

Puntos fijos y ligas europeas de fútbol: discusión de las trayectorias de los equipos de fútbol europeos por la densidad de puntos fijos

Fixed points and European soccer leagues: discussing the trajectories of European soccer teams by the density of fixed points

Paulo Mourao

Departamento de Economía. Universidade do Minho (Portugal)

Resumen: Todos los equipos de fútbol visitan sus clasificaciones finales antes de las últimas jornadas de liga. Este trabajo analiza la trayectoria de los equipos de fútbol europeos hacia sus posiciones finales en las temporadas entre 2004/2005 y 2012/2013. Se estudiaron seis ligas profesionales de fútbol (Inglaterra, Alemania, Italia, Países Bajos, Portugal y España). Utilizando el teorema del punto fijo y regresiones cuantílicas, este artículo concluye que la liga holandesa es la que se caracteriza por el mayor valor para su punto fijo, exhibiendo más cambios de ranking que los otros con el mismo porcentaje de partidos jugados. Las ligas inglesas e italianas fueron identificadas como las ligas en las que una mayor proporción de equipos tendió a estabilizarse antes en las que serían sus posiciones finales en la clasificación.
Palabras clave: Fútbol, Teorema del Punto Fijo, Regresión Cuantílica.

Abstract: All soccer teams visit their final standing ranks before their final journeys. This work analyzes the trajectory of European soccer teams toward their final standing ranks for the seasons between 2004/2005 and 2012/2013. We studied six professional soccer leagues (England, Germany, Italy, the Netherlands, Portugal, and Spain). Referring to the fixed point theorem and to quantile regressions, this article concludes that the Dutch league is the one characterized by the largest value for its fixed point, exhibiting more rank changes than the others with the same percentage of played journeys. The English and Italian leagues were identified as the leagues in which a higher share of teams tended to stabilize at their final standings' ranks at the earliest journeys.

Key-words: Soccer, Fixed Point Theorem, Quantile Regression.

1. Introduction

There is a paradox when we watch a sports match. On one side, we prefer matches with uncertain outcomes. On the other side, we anticipate the final result of the match or the final standing of a professional league's season. Therefore, forecasting sports outcomes can be very important for diverse groups of sports agents, such as gamblers, investors and fans.

All of these agents tend to gain from good forecasts. They gain because they can make better choices, they can bet on the closest result, and they can manage their own expectations, minimizing a large range of costs such as individual (anxiety) or managerial (information, which is very important for investors in sports publicity).

In this work, we are going to investigate how the evolution of European soccer team ranks can predict their final standing ranks. We are going to investigate six major professional leagues (the English Premier League, the German Bundesliga, the Italian Serie A, the Dutch League, the Portuguese Liga Profissional, and the Spanish La Liga). We are going to focus on the study of the share of teams that from the 2004/2005 to the 2012/2013 seasons were at their final standings' ranks after each journey.

Observing this share of teams will allow us to discuss the changes of standings across European leagues with reference

to the theorem of fixed point. Additionally, this study will allow for exploring the empirical potentialities of quantile regressions in order to observe how the heterogeneity of data can enhance detailed empirical results.

The remainder of this paper proceeds as follows. Section 2 identifies the focus of our study, namely, soccer teams' orbits towards their final ranks and their density functions considering their final standings after each journey in a professional league. Section 3 is the empirical section, which discusses the data, estimation methods, and results. Section 4 concludes the paper, revealing policy implications and possibilities for additional work.

2. From soccer teams' orbits to the density functions of their ranks

2.1. Identifying teams' orbits

Currently, for most professional soccer leagues, the competition structure is based on a double round-robin system¹; each victory is awarded 3 points, each draw 1 point, and each loss 0 points. From the first to the last journey in a season, each professional team exhibits a particular evolution in its rank among all of the competing teams in the league.

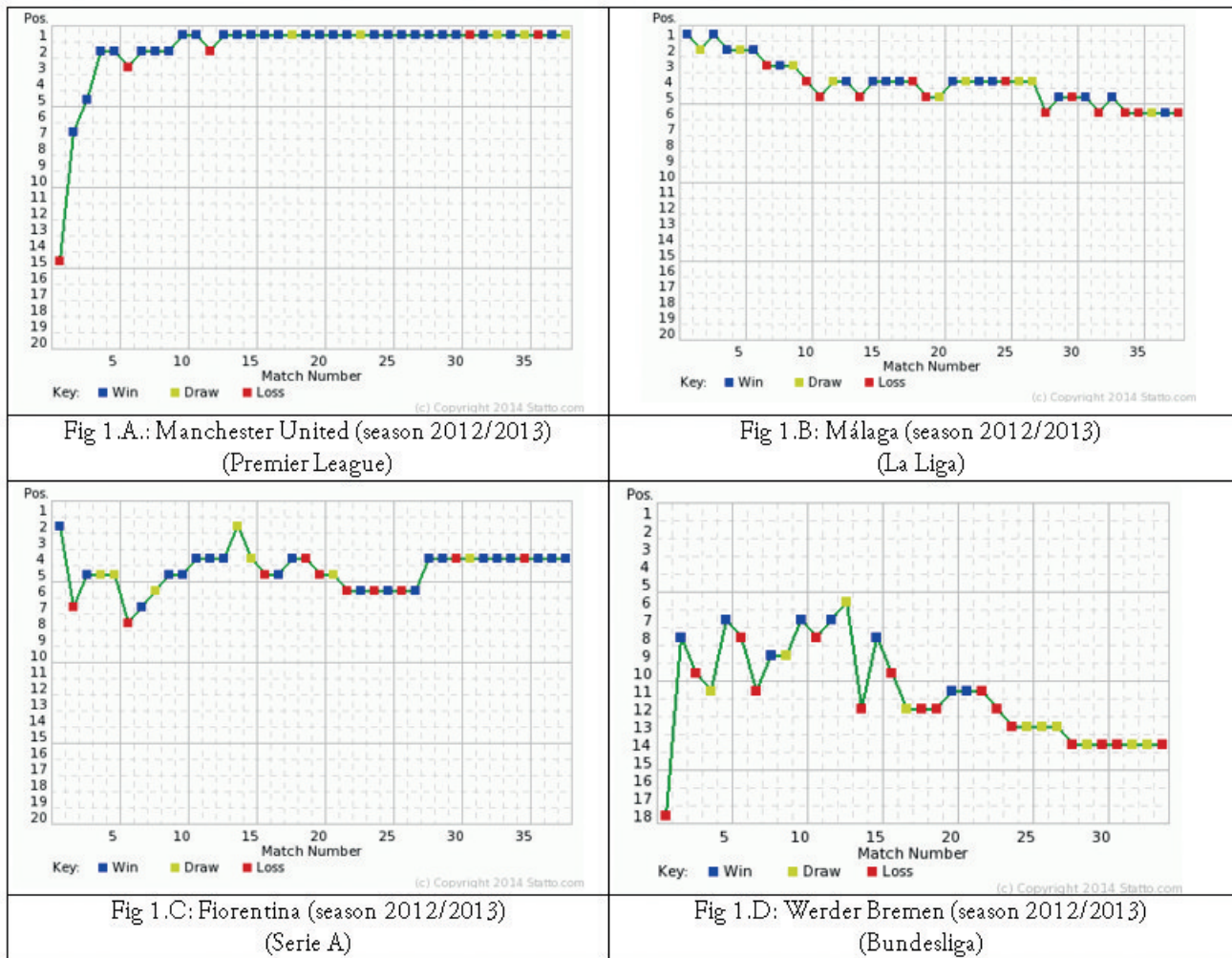
Dirección para correspondencia [Correspondence address]: Paulo Mourao.
E-mail: paulom@eeg.uminho.pt

¹ That is, each team plays two matches against all other competing teams, one at the home stadium and the other at the other team's home stadium.

When we observe the evolution of a randomly chosen professional soccer team's rank from the first journey to the final one, we obtain figures such as Figure 1A (Manchester Uni-

ted), Figure 1B (Málaga), Figure 1C (Fiorentina), or Figure 1D (Werder Bremen)

Figures 1A-1D. League progress of 4 teams at 4 European soccer leagues.



Source: <http://www.statto.com>

These figures suggest the existence of a fixed point at each path that is described by the rank's evolution. Clearly, the first suggestion is that the final standing rank is the fixed rank to which each path converges. Interestingly, we conclude that a 'fixed point' was visited (or occupied) by each team multiple times before the final journey.

Following Marteli (1999) and Druck et al (2001), we can consider this rank evolution as describing each team's orbit towards its final standing.

Let $F : N \rightarrow N$. Therefore, we identify F as a functional iteration that describes the evolution of a soccer team's rank in its respective league after the n -th journey.

That is,

$$F^n(x) = (F \circ F^{n-1})(x) = F(F^{n-1}(x)), \text{ for } n \geq 1$$

Let r_0 identify the rank of a given team. Also assume that $r_0 \in N$, where N is the set of natural numbers. We define the orbit of r_0 under F to be the sequence of points

$$r_0, r_1, r_2, \dots, r_n, \dots, r_E$$

such that $r_{n+1} = F(r_n)$, for $0 \leq n \leq E$, where E is the maximum number of journeys in a season. This also means that $r_{n+1} = F(r_E)$ for $n > E$ (that is, the team is ranked at its final standing's position r_E when the season ends).

The orbit of r_0 under F is the sequence of iterations

$$r_0, F(r_0), F^2(r_0), F^3(r_0), \dots, F^n(r_0), \dots, F^{E-1}(r_0)$$

We identify r_0 as the seed of the orbit. Let us identify r_E as the final standing's rank for a given team. Following Melo and Van Strien (1992, p. 26), we identify r_E as the point to which the orbit of r_0 under F converges.

Therefore, identifying the final journey as E , we can state

$$\lim_{n \rightarrow E} F^n(r_0) = r_E$$

We demonstrate that r_E is a fixed point of F with the following (Frame and Mandelbrot, 2002, p. 37):

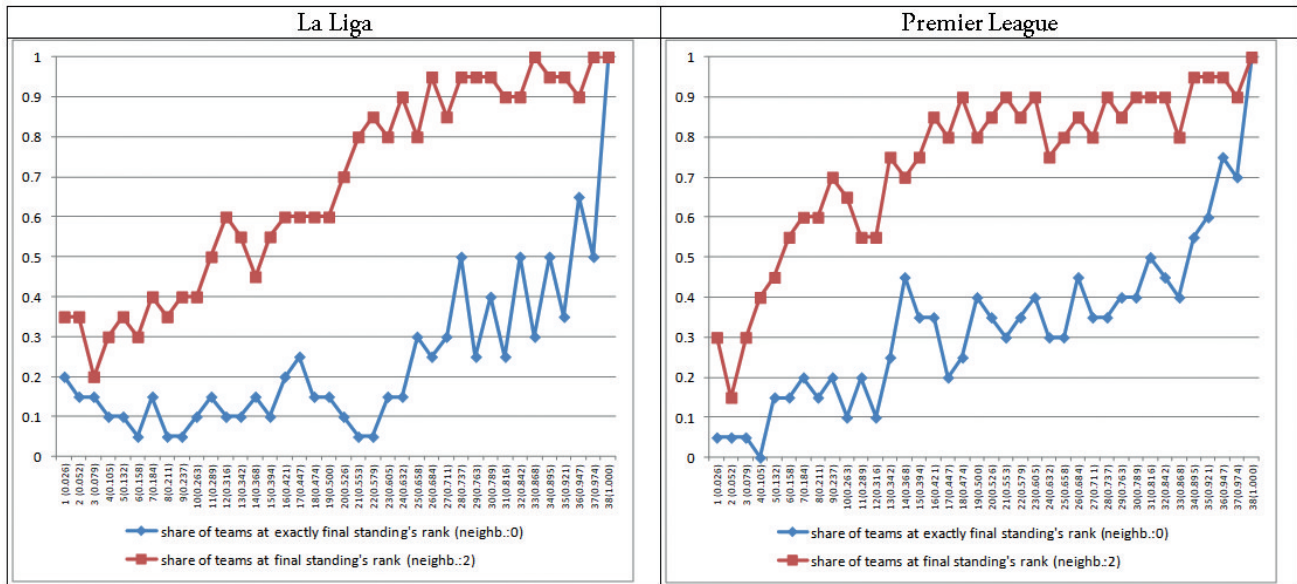
$$\begin{aligned} \lim_{n \rightarrow E} F^n(r_0) = r_E &\Rightarrow F(\lim_{n \rightarrow E} F^n(r_0)) = F(r_E) = r_E \\ &\Rightarrow \lim_{n \rightarrow E} F^{n+1}(r_0) = F(r_E) \\ &\Rightarrow r_E = F(r_E) \end{aligned}$$

2.2. The density of teams at their final standings' ranks after each season's journeys

Recalling Figures 1A–1D, we can confirm that the orbits of the 4 different teams passed their fixed points multiple times before the last journey. For instance, Manchester United (Fig 1A) occupied the first position at 28 of the 38 journeys, and Málaga (Fig 1B) was ranked 6th (its final standing's position) at 5 of the 38 journeys. Therefore, for each journey in each league, we can compute the density of fixed points as the share of competing teams that are located at their final standings' ranks after a given journey j .

Figures 2a and 2b represent these shares for the Spanish and the English teams in the 2012/2013 season considering two positions: at the exact position and within the range between two ranks above and below the exact position.

Figures 2A and 2B. Share of teams, after each journey, close to their final standings' ranks, season 2012/2013 (cases of La Liga and of Premier League).



Note: at the abscises axis, it is represented the number of past journeys (between parentheses it is represented the percentage of past journeys related to the total number of the season's journeys).

Commenting on Figures 2A and 2B, we observe the following: First, nearly all of the journeys in the two represented leagues had at least one team that was ranked at its final standing's rank after that journey. The exception occurred with the 4th journey of the 2012/2013 edition of the Premier League, which recorded no team ranked at its exact

final standing's position. As expected, the shares of the total number of teams that ranked at their final standings' positions exhibited increasing trends as long as the journeys were being played, independent of the neighborhood. As was also expected, if we enlarge the neighborhood, more teams (or a higher share of competing teams) were computed as ranked

at their final ranks. For instance, using the example of the Spanish soccer league, 10% of the teams (i.e., 2 teams) were exactly ranked at their final standings' positions after the 15th journey; however, if we consider all of the teams that were ranked between two ranks above and two ranks below their final standings' positions, we count 45% of the teams (i.e., 9 teams) after the same 15th journey. It is also interesting to observe that from this 15th journey until the last journey (the 38th), more than half of the competing Spanish teams were located in the short space between two ranks above and two ranks below their final standings' ranks. The English teams stabilized in the same ranged space just after the 6th journey of the 2012/2013 season.

We interpret Figures 2A and 2B following Mullin (1996, p.107) That is, these figures describe the density of fixed points for the set of competing teams in each professional soccer league given the percentage of journeys that had been played. As the percentage of journeys increases, we tend to observe in these two leagues an increased percentage of teams ranked at their final standings' positions.

The study of the functions related to the density of fixed

Clark Model (1951)/Dunham et al Model (2002): $\ln Y = \alpha - \beta * X + e$

Sherratt-Tanner Model (1963): $\ln Y = \alpha - \beta * X^2 + e$

Newling Model (1966): $\ln Y = \alpha - \beta * X - \chi^2 + e$

Ricker Model (1954): $\ln(1 + Y) = \alpha - \beta * \ln X - \chi * X + e$

We recall that we are interested in analyzing the density of teams at their fixed points throughout a regular season (that is, for our case, Y identifies the percentage of teams at their final standings' positions after each journey, and X identifies the percentage of journeys that had been run). Therefore, we must choose which of the previous functional forms better fits our data.

Using data from <http://www.statto.com>, we collected the data for the seasons between 2004/2005 and 2012/2013 for the six main European leagues: England, Germany, Italy, the Netherlands, Portugal, and Spain.

To define the more appropriate functional form, we followed Greene (2002) and Gujarati (2000) and used three conventional tests for choosing the functional form for our model that was to describe the density of teams in each league that were at their final standings' ranks and after each journey: the RESET test (Gujarati, 2000), the PE test (MacKinnon, White and Davidson, 1983), and the residual sum of squares after the Box-Cox transformation.

All of these tests suggested a linear relationship between

points has wide applications in many subjects; economics, urban planning, ecology, population dynamics, physics, and random dynamics are among the various subjects with a special focus on the study of the density of fixed points (Petters and Wicklin, 1998). Typically, researchers who focus on density functions choose to debate the concentration of points depending on a metric. The rationale is given by Mullin (1996), who suggests that the center, the core or the nucleus is by definition the space of the density of fixed points. As we move from that center, the density is lower.

Some widely discussed suggestions for modeling the relationship between the density of fixed points and the distance to the center have come from Clark (1951), Sherratt-Tanner (1963), Newling (1966), Dunham et al (2002), and Ricker (1954). For comprehensive lists, we suggest Haefner (1996) or Cade and Noon (2003).

We are going to list the most common functional forms tested in the literature, identifying as Y the density of points and as X the distance to the densest place (e identifies the error term, estimated under the classical hypotheses of a zero mean and of independently identically distributed variances):

our Y (the percentage of teams at their final standings' positions after each journey) and our X (the percentage of journeys that had been played). Full results are available on request.

Therefore, we are going to analyze the following specification for each league j:

$$Y_i^j = \alpha^j + \beta * X_i^j + \lambda * Z^j + e_i^j \quad (\text{equation 1})$$

In equation 1, Y_i^j identifies the percentage of teams in league j that are positioned at their final standings' ranks after the i-th journey. α identifies the estimated intercept, and X_i^j identifies the percentage of journeys in league j that had been played after the i-th journey. Z^j identifies the neighborhood of the final standing rank that we are using (from 0, exact position, to 6, between 6 positions above and 6 positions below the one reached at the final standing). e_i^j identifies the error terms, estimated under the classical hypotheses.

The next section exhibits the empirical effort to estimate equation 1.

3. Empirical section

3.1. Data and estimation methods

To study these six major European soccer leagues (England, Germany, Italy, the Netherlands, Portugal, and Spain), we recall that we collected data from <http://www.statto.com>. We collected the data for the seasons between 2004/2005 and 2012/2013. However, as you can check, there are more observations for each league than solely those that resulted from multiplying the number of journeys by the number of seasons. This occurred because we also included in our dependent variable the 7 neighborhoods considered (from 0, the exact position of the last standing, to 6, in which we observed the percentage of teams that were ranked between six ranks above and six ranks below their last standings).

Table 1 exhibits the descriptive statistics for our data. We can observe that the Premier League (the professional English soccer league) exhibited the highest value for the mean of the variable “percentage of teams at their final standings’ ranks”, although it was not significantly different from the means obtained for the other leagues. The “percentage of journeys” ranged from 0.026 (for the first journey of the round-robin competition involving 20 teams, as in the English, Italian and Spanish cases) to 0.974 (related to the 37th journeys of those cases). If we concentrate on the distance between quantile 75 and quantile 25, we observe that the leagues that showed more widely dispersed values related to the percentages of teams at their final standings’ ranks were the Portuguese (0.563) and the Dutch (0.556).

Table 1. Descriptive statistics (since the first journey to the previous before the last journey, for each season).

| | | Number of observations | Mean | Standard Deviation | Minimum | Maximum | Quantile75-quantile25 |
|------------------------------|---|------------------------|-------|--------------------|---------|---------|-----------------------|
| England (04/05-12/13) | % of teams at the final standing’s rank | 2331 | 0.670 | 0.305 | 0 | 1 | 0.550 |
| | % of journeys | 2331 | 0.500 | 0.281 | 0.026 | 0.974 | 0.474 |
| | Neighborhood | 2331 | 3 | 2.237 | 0 | 6 | 4 |
| Germany (04/05-12/13) | % of teams at the final standing’s rank | 2079 | 0.631 | 0.301 | 0 | 1 | 0.500 |
| | % of journeys | 2079 | 0.500 | 0.280 | 0.029 | 0.971 | 0.471 |
| | Neighborhood | 2079 | 3 | 1.803 | 0 | 6 | 4 |
| Italy (04/05-12/13) | % of teams at the final standing | 2331 | 0.631 | 0.299 | 0 | 1 | 0.500 |
| | % of journeys | 2331 | 0.500 | 0.281 | 0.026 | 0.974 | 0.474 |
| | Neighborhood | 2331 | 3 | 2.237 | 0 | 6 | 4 |
| Netherlands (04/05-12/13) | % of teams at the final standing’s rank | 2079 | 0.663 | 0.304 | 0 | 1 | 0.556 |
| | % of journeys | 2079 | 0.500 | 0.280 | 0.029 | 0.971 | 0.471 |
| | Neighborhood | 2079 | 3 | 1.802 | 0 | 6 | 4 |
| Portugal (04/05-12/13) | % of teams at the final standing’s rank | 1827 | 0.633 | 0.316 | 0 | 1 | 0.563 |
| | % of journeys | 1827 | 0.500 | 0.279 | 0.029 | 0.971 | 0.467 |
| | Neighborhood | 1827 | 3 | 1.802 | 0 | 6 | 4 |
| Spain (04/05-12/13) | % of teams at the final standing’s rank | 2331 | 0.640 | 0.304 | 0 | 1 | 0.500 |
| | % of journeys | 2331 | 0.500 | 0.281 | 0.026 | 0.974 | 0.474 |
| | Neighborhood | 2331 | 3 | 1.802 | 0 | 6 | 4 |

Note: For each seasons’s journey, there were 7 observations, related to the analyzed 7 neighborhoods (from the case in which a certain number of teams were at their final standings – neighborhood of ‘0’ – to the case in which the teams were within a range between six positions above and six positions below – neighborhood of ‘6’).

Following Okada and Samreth (2012), the quantile regression approach is more appropriate than is ordinary least squares (OLS) if the data exhibit such significant inter-quantiles distances as ours. Therefore, as an alternative to the OLS estimates with robust standard errors (from 10,000 bootstrapping repetitions), we must run quantile regression estimations.

A number of studies discuss the quantile regression approach and the quantile estimator (Li et al, 2012; Kim et al, 2002; Huber and Melly, 2011; Okada and Samreth, 2012). We opted for the model presented by Okada and Samreth (2012) and by Kim et al (2002).

According to this model, the quantile estimator is achieved when the researcher solves the optimization problem for

the θ th quantile (i.e., θ between 0 and 1), considering y_i as the dependent variable and x_i the vector ($k \times 1$) of the explanatory variables:

$$\min_{\beta \in R^k} \left[\sum_{i \in \{i: y_i \geq x_i' \beta\}} \theta |y_i - x_i' \beta| + \sum_{i \in \{i: y_i < x_i' \beta\}} (1 - \theta) |y_i - x_i' \beta| \right]$$

3.2. Results

The estimates of our equation 1 are presented in Table 2. We observe the percentage of journeys that exhibit estimated coefficients within the range of 0.271 to 0.497 for the six

leagues and for the ordinary least squares estimations with robust standard errors. Most of these coefficients are greater than 0.400, indicating that as the competition evolves, more teams approach their final standings' positions (which follows our main argument that rE is a limit fixed point for all the teams). Considering, for instance, the English case, if the percentage of journeys increased by 1 percent point, we would expect that more than 0.47% of the competing teams would be located at their final standings' ranks. Concretizing for the Premier League (20 competing teams, 38 journeys), our estimates find that after each four journeys (approximately 10% of the 38 total journeys), one more team is expected to be located at its final standing's position.

Table 2. Quantile regressions for determinants of percentage of teams at the final standing's rank.

| | England | | | | Germany | | | |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile |
| Intercept | 0.115*** (0.009) | 0.037*** (0.012) | 0.093*** (0.013) | 0.235*** (0.014) | 0.061*** (0.008) | -0.001 (0.009) | 0.063*** (0.015) | 0.139*** (0.013) |
| % of journeys | 0.469*** (0.013) | 0.400*** (0.018) | 0.493*** (0.018) | 0.490*** (0.020) | 0.497*** (0.013) | 0.383*** (0.014) | 0.493*** (0.021) | 0.567*** (0.019) |
| Neighborhood | 0.107*** (0.002) | 0.110*** (0.002) | 0.110*** (0.002) | 0.098*** (0.002) | 0.129*** (0.002) | 0.137*** (0.002) | 0.128*** (0.003) | 0.122*** (0.003) |
| N.Obs / ps.-R2 | 2331 / 0.801 | 2331 / 0.618 | 2331 / 0.568 | 2331 / 0.457 | 2079 / 0.812 | 2079 / 0.622 | 2079 / 0.576 | 2079 / 0.504 |

| | Italy | | | | Netherlands | | | |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile |
| Intercept | 0.133*** (0.010) | 0.030*** (0.011) | 0.100*** (0.013) | 0.288*** (0.016) | 0.233*** (0.017) | 0.103*** (0.011) | 0.161*** (0.021) | 0.250*** (0.010) |
| % of journeys | 0.439*** (0.016) | 0.380*** (0.017) | 0.475*** (0.020) | 0.475*** (0.023) | 0.271*** (0.028) | 0.270*** (0.016) | 0.398*** (0.030) | 0.472*** (0.015) |
| Neighborhood | 0.098*** (0.002) | 0.107*** (0.002) | 0.100*** (0.002) | 0.288*** (0.016) | 0.120*** (0.003) | 0.137*** (0.003) | 0.132*** (0.005) | 0.125*** (0.002) |
| N.Obs / ps.-R2 | 2331 / 0.733 | 2331 / 0.581 | 2331 / 0.523 | 2331 / 0.409 | 2079 / 0.567 | 2079 / 0.473 | 2079 / 0.453 | 2079 / 0.403 |

| | Portugal | | | | Spain | | | |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile | Robust OLS | 0.25 quantile | 0.50 quantile | 0.75 quantile |
| Intercept | 0.133*** (0.012) | 0.060*** (0.013) | 0.144*** (0.017) | 0.215*** (0.015) | 0.092*** (0.008) | 0.017* (0.010) | 0.069*** (0.011) | 0.177*** (0.013) |
| % of journeys | 0.432*** (0.011) | 0.346*** (0.020) | 0.383*** (0.027) | 0.545*** (0.023) | 0.434*** (0.012) | 0.358*** (0.014) | 0.454*** (0.017) | 0.481*** (0.020) |
| Neighborhood | 0.124*** (0.003) | 0.131*** (0.003) | 0.123*** (0.004) | 0.106*** (0.003) | 0.110*** (0.001) | 0.114*** (0.002) | 0.113*** (0.002) | 0.104*** (0.002) |
| N.Obs / ps.-R2 | 1827 / 0.809 | 1827 / 0.632 | 1827 / 0.604 | 1827 / 0.503 | 2331 / 0.823 | 2331 / 0.650 | 2331 / 0.594 | 2331 / 0.492 |

The Dutch league exhibits very interesting results if we only focus on the OLS estimates. We can observe that the estimated intercept has the larger value across the estimated intercepts for the six leagues and that the effect of the evolution of the percentage of journeys has the lower value (0.271). Our

interpretation converges with those of Koning (1999), Nagsbandi et al (2011), and Pawlowski and Budzinski (2013), which state that this league is the one of the observed six professional competitors with the most uncertain final standing positions, independent of the number of past journeys. In

reality, our estimates find that after only eight journeys, one more team is expected to be at its final standing's rank.

Applying the intermediate value theorem (Carter, 2001, p.258), we can confirm that our trajectories are continuous. According to Schauder's theorem, our estimated trajectories have their own fixed points:

"Let S be a nonempty, compact, convex subset of a normed linear space. Every continuous function $f: S \rightarrow S$ has a fixed point" (Carter, 2001, p. 258)

In our case, the fixed points occur when the estimated share of teams that are close to their fixed points converges with the percentage of journeys that have been run.

Although this interpretation is not our main objective, we can use this measure as an alternative measure of the competitive balance of the league—more competitive leagues are associated with fixed points that are located closer to point (1,1). Considering robust OLS estimates and a neighborhood of zero, this value will give us the estimated percentage of the number of journeys in which it is expected that the same percentage of teams will be at their final standings' ranks.

Table 3 gives us these "fixed points," which we interpret as measures/indexes of global competitive balance for the leagues. Given that this fixed point signals the start of the teams' convergences toward their final standings' positions, a lower value for each league reflects less competitive balance in the league (that is, more teams have their close final standings' ranks defined after earlier journeys than in a league characterized by a higher fixed point).

Table 3. Fixed points for six European soccer leagues (neighborhood of 0).

| Leagues | Fixed Point (Robust OLS), neighborhood: 0 |
|-------------|--|
| England | 0.217 (approximately the 8 th journey) |
| Germany | 0.121 (approximately the 4 th journey) |
| Italy | 0.237 (around the 9 th journey) |
| Netherlands | 0.320 (approximately the 11 th journey) |
| Portugal | 0.234 (approximately the 7 th journey) |
| Spain | 0.163 (approximately the 6 th journey) |

For instance, we found 0.217 to be the fixed point for England. That is, after 21.7% of the total number of journeys (i.e., roughly the eighth journey), 21.7% of the teams (i.e., 4 teams) are expected to be at the same rank where they will be at the final journey.

The next table (Table 4) gives us the fixed point for each league but considering a neighborhood of two ranks. This weakens our previous restriction (used in Table 3) that the teams must be at exactly their same final standings' ranks. With a neighborhood of two positions, we admit that each

team may be ranked between two ranks above and two ranks below its final rank in a given journey during the season.

Table 4. Fixed points for six European soccer leagues (neighborhood of 2).

| Leagues | Fixed Point (Robust OLS), neighborhood: 2 |
|-------------|---|
| England | 0.620 (24 th journey) |
| Germany | 0.634 (22 nd journey) |
| Italy | 0.586 (22 nd journey) |
| Netherlands | 0.724 (24 th journey) |
| Portugal | 0.671 (20 th journey) |
| Spain | 0.551 (21 st journey) |

Based on Table 4, we expect that 67.1% of the competing Portuguese teams (approximately 11 of the 16 teams) will be ranked between two ranks above and two ranks below their final standings' positions after the 20th journey. Once again, the Dutch league is expected to be the most uncertain/competitive because after only the 24th journey, 72.4% of the 18 teams (i.e., 13 teams) are estimated to be ranked between two ranks above and two ranks below their final standings' ranks.

If we focus now on the results obtained by the quantile regressions (returning to Table 2, the last three columns for each league), we observe that, as expected, the median regression (the 0.50 quantile) provides estimates that are statistically significantly close to the robust OLS estimates. The estimated intercepts, depending on the quartiles, are associated with the quartiles of each league's percentage of teams at their final standings' positions after each journey. We observe that the estimated intercept for the Dutch League is the highest among the six leagues. Combining this result with the slight impact of the number of past journeys (Table 2), we can also claim that the number of rank changes in the Dutch league for a significant number of journeys after the start of the league is higher than the number of rank changes for the other leagues².

2 Given a number of past journeys and assuming the same neighborhood, we claim that the number of rank changes is higher for a given league if the estimated value for the percentage of teams at their final standings is lower than another league's estimated value for the respective percentage of teams at their own final standings. For an even number n of teams, there are $n/2$ matches at each journey. If all of the teams changed their previous ranks, there would be n rank changes. If after a second journey, all of these teams were ranked at their final standings, the percentage of teams at their final standings would be 100%. Assume that after a third journey, only two teams changed their ranks. Then, the percentage of teams at their final standings would be $(n-2)/n$, which is near 100%. By a reverse possibility, now assume that all but 2 teams changed their ranks after three journeys. Then, the percentage of teams at their final standings would be $2/n$, and there would be $n-2$ rank changes after the third journey (which is higher than 2 for any even number n greater than 4). Therefore, lower percentages of teams at their final (or "fixed-point") standings correspond to more past rank changes and to greater levels of competitive balance.

The impact of the percentage of journeys on the percentage of teams at their final standings' ranks tends to increase as the quantile increases (Figures 3–5 plot, respectively, the estimated intercepts, the estimated coefficients for the percentage of teams at their final standings' positions and the estimated neighborhood's impact across the three quanti-

les). Following Table 5, we can reject the null hypothesis of the same coefficients across quartiles. This is an important conclusion because it shows us the heterogeneity across the observations that were collected in each league and the sign of the marginal effect of increasing the explanatory variables (Greene, 2002).

Figure 3. Plots of the estimated Intercepts on percentage of teams at the final standing's rank.

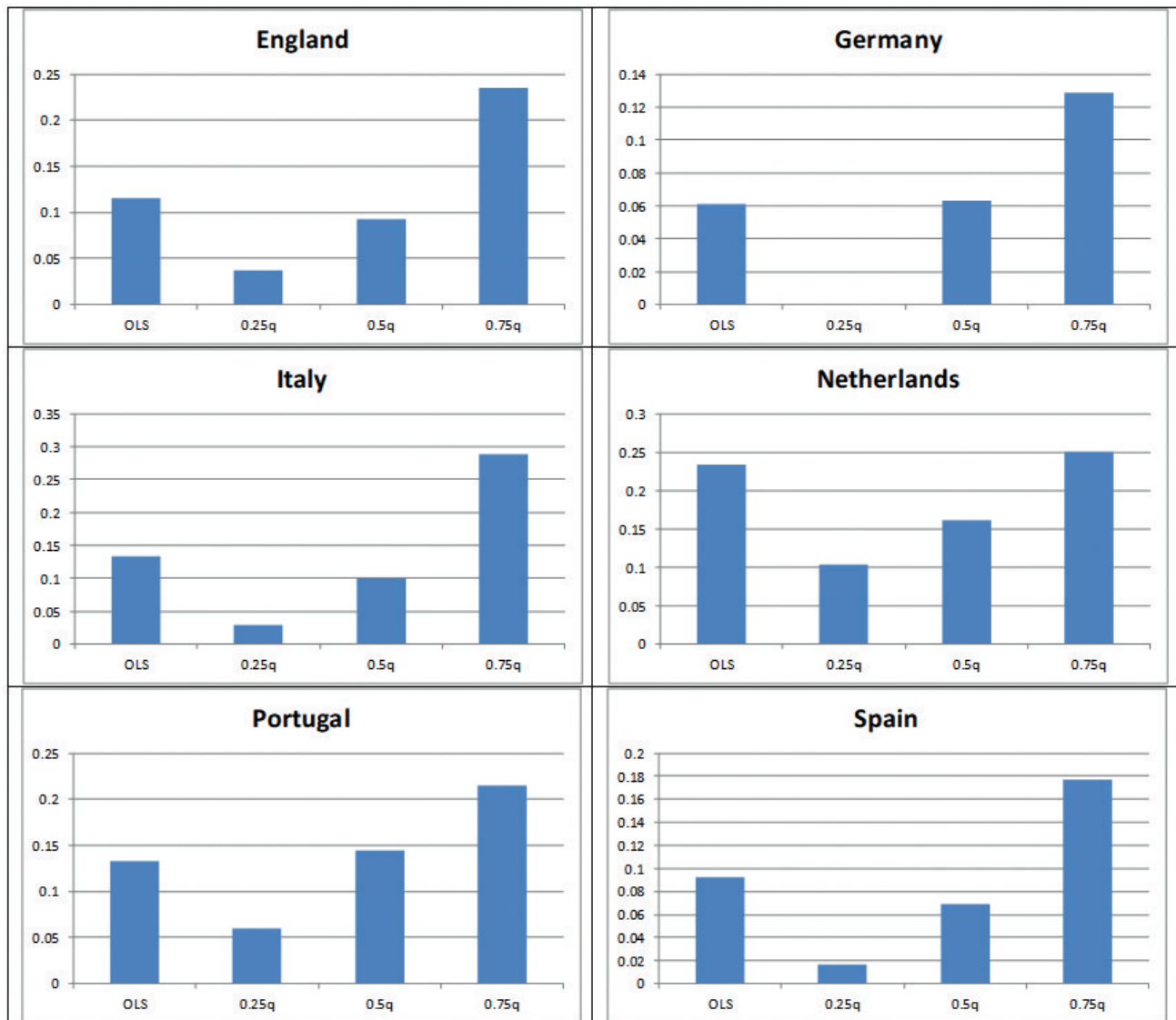


Figure 4. Plots of the Impact of percentage of journeys on percentage of teams at the final standing's rank.

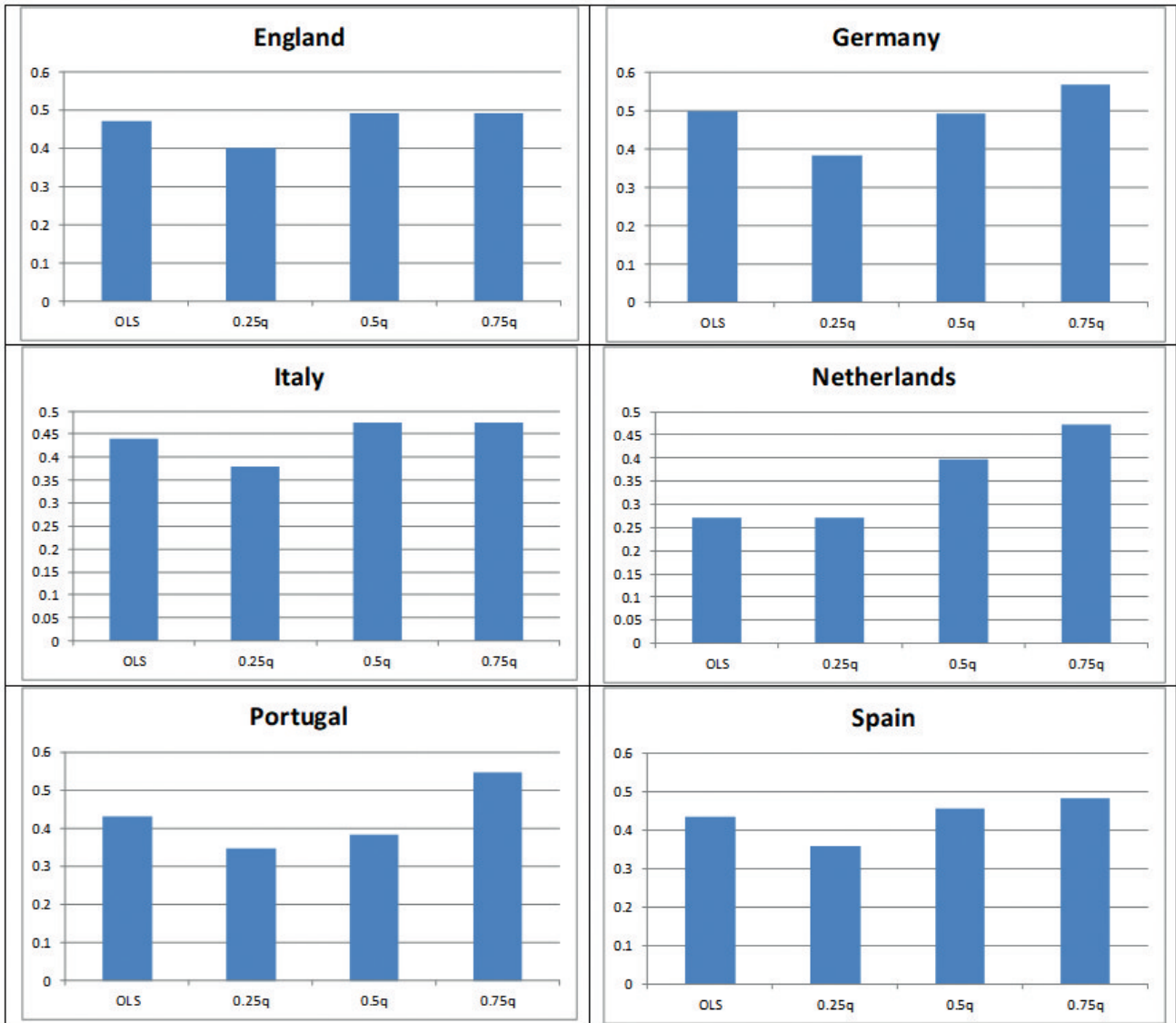


Figure 5. Plots of the Impact of Neighborhood on percentage of teams at the final standing's rank.



Table 5. Tests of equality of coefficients across quantiles (H0: 0.25 quantile = 0.75 quantile).

| | % of journeys | Neighborhood |
|-------------|------------------------------------|-------------------------------------|
| England | F-stat (1,2331)=3.34 p-value=0.068 | F-stat (1,2331)=4.24 p-value=0.040 |
| Germany | F-stat (1,2079)=7.80 p-value=0.005 | F-stat (1,2079)=4.16 p-value=0.042 |
| Italy | F-stat (1,2331)=3.32 p-value=0.069 | F-stat (1,2331)=11.23 p-value=0.001 |
| Netherlands | F-stat (1,2079)=7.80 p-value=0.005 | F-stat (1,2079)=9.12 p-value=0.003 |
| Portugal | F-stat (1,1827)=6.56 p-value=0.010 | F-stat (1,1827)=6.56 p-value=0.010 |
| Spain | F-stat (1,2331)=4.71 p-value=0.030 | F-stat (1,2331)=3.53 p-value=0.060 |

The difference between quantiles is especially significant for the estimated coefficients for the 0.25 and 0.75 quantiles for the impact of the percentage of journeys that were observed for the German (F-statistic: 7.80), Dutch (F-statistics: 7.80),

and Portuguese cases (F-statistics: 6.56). Essentially, the percentage of teams ranked in certain neighborhoods within their own final standings' ranks are more disperse for these

three leagues than they are for other leagues (as confirmed by Table 1, related to the descriptive statistics).

The impact of the neighborhood is, as expected, positive. As the neighborhood of the ranks expands, more teams approach their final standings' positions. The estimated value of 0.107 for the English case indicates that a one-unit increase in the neighborhood leads to more than 10.7% of teams (that is, two more teams for the English league) being ranked close to their final ranks at a given journey during the season.

In the cases of Germany, the Netherlands and Spain (Figure 5), the decreasing coefficients estimated for neighborhood that are attributable to the increase in quartiles suggest the presence of a form of decreasing marginal effects. We interpret that with enlarged neighborhoods, these three leagues find it more difficult to exhibit significant changes in their teams' standings. That is, we do not obtain a significantly higher share of teams that are close to their final standings' ranks solely because we enlarge the considered neighborhood.

We also observed the estimated coefficients for the neighborhood increases as quartiles increased. By a reverse interpretation from that of the previous three cases, we interpret that with enlarged neighborhoods, the Italian league shows increasing values for the percentage of teams that are ranked close to their final standings' positions.

A discussion on the implications of the results

The results that we obtained confirmed three main facts for European soccer leagues:

- i) It is very rare that after a journey, no team is positioned at its exact final standing's rank;
- ii) Until the last journey, all of the professional leagues showed changes in their teams' ranks (more likely to occur during the second third of the journeys);
- iii) In all leagues, more than 50% of teams are located within a short final standing's rank's neighborhood (comprising between two ranks above and two ranks below the final standing's position) after the first third of the journeys.

Therefore, if we observe that such a significant number of teams are located near their own final standings' ranks at early journeys, there remain two relevant questions:

- Why does this happen? and
- Why are there so many journeys if the main positions and the teams' ranks become reasonably defined during the middle of the season?

We are aware that each of these questions deserves a proper paper. However, at least for tracking purposes for further research, we can now suggest that, following Stefani (1997) and Lasek et al (2009), there are two main reasons that the

standings' ranks after each journey converge to the final standings' ranks.

The first reason (Stefani, 1997) is that all competing teams attempt to stabilize at their own most appropriate positions as soon as possible. Being stable at a position that at least converges to the team's historical median rank minimizes the expectation shocks of fans, supporters and team investors. It eases the squad's management throughout the season and during other competitions (such as national or international leagues and cups).

The second reason relates to the recognition that a team's quality profile results from the sum of the quality profiles of the team's players. Therefore, the relative rank of a team or a competitor is reasonably defined after a small number of early journeys (Lasek et al, 2009). For instance, chess tournaments use the 'Swiss system' to pair players with similar numbers of accumulated points; this system does not use the round-robin system of soccer leagues but it opts for a significantly shorter number of matches (typically $\log_2 N$, where N is the number of players³).

These reasons strictly relate to the answers to the second question of this discussion: Why do we like so many journeys in a regular soccer season? For two additional reasons: cultural and purely economic. The first reason relates to cultural institutions. Across the European leagues and for generations for nearly one hundred years, we have learned to enjoy lengthy seasons, which allow millions of soccer fans, supporters and investors to change expectations throughout the journeys, to follow newspaper standings, to surround themselves in the soccer spectacle nearly every day all season long (Fahrmeir and Tutz, 1994; Fallahi et al, 2011). The second reason is mostly economic. More journeys bring more revenues to teams and, primarily, to the competitions' organizers (European Commission et al, 2004). More journeys also tend to improve players' athletic performance, which generates higher transfers to the selling clubs (Dobson, 2001). Finally, more journeys (clearly associated with higher costs and expenses) also generate a greater multiplier effect (Hamill, 2001) for all of the teams and for the soccer industry as a whole (Parrish, 2013).

4. Conclusions and Additional Work

In this article, we analyzed the evolution of the ranks of professional soccer teams in European leagues from the 2004/2005 to the 2012/2013 seasons. We studied this evolution with reference to the analysis of fixed points. We demonstrate that the final standings ranks of each soccer team are fixed points for the orbit described by the sequence of

³ For instance, 30 players will only have 5 paired matches (not the 58 round-robin paired matches), with a reasonably well-defined final ranking (also improved by the break-even factor, the *buchholz*).

iterations/journeys. We observed that all of the professional soccer teams occupied their final standings' ranks multiple times before the last journey and that different European soccer leagues exhibit different shares of teams that are at their own final standings' ranks after each journey.

Referring to quantile regressions, we observed that the Dutch league exhibited the largest value for the coordinates of its fixed point considering the six leagues that were studied (England, Germany, Italy, the Netherlands, Portugal and Spain). We obtained estimates that demonstrated that only at the late journeys did a significant share of Dutch teams stabilize at their ranks; that is, before those journeys, the Dutch teams had more rank changes than did the other European leagues. As an additional observation, the Italian Serie A and the English Premier League are the two leagues in which most of the teams tend to stabilize at their final standings' ranks at earlier journeys.

These results suggest some final observations and policy implications. First, as we observed, it is very rare that after a random journey, no team is at its final standing's position; this is interpreted to indicate that each team gains from being ranked as high as possible as soon as possible. Second, it is also true that journey after journey, there were always rank

changes, demonstrating that we must wait until the last journey to receive the final definite standings. However, when we considered close but not exact positions (i.e., when we analyzed the neighborhoods of the final standings' ranks), we concluded that for all of the leagues, most of the competing teams are located in their final ranks' neighborhoods after the first 33% of the season. This occurs for the neighborhoods two ranks above and two ranks below each team's final standing's rank.

We are aware that this original work opened a very significant space for further developments. First, we intend to collect more data, on more European professional leagues (for instances, for the French, Ukrainian, and Russian leagues) and for the seasons in which each victory was awarded only 2 points. Second, we also want to detail the changes within each rank's quartile (attempting to answer questions such as: "Are there more changes after 50% of the journeys for the top-5 ranked teams or the last-5 positioned teams?"). Finally, we want to extend this methodological analysis to other sports, not only team sports and their standings' tables but also expanded to study the relevance of fixed points (such as the final position of a racing competitor) to motor races, marathons, and professional cycling.

References

- Cade, B. and B. Noon (2003). "A gentle introduction to quantile regression for ecologists". *Frontiers in Ecology and the Environment*, 1(8), 412-402.
- Carter, M. (2001). *Foundations of mathematical economics*. Cambridge MA: Massachusetts
- Clark, C. (1951). "Urban population densities." *Journal of Royal Statistical Society*. 114 (2). 121-134
- Dobson, S. (2001). *The Economics of Football*. New York: Cambridge University Press, 2001. 458 p.
- Druck, S., Fang, F. and S. Firmo (2001). "Fixed Points of discrete nilpotent group actions on S^2 ". arXiv:math/109015 v1 (3 september 2001)
- Dunham, J., Cade, B. and J. Terrell (2002). "Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles." *Trans Am Fish Soc* 131: 86-98.
- European Commission, Ofcom, and Human Capital – Media Strategy and Research (2004). *Premier League Football - Research into viewing trends, stadium attendance, fans' preferences and behaviour, and the commercial market*. COMP/C/38.173 – FAPL. Available through http://ec.europa.eu/competition/antitrust/cases/dec_docs/38173/38173_104_7.pdf
- Fahrmeir, L. and Tutz, G. (1994) "Dynamic stochastic models for time-dependent ordered paired comparison systems.", *J. Am. Statist. Ass.*, 89, 1438-1449.
- Fallahi, A., Asadi, H. and M. Khabiri (2011). "The Comparison of the Importance of the Factors Affecting Spectators' Attendance in Football Matches in Iranian Professional League Within Age Groups". *World Journal of Sport Sciences* 4 (2): 159-165, 2011
- Frame, M. and B. Mandelbrot (2002). *Fractals, Graphics, and Mathematics Education*. The Mathematical Association of America, Washington DC
- Greene, W. (2002). *Econometric Analysis*. Prentice Hall, New York
- Gujarati, D. (2000). *Econometria Básica*. Makron Books, Rio de Janeiro
- Haefner, J.W., (1996). *Modeling Biological Systems: Principles and Applications*. Chapman & Hall, New York, NY
- Hamil, S. (2001). *The Changing Face of the Football Business*. London: F. Cass, 2001.
- Huber, M. and B. Melly (2011). "Quantile Regression in the Presence of Sample Selection," *Economics Working Paper Series 1109*, University of St. Gallen, School of Economics and Political Science.
- Kim, T., Kim, H. and S. Hur (2002). "Asymptotic properties of a particular nonlinear regression quantile estimation," *Statistics & Probability Letters*, Elsevier, vol. 60(4), pages 387-394, December.
- Koning, R., (1999). "The competitive balance in Dutch soccer," *Research Report 99B04*, University of Groningen, Research Institute SOM (Systems, Organisations and Management). Institute of Technology Press.
- Lasek, J., Szlavik, Z., and S. Bhulai (2009). "The predictive power of ranking systems in association football". *Int. J. of Applied Pattern Recognition*, 1 (1). 27-46
- Li, L., Lee, L., Chen, C., and B. Guo (2012). "On unbiased optimal L-statistics quantile estimators," *Statistics & Probability Letters*, Elsevier, vol. 82(11), pages 1891-1897.
- MacKinnon, J., White, H., and R. Davidson (1983). "Tests for model specification in the presence of alternative hypotheses : Some further results," *Journal of Econometrics*, vol. 21(1), pages 53-70.
- Marteli, M. (1999). *Introduction to Discrete Dynamical Systems and Chaos*. Wiley Publisher, New York.
- Melo, W. and S. Van Strien (1992). *One-Dimensional Dynamics*. Springer-Verlag, Berlin
- Mullin, A. (1996). "On the Smarandache function and the fixed-point theory of numbers". In C. Dumitrescu, V. Seleacu, eds. (1996). *Smarandache notions vol. 7*. Infinite Study, American Research Press, New York.
- Naghshbandi S, Yousefi B, Etemad Z, and M Moradi (2011). *The com-*

- parison of competitive balance in football Premier Leagues of England, Germany, Spain, France, Italy and Iran: A case study from 2009-2010 Season. *J. Hum. Sport Exerc.* Vol. 6, No. 4, pp. 673-681
25. Newling, B. (1966). "The spatial variation of urban population densities." *Geographical Review.* 59 (3). 342-359
26. Parrish, C. (2013). "Soccer Specific Stadiums and Designated Players: Exploring the Major League Soccer Attendance Assumption". *International Journal of Sport Management Recreation & Tourism*, Vol.12, p.57-70, 2013
27. Pawlowski, T and O Budzinski (2013), "The (monetary) value of competitive balance for sport consumers - A stated preference approach to European professional football", *International Journal of Sport Finance* 8(2), 112-123.
28. Petters, A. and R. Wicklin (1998). "Fixed points due to gravitational lenses". *Journal of Mathematical Physics* 39 (2), 1011-1023
29. Ricker, W. E. (1954) "Stock and Recruitment" *Journal of the Fisheries Research Board of Canada*, 11(5): 559-623
30. Sherratt, M. and H. Tennant (1963). "Population density pattern in eight Asian cities". *Proceedings of Association of American Geographers.* 79-68
31. Stefani, R. (1997) "Survey of the major world sports rating systems". *Journal Applied Statistics*, 24, 635-646