the young player is selected for several matches, he is likely to become a permanent member of the first-team squad in the following season. It is also customary to give young players their début in end-of-season matches, if the outcome of the match is not vitally important.

On the other hand, many clubs will buy and/or sell players during the permitted times of the year, thereby implying that other alternatives for the cost function (such as using the total number of players who have played in matches during the season) present greater problems. Consequently, we advocate here that the maximum possible number of players registered by the clubs is the most feasible criterion.

[6] Cf.: General Rules of the Royal Spanish Soccer Federation (Royal Spanish Soccer Federation. February 2007 edition), Book XI: Article 96. 1. Teams in first and second football league can register the maximum of 25 players in their team. Non registered players are not allowed to play in the season.

Finally, in cost functions, the state of the technology is often included. In contrast to other industries, there is a maximum output that can be achieved by all decision-making units (i.e., total points in dispute). Therefore, the clubs in the sample as a whole are not able to increase the output. Another noteworthy issue is the fact that the years of our sample coincide with a highly substantial increase in the salaries paid to players. Given that in order to achieve more output, *ceteris paribus*, the cost needs to be increased, we expect the coefficient of the time trend to be positive.

5. A stochastic frontier latent class model

In this paper, we adopt the stochastic frontier approach, which came to prominence in the late-1970s as a result of the work of Aigner, Lovell and Schmidt (1977), Battese and Corra (1977) and Meeusen and van den Broeck (1977). In this framework, it is assumed that the residuals have two components (noise and inefficiency). Thus, a stochastic cost function may be written as:

$$ln\, C_{it} = f\left(x_{it}\right) + \varepsilon_{it}\,;\ \varepsilon_{it} = v_{it} + u_{it} \qquad (2)$$

where C_{it} represents the cost of the decision-unit *i* under analysis in the *t*-th period, *f(x)* is the functional form, x_{it} is a vector of variables including input prices and output quantity, and ε is the error term which is composed by two components. The symmetric component, *v*, captures statistical noise and is assumed to follow a distribution centered at zero, while *u* is a non-negative term that reflects inefficiency and is assumed to follow a one-sided distribution (i.e. truncated normal, half-normal, exponential). Moreover, the two components *v* and *u* are assumed to be independent of each other.

Since the estimation procedure of equation (2) yields merely the residual ε, rather than the inefficiency term *u,* the latter must be calculated indirectly, using the Jondrow et al. (1982) formula, which is the conditional expectation of u_{it}, conditioned on the realized value of ε_{it}.

Following Greene (2001), we can write equation (2) as a latent class model:

$$ln C_{it}\big|_j = f(x_{it})\big|_j + v_{it}\big|_j + u_{it}\big|_j \tag{3}$$

where subscript *i* denotes the firm, *t* indicates time and *j* represents the different classes. The vertical bar means that there is a different model for each class *j.* It is assumed that each club belongs to the same group in all periods[7].

An important issue in these models is how to determine the number of classes. The usual procedure is to estimate several models with different numbers of groups and then use a statistical test in order to choose the preferred model. Greene (2005) proposed testing 'down', where beginning from a *J** known to be at least as large as the true *J*, one can test down, given that the *J-1* class model is nested with the *J* class model imposing $\theta_j = \theta_{j-1}$, based on likelihood ratio tests[8]. An alternative is to use information criteria, such as the Akaike Information Criterion (AIC) or the Schwarz Bayesian Information Criterion (SBIC). These statistics are calculated using the following expressions:

$$SBIC = -2 \cdot \log LF(J) + \log(n) \cdot m \tag{4}$$

$$AIC = -2 \cdot \log LF(J) + 2 \cdot m \tag{5}$$

where *LF(J)* is the value that the likelihood function takes for *J* groups, *m* is the number of parameters used in the model and *n* is the number of observations. The favored model will be that for which the value of the statistic is lowest.

6. Data and results

Since the early 1990s, most of the Spanish clubs have adopted corporate status (known as SAD, or Sociedad Anónima Deportiva), abandoning their previous non-profit

[7] Further technical details on the estimation procedure are provided in the appendix.

[8] The statistic is constructed as $-2 \cdot (\log LF_r - \log LF_u)$, where LF_r and LF_u are the log-likelihood functions evaluated at the restricted and unrestricted estimates. The statistic under the null hypothesis, $\theta_j = \theta_{j-1}$, follows a chi-squared distribution with degrees of freedom equal to the number of restrictions being tested. If the null hypothesis is rejected, the preferred model is that with J classes.

status and thereby being obliged to publish their financial statements regularly. Nevertheless, the task of gathering our data panel was not straightforward. Some clubs did not publish the required information, whereas four other clubs (*F.C. Barcelona*, *Real Madrid*, *Athletic de Bilbao* and *Osasuna*) have retained club status and do not have any legal obligation to disclose their accounts to the general public. To estimate the cost function, we have used an unbalanced panel of the Spanish clubs competing in the First Division in seasons 1995/96 to 2004/05.[9]

The specification of the cost function follows microeconomic theory (Varian, 1987). The costs are regressed in input prices and output descriptors. With regard to output, despite the fact that other output measures (e.g. league standing, winning percentage) could be employed, we opt to use the number of points achieved in the league. In relation to inputs, as previously explained, the labor price is constructed by dividing total wages by the number of players[10]. On the other hand, the literature acknowledges that not only the labor price, but also at least the capital price must usually be included in the calculations. Accordingly, we also include a proxy for capital price, which is obtained by dividing the amortizations by the value of the total assets, Barros and Leach (2006B, 2007), Kraft, Hofler and Payne (2006).

Table 2 shows the descriptive statistics of the data used in the estimation. The costs are expressed in euros and include personnel expenses, depreciation and amortization, and other operational expenses.

Table 2. Descriptive statistics of the data

Variable	Mean	Standard deviation	Minimum	Maximum
Cost	44,451,600	56,867,600	3,036,630	471,052,000

[9] In almost all cases, the information was directly obtained from the clubs' accounts. In seasons 1996-1997 and 1997-1998, data of some clubs (*Athletic de Bilbao*, *Barcelona*, *Real Valladolid*, *Real Sociedad*, *Atlético de Madrid* and *Sevilla F.C.*) was obtained from Deloitte and Touche, Review of Spanish Football (1996/7 and 1998/9). Angel Barajas kindly provided us with the data for *F.C. Barcelona* and *Athletic de Bilbao* in seasons 2000-2001, 2001-2002 and 2002-2003; and for *Real Madrid* in the season 2000-2001.

[10] We are not able to disentangle between players' salaries and other staff salaries. However, it should be noted that by far the largest proportion of personnel expenses corresponds to players' earnings. Therefore, we believe that standardizing by the number of players is a good procedure to construct a proxy variable for the price of the input quality of players. We use as the number of players the maximum permitted number of players; that is, 25.

Points	54	12	27	90
Labor price	837,714	1,010,950	67,101	5,566,800
Capital price	0.18	0.16	0.01	1.49

Note: The monetary variables are expressed in constant euros, year 2000

The empirical specification of the cost function is the translog. We have chosen a flexible functional form in order to avoid imposing unnecessary *a priori* restrictions on the technologies to be estimated. Each explanatory variable is divided by its geometric mean. In this way, the translog can be considered as an approximation to an unknown function and the first order coefficients can be interpreted as the production elasticities evaluated at the sample geometric mean. We also include a time trend as well as a squared time trend in order to obtain some temporal changes. The equation to estimate is:

$$\begin{aligned} ln\left(\frac{Cost_{it}}{y_{it}}\right) = \beta_0\big|_j + \beta_y\big|_j \, ln\, y_{it} + \sum_l \beta_{wl}\big|_j \, ln\, w_{ilt} + \frac{1}{2}\beta_{yy}\big|_j \, ln(y_{it})^2 + \\ \frac{1}{2}\sum_l \beta_{wwl}\big|_j \, ln\, w_{ilt}^2 + \sum_l \beta_{wly}\big|_j \, ln\, w_{ilt} \, ln\, y_{it} + \delta_t\big|_j \cdot t + \delta_{tt}\big|_j \cdot t^2 + v_{it}\big|_j + u_i\big|_j \end{aligned} \quad (6)$$

where *y* is the output measured as points, *w* denotes input price, *t* is a time trend, *v* is a random error which reflects the statistical noise and is assumed to follow a normal distribution centered at zero, while *u* reflects inefficiency and is assumed to follow a half-normal distribution. In order to test the number of groups we used the testing-down procedure proposed by Greene (2005), where beginning from a *J** known to be at least as large as the true *J*, one can test down, given that the *J-1* class model is nested with the *J* class model imposing $\theta_j = \theta_{j-1}$, based on likelihood ratio tests, as well as the SBIC and AIC. The model with four groups does not converge, likewise the model with two groups is rejected against the model with three groups using the three criteria and therefore, we estimate the model with three groups[11]. Moreover, since none of the variables used as separate variables was significant, we decided to adopt the model in which none of them was included. Following this, the latent class model uses the goodness of fit to create the groups.

[11] The program used to estimate the model was Limdep 9.0.

First we present the groups formed by the latent class model using the posterior probabilities of class membership in Table 3. Subsequently, we present the results of the stochastic frontier latent class model in Table 4. Similarly, we have also estimated a standard stochastic frontier with panel data which assumes that the technology is the same for the entire sample, in order to compare it with the latent class model.

Table 3. Group composition in alphabetical order

	Group 1	Group 2	Group 3
	Atlético de Madrid	Albacete Balompié	C.D.Numancia
	C.D. Tenerife	Athletic de Bilbao	Club Atlético de Osasuna
	Celta de Vigo	C.D. Alavés	Málaga C.F.
	Deportivo de la Coruña	S.D. Compostela	Rayo Vallecano
	F.C. Barcelona		Real Valladolid
	R.C.D. Español		
	R.C.D. Mallorca		
	Racing de Santander		
	Real Betis Balompié		
	Real Madrid		
	Real Oviedo		
	Real Sociedad		
	Real Zaragoza C.D.		
	Sevilla F.C.		
	U.D. Salamanca		
	Valencia C.F.		
	Villarreal C.F.		
Observations	115	18	16

We can observe that the most important clubs appear in the first group (i.e. *Real Madrid*, *F.C. Bar*celona, *Valencia*, *Deportivo de la Coruña*, *Atlético de Madrid*). The second group consists of clubs employing many home-grown players who were nurtured in the junior teams. Finally, the third group comprises clubs enduring low-budget constraints.

Table 4. Estimation results

	Standard SF.	Latent class model		
		Group 1	Group 2	Group3
Constant	12.44*** (0.12)	12.21*** (0.04)	12.91*** (0.06)	11.58*** (0.15)
Points	-0.90*** (0.25)	-0.76*** (0.08)	-0.71*** (0.04)	-0.71 (0.53)
Labor price	0.86*** (0.07)	0.81*** (0.02)	1.04*** (0.02)	0.68*** (0.08)

Capital price	0.13*** (0.05)	0.15*** (0.03)	0.06*** (0.02)	0.18 (0.15)
Points^2	0.55 (0.95)	0.48*** (0.19)	2.13*** (0.37)	8.31** (3.32)
Labor price^2	0.07 (0.08)	0.23*** (0.02)	-0.03 (0.07)	0.29** (0.12)
Capital price^2	0.04 (0.06)	0.01 (0.03)	0.14*** (0.02)	-0.01 (0.22)
Points*Labor price	-0.16 (0.19)	-0.39*** (0.06)	-0.17* (0.10)	-0.21 (0.38)
Points*Capital price	-0.02 (0.13)	-0.09 (0.12)	-0.08* (0.05)	0.51 (0.88)
Labor price*Capital price	-0.03 (0.03)	-0.02 (0.04)	-0.07** (0.03)	-0.15 (0.14)
Trend	0.17*** (0.04)	0.30*** (0.03)	0.00 (0.02)	0.34*** (0.05)
Squared trend	-0.01*** (0.00)	-0.02*** (0.00)	0.00 (0.00)	-0.02*** (0.00)
$\sigma = \left(\sigma_v^2 + \sigma_u^2\right)^{1/2}$	2.29 (2.11)	0.28 (0.02)	0.03*** (0.01)	0.11*** (0.02)
$\lambda = \sigma_u / \sigma_v$	0.32** (0.18)	248.98 (219.42)	168.31 (292.96)	231.02 (530.53)
Log Likelihood Function	51		112	
Observations	149		149	

Note: * ,**,*** indicate significance at the 10%, 5% and 1% levels, respectively.

Standard errors are shown in parenthesis.

The estimated coefficients have the expected signs, as all price elasticities are positive. It is noteworthy that we have tested the homogeneity in prices, rather than impose it. The results show that only the homogeneity in prices for the second group is rejected. An important result that supports the latent class model estimation is that the differences of the input prices among groups are significant, with the only difference in capital price appearing in Group 3. Similarly, it can be observed that the group with the highest labor elasticity is Group 2. The costs increase with the trend, but at a decreasing rate, for Groups 1 and 3. Another interesting outcome is that the estimated coefficients of the standard stochastic frontier are between the minimum and the maximum of the latent class model coefficients, with the sole exception of the output measure. This result suggests that to some extent, the standard stochastic frontier estimates an average of the latent class model technologies.

Finally, Table 5 shows the cost average cost efficiency for each team across seasons. The cost efficiency is defined as the ratio between the minimum cost and the actual cost, and takes values between 0 and 1. According to this definition, the closer to 1 is the efficiency measure, the more efficient must the club be considered. Given that the dependent variable is expressed in logarithms, it was calculated as:

$$EC = exp(-\hat{u}) \quad (7)$$

where the estimated value of the inefficiency ($\hat{u}$) is separated from the random error term ($\hat{v}$) using the Jondrow et al. (1982) formula.

Table 5. Average cost efficiency across seasons

Team	Standard SF	Latent Class Model	Difference
Atlético de Madrid	0.56	0.66	0.10
Real Betis Balompié	0.55	0.68	0.13
Villarreal C.F.	0.62	0.72	0.10
Celta de Vigo	0.58	0.73	0.15
U.D. Salamanca	0.71	0.76	0.05
C.D. Tenerife	0.69	0.77	0.08
Racing de Santander	0.74	0.82	0.08
F.C. Barcelona	0.67	0.83	0.16
Valencia C.F.	0.70	0.83	0.13
R.C.D. Mallorca	0.72	0.85	0.13
Real Zaragoza C.D.	0.76	0.85	0.09
R.C.D. Español	0.77	0.87	0.10
Real Madrid	0.70	0.87	0.17
Real Sociedad	0.75	0.87	0.12
Real Oviedo	0.78	0.89	0.11
Real Valladolid	0.94	0.89	-0.05
Deportivo de la Coruña	0.83	0.90	0.07
Sevilla F.C.	0.86	0.93	0.07
Málaga C.F.	0.90	0.94	0.04
Albacete Balompié	0.93	0.96	0.03
C.D. Alavés	0.97	0.97	0.00
Athletic de Bilbao	0.92	0.99	0.07
Club Atlético de Osasuna	0.85	0.99	0.14
Rayo Vallecano	0.77	0.99	0.22
S.D. Compostela	0.96	0.99	0.03
C.D. Numancia	0.85	1.00	0.15
Total	0.75	0.85	0.10

The clubs with the lowest cost efficiencies are *Atlético de Madrid* and *Real Betis*. It should be emphasized that these two clubs were in the hands of two of Spain's richest men throughout the period analyzed: Jesus Gil, who owned *Atlético de Madrid* and Manuel Ruiz de Lopera who was, and still is, the president of *Real Betis Balompié.* In both cases, the concentration of control over the decision-making within the club was extremely high, which may easily result in the fact that they considered their clubs' expenses to be of much less importance than their sporting achievements. Despite this, the league success achieved by these two clubs was in fact minimal, in return for the amounts of their corresponding expenses, which could be due to the singular character of the owners. This would imply that the finding in relation to low cost efficiency is reliable. On the other hand, the club with the highest cost efficiency in the latent class model is *C.D. Numancia*. This club, which was present in only two of the analyzed seasons, had a very low budget. In the two seasons in question, *C.D. Numancia* won more than 40 points, which constitutes a very respectable achievement, in view of its budget. Hence, it can also be considered a reliable finding.

Concerning the comparison between the standard stochastic frontier and the latent class model efficiencies, it is worth indicating that only *Real Valladolid* presents a greater cost efficiency in the standard model than in the latent class model. The average difference is 0.10, which is rather significant difference. Likewise, the differences are not consistent for all teams, implying that the latent class model efficiencies are not merely a constant added to the standard stochastic frontier model.

7. Conclusions

This article has proposed a simple framework for the comparative evaluation of Spanish First Division soccer clubs and the rationalization of their operational activities. The analysis was conducted by means of the implementation of a stochastic frontier latent class model that, in determining the relative efficiencies, allows the incorporation of a broad variety of inputs and outputs while permitting researchers to account for segments in the sample and the existence of heterogeneity in the data. The main limitation of this paper derives from the data set, since the available data span is relatively short.

This study has drawn attention to the identification of segments among clubs, suggesting that business strategies should be defined for each segment, in such a way that such strategies are adapted to the characteristics of the clubs. In any event, in order to offer more conclusive policy prescriptions, a larger data set would be required. Indeed, the limitations of the present paper suggest directions for new research. Additional research is needed to confirm the results of this paper, as well as to clarify some of the issues identified here. Research that takes into account the presence of heterogeneity should also be expanded to examine sports leagues in other countries.

References

Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21-37.

Ascari, G. and Gagnepain, P. (2006). Spanish Football. *Journal of Sports Economics*, 7, 1, 76-89.

Ascari, G. and Gagnepain, P. (2007). Evaluating rent dissipation in the Spanish football industry. *Journal of Sports Economics*, 8, 5, 468-490.

Barros, C.P. (2003). Incentive regulation and efficiency in sports organisational training activities. *Sport Management Review*, 6, 1, 33-52.

Barros, C.P. and Santos, A. (2003). Productivity in sports organisational training activities: A DEA study. *European Journal of Sport Management Quarterly*, 1, 46-65.

Barros, C.P. and Leach, S. (2006A). Performance evaluation of the English premier football league with Data Envelopment Analysis. *Applied Economics*, 38, 12, 1449-1458.

Barros, C.P. and Leach, S. (2006B). Analyzing the performance of the English Football League with an econometric frontier model. *Journal of Sports Economics*, 7, 4, 391-409.

Barros, C.P. and Leach, S. (2007). Technical efficiency in the English Football Association Premier League with a stochastic cost frontier. *Applied Economics Letters*, 14, 731-741.

Battese, G.E. and Corra, G.S. (1977). Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics*, 21, 169-179.

Carmichael, F., Thomas, D. and Ward, R. (2001). Production and efficiency in Association Football. *Journal of Sports Economics*, 2, 3:228-243.

Charnes, A., Cooper, W., Lewin, A. and Seiford, L. (1995). *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston.

Dawson, P., Dobson, S. and Gerrard, B. (2000A). Stochastic frontier and the temporal structure of managerial efficiency in English soccer. *Journal of Sports Economics*, 1, 341-362.

Dawson, P., Dobson, S. and Gerrard, B. (2000B). Estimating coaching efficiency in professional team sports: evidence from English Association Football. *Scottish Journal of Political Economy*, 47, 4, 399-421.

Espitia-Escuer, M. and García-Cebrian, L.I. (2004). Measuring the efficiency of Spanish First-Division soccer teams. *Journal of Sports Economics*, 5, 4, 329-346.

Fizel, J.L. and D'Itri, M.P. (1997). Managerial efficiency, managerial succession and organizational performance. *Managerial and Decision Economics*, 18, 295-308.

García, J. and Rodríguez, P. (2002). The determinants of football match attendance revisited: empirical evidence from the Spanish Football League. *Journal of Sports Economics*, 3, 1, 18-38.

García-del-Barrio, P. and Pujol, F. (2007). Hidden monopsony rents in winner-take-all markets. *Managerial and Decision Economics*, 27, 1313-27.

Gerrard, B. (2001). A new approach to measuring player and team quality in professional team sports. European Sport Management Quarterly, 1, 219-233.

Greene, W. (2001). New developments in the estimation of stochastic frontier models with panel data. Efficiency Series Paper 6/2001, Dept. of Economics, Univ. of Oviedo.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126, 269-303.

Greene, W. (2007). The Econometric Approach to Efficiency Analysis, in Fried, H, Lovell, K. and Schmidt, S. *The Measurement of Productive Efficiency and Productivity Growth*, Oxford University Press, New York, Oxford.

Haas, D.J. (2003A). Technical efficiency in the Major League Soccer. *Journal of Sports Economics*, 4, 3, 203- 215.

Haas, D.J. (2003B). Productive efficiency of English Football teams – A Data Envelopment Approach. *Managerial and Decision Economics*, 24, 403-410.

Hofler, R.A. and Payne, J.E. (1997). Measuring efficiency in the National Basketball Association. *Economics Letters*, 55, 293-299.

Jondrow, J., Lovell, C.A.K., Materov, I. and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics,* 19, 233-238.

Kahane, L.H. (2005). Production efficiency and discriminatory hiring practices in the National Hockey League: a stochastic frontier approach. *Review of Industrial Organization*, 27, 1, 47-71.

Kraft, E.; Hofler, R. and Payne, J. (2006) Privatization, foreign bank entry and bank efficiency in Croatia: A form-flexible function stochastic cost frontier analysis. Applied Economics, 38, 2077-2088.

Kumbhakar, S. and Lovell, K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from a Cobb-Douglas production function with composed error. *International Economic Review*, 18, 435-444.

Orea, L. and Kumbhakar, S. (2004). Efficiency measurement using stochastic frontier latent class model. *Empirical Economics*, 29, 169-183.

Porter, P. and Scully, G.W. (1982). Measuring managerial efficiency: the case of baseball. *Southern Economic Journal*, 48, 642-650.

Scully, G.W. (1994). Managerial efficiency and survivability in professional team sports. *Managerial and Decision Economics*, 15, 403-411.

Slack, T. (1997). *Understanding Sport Organizational Change: The Application of Organization Theory*. Champaign, Ill: Human Kinetics.

Varian, H.R. (1987) Intermediate Microeconomics: A Modern Approach, (N.Y.: Norton and Co).

Zak, T.A., Huang, C.J. and Siegfried, J.J. (1979). Production efficiency: the case of professional basketball. *Journal of Business*, 52, 379-92.