Does diversity in the payroll affect soccer teams' performance? Evidence from the Italian Serie A

Caruso, Raul and Carlo, Bellavite Pellegrini and Marco, Di Domizio

December 2016

Does diversity in the payroll affect soccer teams' performance? Evidence from the Italian *Serie A*

by

Raul Caruso, Catholic University of the Sacred Heart

Carlo Bellavite Pellegrini, Catholic University of the Sacred Heart

Marco Di Domizio, University of Teramo

Abstract

This paper empirically investigates the impact of diversity in wage levels of players on seasonal performances of teams in the top Italian soccer league, namely the Serie A . We explore the payroll of 32 professional football teams in the Italian Serie A to compute three measures of diversity and concentration in wage levels, namely the Gini, the Shannon and the Simpson indexes from season 2007/08 to 2015/16. We use the percentage of points achieved by teams as dependent variable, and then we employ panel data techniques estimating random and fixed effect models. We find that only the Simpson index is significantly associated with sport performance. In particular, it appears that sport performance improves as diversity in payroll decreases.

JEL CODE: Z20, Z22, **Z21, L83, J49.**

KEYWORDS: diversity in wage level, inequality, payroll and sport performance, Italian serie A

Introduction

This paper analyses whether sport success is correlated with diversity in wage levels among players in Italian Serie A between season 2007/2008 and season 2015/2016. By diversity we mean a quantitative measure that depends on the number of different types in a population and their relative abundance. Whenever types differ by wage levels, the concept of diversity overlaps with that of inequality. Therefore, for sake of simplicity we investigate whether diversity in the payroll affects the performance of soccer teams in the Italian Serie A.

In fact, economic theory devotes substantial attention to diversity in the payroll (i.e. distribution of salaries) and its impact on the productivity of employees and firms. Since the pioneering contribution of Lazear and Rosen (1981) and Lazear (1989), despite the expected benefits of stimulating competition among employees by wages differentiation, a large body of literature also stresses the likely negative impact of this pay strategy on productivity.[i] The latter point is in line with the *cohesion theory* (Levine, 1991). The *cohesion theory* states that firms can increase productivity by narrowing wage dispersion among workers since this policy improves the cohesiveness of the group. Here cohesiveness has to do with *i*) the within-group harmony, *ii*) the force that keeps the members from leaving the group, *iii*) the capacity of the group to maintain integrity, *iv*) the extent to which the members reinforce each others' expectations regarding the value of maintaining the identity of the group (Stogdill, 1972). This theory is also related to the fair wage-effort hypothesis as expounded in Akerlof and Yellen (1990), according to which efforts of workers decrease as their wage falls short of their expectations. Another theoretical linkage can be found in Cubel and Sanchez-Pages (2012) that present a model conflict in which a group of agents defend their individual income from an external threat by pooling their efforts against it. The winner of this confrontation is determined by a contest success function where members' efforts display a varying degree of complementarity. Individual effort is costly and its cost follows a convex function. The latter assumptions, in particular, are in line with the theoretical model of sport tournaments. The theoretical outcome is a natural relationship between the group's probability of victory and the Atkinson index of inequality. If members' efforts are complementary or the cost function convex sufficient, less diversity within the group increases the likelihood of victory against the external threat. In other words in the presence of convex costs a larger inequality in the expected payoffs leads to fewer efforts so reducing the probability of a win. This can be applied easily to a sport scenario which is often characterized by convex costs.

In the wake of this, we investigate which theory holds for the Italian soccer top league, namely the *Serie A*. In practice, this paper analyzes the impact of diversity in wage levels among players on seasonal performance. Therefore, by means of panel data techniques we estimate the impact of diversity of wages on the performance. We find that an index of diversity – the Simpson index – contributes to explain sport performances. In particular, there is a negative association between the Simpson index and the sport performance; in other words, the higher the diversity in the payroll, the lower the sport performance. It is worth noting that the result is not clear-cut. In particular, we show that the diversity of wage levels among players is statistically significant only when measured by the *Simpson* index. In fact, we have employed three different indexes of concentration/diversity, namely the *Gini,* the *Shannon* and the *Simpson*.[ii] Only the Simpson index appears to be statistically and significantly associated with the dependent variable of sport success.

As dependent variable we have employed the percentage of points achieved by teams as the proxy for the performance (dependent variable), and consider as covariates a set of variables drawn from the related literature

The paper is organized as follows. In a first section a survey of the related literature on diversity in wage levels and team performance is presented. The following section is dedicated to the description of the dataset and of the variables. The last section presents the econometric model and the related results providing some conclusive remarks and an agenda for future research will follow.

The payroll-performance relationship: a brief survey

The performance and related success in team sports depend on several elements; the most important is indisputably the availability of talent. This hypothesis is supported by a wide range of papers that have investigated the talent-performance relationship for several professional team sports, both theoretically and empirically. Since pioneering theoretical papers by El Hodiri and Quirk (1971) and Scully (1974), the probability of success had always been associated with teams' talent availability. There, the winning ratio between two representative teams was approximated by the units of talent ratio, and used to compare the competitive balance equilibria associated to different player labour market schemes. The mentioned contributions were eventually enriched by a number of empirical contributions where, with few exceptions (Franck and Nüesch, 2010), the payroll is considered the better approximation of talent.[iii]

However, the payroll-performance relationship has not been explored only considering the nexus *more wages – more talent – more victories*. In fact, the sports economic literature focuses also on the wages distribution perspective and on its potential effect on teams' performance.

In the literature on industrial organization, the distribution of wages among employees has been scrutinized as it can be supposed to affect the productivity of companies and workers. In fact, on the one hand some wage inequality may trigger more effort and an increased productivity (*tournament theory*) (Lazear and Rosen, 1981; Ramaswamy and Rowthorn, 1991). In other words, differentiated reward schemes would enhance competition between employees so improving firms' productivity.

As noted above, alternative to the tournament approach is the *cohesion theory* as expounded in Akerlof and Yellen (1990) and Levine (1991). According to this approach firms are expected to improve their productivity by equalizing the wages because a lesser degree of wage dispersion may contribute to strengthen relationships and solidarity among employees.

Applying this line of reasoning to team sports organization, the empirical literature has devoted a substantial effort to validate either the *tournament* or the *cohesion theory*. The relationship between professional players' wages distribution and the team overall performances has been explored for different professional team sports. One of the most explored has been, traditionally, baseball, particularly Major League Baseball (MLB). In this context the empirical investigations discovered a negative relationship between team success and wages dispersion, so supporting the *cohesion theory*. Richards and Guell (1998) find a negative association between wage distribution, measured by the variance of salaries, and on field performance, proxied by the winning percentage. The same negative relationship between team performance and pay structure emerges in Bloom (1999), where the dispersion among salaries is measured by the Gini coefficient, although pay dispersion is able to explain only 5 per cent of the variation in the winning percentage. The *cohesion theory* is validated in the baseball context also by Depken (2000) in an empirical study covering the period from 1985 to 1998, where the wage dispersion is measured by the Herfindahl-Hirschman index (HHI), and by Wiseman and Chatteriee (2003) in their analysis covering the period 1985-2002, that used the same dependent variable and the Gini coefficient as a proxy for the salary dispersion. De Brock et al. (2004) used unconditional and conditional measures of wage inequality, where the first does not take into account the salary dispersion for observable differences in the performance of individual players, while the second does. At the first stage the dispersion is measured by the HHI, and team performance by the won-lost percentage; the sign of the relationship is negative, confirming the *cohesion theory*. Not dissimilar results emerge at the second stage, when controls are introduced for individual performances and team effects are directly estimated. Again, Jewell and Molina (2004) focus on MLB teams' records for the period 1985-2000; they use a Cobb-Douglas stochastic production frontier random effects panel data estimation, and find a negative relationship between salary dispersion, measured by a predicted Gini index, and the performance, approximated by the winning percentage.

The same result is reached by Avrutin and Sommers (2007) in an empirical investigation based on data from 2001 to 2005, using standard measures for team performances (win percentage) and salary distribution (Gini coefficient). More recently Annala and Winfree (2011) explore data on the MLB from season 1985 to 2004. They use the same measures of sport performance and payroll inequality, and find statistically significant negative relationships for the period 1985-2004, both with pooling data and team fixed effects approach. Moreover, Breunig et al. (2014) develop a theoretical model, built up from a Contest Success Function (CSF), to analyse the effect of wage dispersion on individual effort, and then on team performance. The model allows for both positive and negative effects of wage inequality on performances, *tournament* or *cohesion theory* dominance respectively, depending on the key parameters to be estimated. They estimate by an OLS regression a model where the winning percentage is the dependent variable, and the average salary and Gini coefficient – alternatively the HHI – are the covariates, together with dummies to control for team fixed effects. The dataset includes observations collected from MLB in the period 1985-2010; they conclude that wage dispersion is negatively associated to players' effort, and then to team performance, supporting the idea of Annala and Winfree (2011).

At last Tao et al. (2016) have estimated a dynamic panel model based on MLB matches from 1985 to 2013 adopting the winning percentage as the proxy for team performance. The authors include the lagged dependent variable, and a number of control variables associated to the wage dispersion and to team specific controls. The empirical estimation confirms the negative association between wage dispersion and team performance although the statistical significance depends on the proxy for pay dispersion, if Gini or HHI, and on the selection of team payroll level or team payroll relative position as covariate.

Different results arise from studies exploring this topic within the *National Basketball Association* (NBA). Simmons and Berri (2011) approach the wages inequality issue by using a conditional measure, so that they estimated two measures of pay inequality, *Gini Predicted* and *Gini Residuals*, to be used successively in the main equation of performance-wages inequality relationships. The *Gini Predicted* relates to the dispersion "of" expected salaries, while the *Gini Residuals* relates to the dispersion "around" the expected salaries. In their fixed effect model estimation they find a positive effect of wage dispersion on performance since player productivity is positively related to the dispersion "of" expected salaries. This result validates the tournament theory for the NBA as previously suggested by Frick et al. (2003). Other studies on NBA found inconclusive results (Berri and Jewell, 2004; Katayama and Nuch, 2011) such as for the *National Hockey League* (NHL) (Sommers, 1998; Kahane, 2012).

Interesting results appear from the National Football League (NFL) context. As example, Frick et al. (2003) in the above mentioned contribution, find a negative but not significant relationship between pay dispersion and team performance. On the contrary, Mondello and Maxcy (2009) find the negative relationship to be statistical significant in an empirical investigation based on 254 observations from 2000 and 2007. They use the coefficient of variation as a measure of pay inequality, considering for the first time ever that part of salary associated to the bonuses. Together with the negative association between dispersion and performance, they also find a positive relationship between dispersion and total revenues. This implies a potential conflict for management between on field performance and total revenue goals in terms of wage structure strategies.

If we concentrate on the soccer team environment, Franck and Nüesch (2011) allow for potential non-linearity in the relationships between wage inequality and performance. Their empirical analysis is based on teams appearing in the first German soccer league from the season 1995/96 to 2006/07. They use as dependent variable the winning ratio and a modified league standing at the end of the season, alternatively. As proxy for pay dispersion they consider the Gini and the coefficient of variation (CV), both linear and squared, alternatively. Adding other control variables such as the talent heterogeneity, the wage expenditures and the roster size, they estimate a 2SLS model with team fixed effects. The coefficients associated to the wage dispersion variables appear to be statistically significant, both Gini and CV, both if linear and squared. In particular, a U-shaped relationship emerges for which teams having either a high or low level of pay dispersion are more successful than teams with a medium level of wages distribution. Again, Coates et al. (2016) focus on *Major League Soccer* (MLS) in North America and provide support for the cohesion theory. Studying data on 19 teams playing in the MLS from 2005 to 2013, they estimate a model using the points achieved as a measure of the on field seasonal performance, the Gini and CV, alternatively, as proxy for pay dispersion – both linear and squared, and among others, they add the relative wage as control variable. From their estimation it emerges that performance in MLS is negatively correlated with the increase in salary inequality, and non-linearity in the wages dispersion-performance is excluded.

Finally, Yamamura (2015) analyses the wages dispersion-performance relationship of the Japanese professional football league (J-League); he uses data from 1993 to 2011 distinguishing two periods: developing (1993-1997) and developed stages (1999-2010), where the qualification of the Japanese national team at the World Cup in 1998 is considered as the dividing line. The author estimates both a panel fixed effects model and an Arellano-Bond type dynamic model, to control for endogeneity bias and unobservable fixed team effects. He uses the seasonal winning ratio as a proxy for the on-field performance, the HHI associated to the inter-team annual salary as proxy for the wages dispersion, and other controls focused on players' wages level and age. In the whole period estimation the coefficient associated to the wages dispersion is statistically significant and negative, but only in the fixed effects estimation. When the estimation distinguishes between developing and developed stages it emerges that the cohesion theory holds during the first stage, whereas the on-field performance is not influenced by wages dispersion in the developed stage, both in the fixed effects model and in the dynamic panel approach.

The empirical investigations on the Italian football context are quantitatively poor. Few studies have empirically analysed the determinants of team performance. Di Betta and Amenta (2010) identify 'tradition' as the main factor of success. From their analysis it emerges that, from 1929 to 2009, *Serie A* would be characterized by a 'self-reinforcing mechanism' of supremacy according to which only ten teams can be included in the 'aristocracy' of Italian football. Szymanski (2004), on the contrary, concentrates on wages; he associated payroll and the final standings of 27 Italian professional teams, from 1987 and 2001, and found a strong association between the two variables. Simmons and Forrest (2004) find that teams' performance in *Serie A* – from 1987 to 1999 – increased with relative payrolls, but at a decreasing rate.

With respect to the wages dispersion-performance nexus for the Italian context Bucciol et al. (2014) focus on teams' performance in *Serie A* from 2009 to 2011, collecting a unique dataset of 666 observations at single match level. They estimate a Probit model where the dependent variable is a dummy (1 if team wins), and the pay dispersion is approximated by the *Theil* index. Among covariates they consider team characteristics, the quality of the opponents, coaches' peculiarities and other controls. The novelty of their contribution is the different definition of team; they consider *i*) the *active team members* (ATM) which refers to the players that played at least one minute in the match (weighting for the number of minutes played); *ii*) the *unweighted* ATM; *iii*) the *potential* players, which refers to the starter players and substitute players; *iv*) the whole *roster*. Once differentiated, they calculate the *Theil* index for each definition of team obtaining, from the estimation, conflicting results with respect to the effect of wages dispersion on performance. In particular, the coefficient associated to pay dispersion is negative and statistically significant when the *Theil* index is calculated on *ATM*, both weighted and unweighted; is negative but not statistically significant if calculated on *potential* team, and is positive and statistically significant if calculated on *roster*. In addition, they also find that the negative effect of pay dispersion on performance can be ascribed to a worse individual performance rather than a reduction of team cooperation.

Wage dispersion and performances in Italy: the data

As noted above, we empirically test whether wage dispersion influences the teams' performances. In order to do this, we exploit a dataset on Italian *Serie A* from season 2007/2008 to 2015/16.[iv] Data on wages are drawn from *La Gazzetta dello Sport*. Since 2007, it releases data on players' wages at the beginning of the season. Note that data for the season 2008/2009 are not available. Given relegations and/or promotions to/from the second league (*Serie B*), we have 180 observations on 32 teams for the eight periods considered. It is worth noting that only 11 teams participated in all the seasons under investigation.

Following Franck and Nüesch (2011), the dependent variable, namely the seasonal performance, is the percentage of points achieved by each team at the end of the season, and is computed as follows:

$$points\_pct_{i,t} = \frac{points_{i,t}}{points_t},$$

where $points_{i,t}$ are the points achieved by team i in the season t, and $points_t$ are the maximum points achievable in the season t, so that given the maximum points attainable by each team at the end of the season (114 = 38 matches x 3 points each), *points_pct* denotes the ratio between seasonal cumulated points and 114. Note that we consider the count of points actually obtained by teams without taking into account penalties imposed by the League.[v]

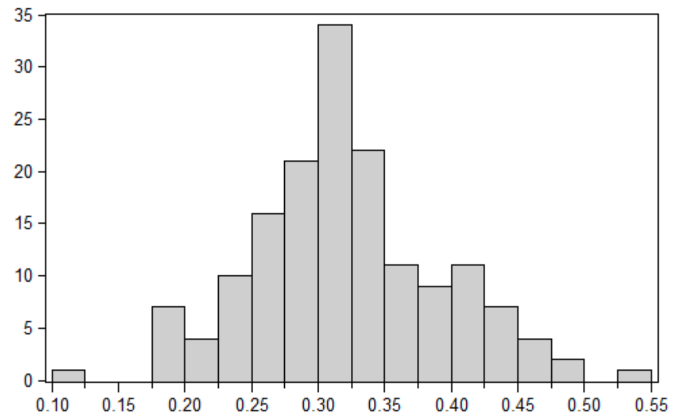
We include our dependent variable in two separate sets. The first relates to the measures of wage dispersion. We followed the main literature on the relationships between sport performances and wage dispersion using the *Gini* index as a measure of salary concentration (Coates et al., 2016; Yamamura, 2015). In addition and alternatively, we use two other measures of inequality: the *Shannon* and *Simpson* indexes. They are often employed in biology to measure the diversity within a group. Suppose that we have a population of N individuals of s species, and n_i is the number of individuals of species i. The *Shannon* and *Simpson* indexes are computed as follows:

$$Shannon = \left[-\left(\sum_{i=1}^{s} p_i \ln p_i\right)\right] / \ln(s),$$

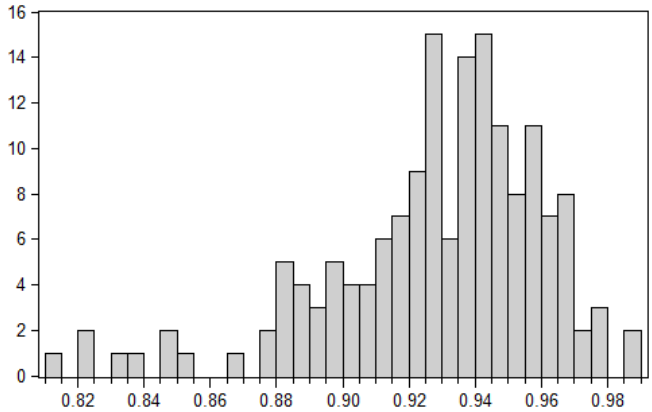
$$Simpson = 1 - \sum_{i=1}^{s} p_i^2,$$

where p_i is the relative abundance (n_i/N) of individuals of species i. Translating biological terms into the football field, if N is the number of players of a team, and s is the different wage levels paid to players, the *Shannon* and *Simpson* indexes give us information about the perceived diversity in terms of monetary salary among players. The two indexes are ranged between 0 and 1, and are increasing in diversity. That is, for example, a perfectly homogenous team in wage levels would have a score of 0. As the Simpson score increases the payroll is characterized by higher diversity. The frequencies distribution of the diversity scores is illustrated in the following graphs 1, 2 and 3.

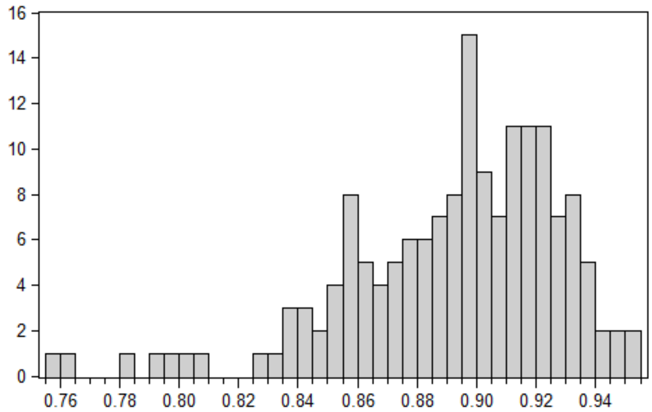
Graph 1. Frequencies distribution of Gini Index - 160 obs



Graph 2. Frequencies distribution of Shannon Index - 160 obs



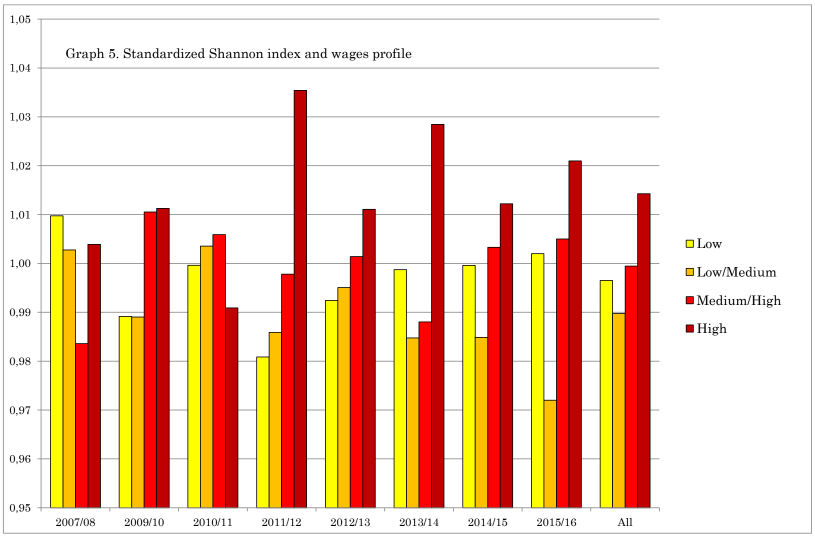
Graph 3. Frequencies distribution of Simpson Index - 160 obs

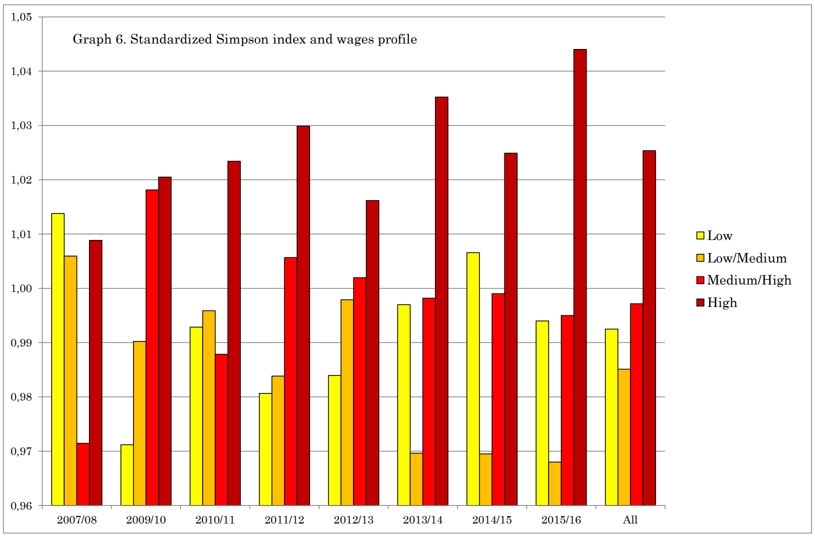


We start with a short visual analysis of the tendency of the three concentration/diversity variables, based on three steps. First we computed the standardized indexes by dividing team-related values for each seasonal average of associated *Gini*, *Shannon* and *Simpson* indexes.

The second step consists of taking into consideration the quartiles of teams' population based on the real wages levels. Needless to say, the first quartile (*Low*) includes the five teams with lower payrolls, the second quartile (*Low-Medium)* includes the five teams with payroll between 16th and 11th, the third quartile (*Medium-High)* includes the five teams with payroll ranged between 10th and 6th, and the last quartile (*High)* includes the teams with the highest five payrolls.

Notice that each quartile is to be considered an open group. In fact, a team included within a group in a certain season can be assigned to an alternative group if its payroll rank changes in time. For example, Fiorentina is included in the *High* real wages group in the first three seasons under observation; in the season 2011/12 that team was replaced by Lazio that was assigned to the *High* group for two seasons, 2011/12 and 2012/13. From 2013/14 on Lazio was replaced in the *High* group by Napoli. The last step consists of calculating the average standardized indexes for each group and season. The results are shown in the graphs 4, 5 and 6.





From the *All* columns we note that both concentration and diversity are higher in the *High* groups, but while the standardized *Gini* index increases with wages clusters, the same is not true for the diversity. If we concentrate on the different seasons, we observe a tendency of diversity to reduce in the medium wages groups, while in the *Low* groups diversity indexes are greater than we expected them to be.

Notice that what is relevant in our measures of diversity has nothing to do with wage levels, but relates only with the different salary concerns of management with respect to players' performance. Roughly speaking, what we are testing is if players' efforts, and then teams' performances, are sensitive to different wage treatment apart from the level of disparity. In the following table 1 we show descriptive statistics about winning teams and related concentration/diversity measures.

Table 1. Winning teams, points percentage and concentration/diversity measures. 2007-2016

	Team	points_pct	Gini	Gini avg	Shannon	Shannon avg	Simpson	Simpson avg
2007/08	Inter	0.746	0.287	0.310	0.918	0.927	0.870	0.883
2009/10	Inter	0.719	0.307	0.308	0.917	0.931	0.854	0.883
2010/11	Milan	0.719	0.396	0.321	0.910	0.923	0.885	0.895
2011/12	Juventus	0.737	0.358	0.329	0.946	0.920	0.928	0.899

2012/13	Juventus	0.763	0.340	0.327	0.953	0.925	0.916	0.899
2013/14	Juventus	0.895	0.328	0.308	0.959	0.927	0.926	0.891
2014/15	Juventus	0.763	0.350	0.320	0.966	0.941	0.929	0.901
2015/16	Juventus	0.798	0.424	0.340	0.951	0.930	0.918	0.886

It seems that there is an inconclusive association between the points percentage of winning teams and our diversity measures. According to *Gini*, a less concentrated (with respect to the seasonal average) payroll is associated with the winning team only for Inter in the seasons 2007/08 and 2009/10, while the opposite is true for the other seasons (for Milan and Juventus). With respect to *Shannon* and *Simpson* we note that winning team indexes are below the seasonal average for Milan and Inter, and above the average for Juventus. This preliminary investigation dictates that more refined investigation is required in order to identify regularities among sport performance and our covariates.

A more sophisticated analysis on the relationships between wages dispersion and sport performances may be based on the correlations among our measures of concentration/diversity and the dependent variable. In the following figures we report three scatter plots in which the seasonal points percentages are aligned on the Y-axis, while the three concentration/dispersion variables, the *Gini*, *Shannon* and *Simpson* indexes are on the X-axis.

Figure 1. Wages dispersion and sport performances: 2007/08-2015/16

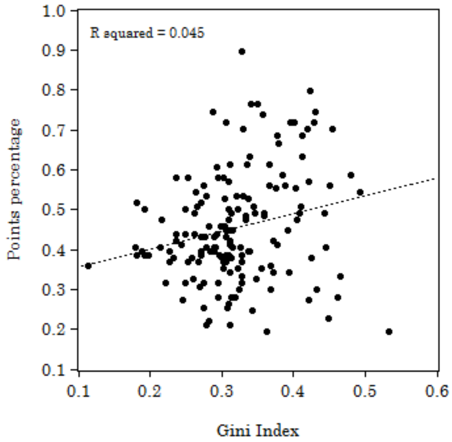


Figure 2. Wages dispersion and sport performances: 2007/08-2015/16

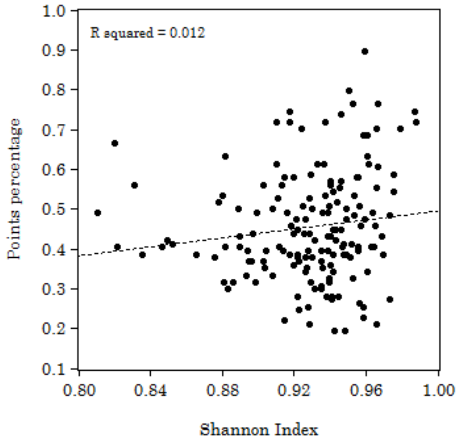
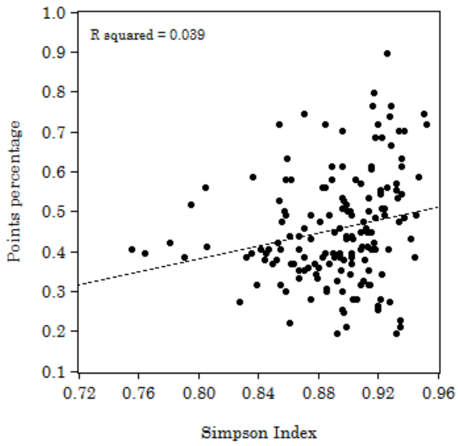


Figure 3. Wages dispersion and sport performances: 2007/08-2015/16



In the scatter plots a positive linear association among the points percentage and our measures of concentration/diversity seems to emerge. This result appears to be inconsistent with our belief, since we expect that the relationships between the sport performances and *Gini* index is of the opposite sign with respect to that associated with the other two measures of diversity, the *Shannon* and *Simpson* indexes, respectively. The low levels of R-squared[vi] reported on the top-left of scatter plots suggest that the concentration/diversity measures are not able to explain a consistent part of the variability of sport performances, if considered alone.

In the light of this, we develop the empirical analysis in two directions; first setting our data in a panel framework taking into account potential teams' fixed effects. Second, adding a set of covariates suggested by the established literature on sports performances (Scully, 1974; Kahn, 2000; Burger and Walters, 2003; Berri and Schmidt, 2010; Simmons and Berri, 2011). The first relates to the player talent as the main factor of sport performances, and wages its best approximation (Hall et al. 2002; Frick, 2007; García del Barrio and Szymanski, 2009).

The dataset includes data on wages provided by La Gazzetta dello Sport. Since conditional measures of performance are to be preferred, we follow Tao et al. (2016) in using the (log) seasonal relative position of team payroll, instead of (log) levels. We divided each payroll of team i in the period t by the average payroll in the season t, so obtaining *relative_wages*. The second covariate (*aristocracy*) proxies the history of the team in the *Serie A* (Di Betta and Amenta, 2010). We use the count of seasons in the top league of each team, including the season under investigation. The third explanatory variable is the average age of the team. It is calculated using data provided by the *Almanacco del Calcio – Panini* (hereafter *Almanacco* for sake of brevity), a yearly publication that provides detailed information about current and past Italian professional football organizations and results. The *Almanacco* also includes details about the teams playing in

Serie A and the composition of rosters. The data are those officially provided by the teams to the League at the end of October, that is after the end of the summer season players' transfer market window.

Following Bucciol et al. (2014) we also included the average age of each roster (*age*) among covariates. In addition, we also computed the square of the age (*age_squared*) searching for any non-linearity in the age structure. The fourth covariate is the town *population* of the city that hosts the team.[vii] Data refer to residents in the town population at the 31 December of the previous year and are provided by ISTAT. In addition, following one more time Bucciol et al. (2014) we introduce among the covariates a variable to capture the seniority of players in the roster. We first take into account the sum of the number of matches played by each player in the Italian first division, and eventually we divide it by the number of players in the roster, so obtaining a variable *experience*. Again, as suggested by Bryson et al. (2014) we introduce the variable *foreigners_pct* defined as the ratio between the number of foreign players and the number of the players in the roster. In fact, the share of foreign players may influence the wage structure. Lastly, a dummy variable *euro_cup* is to highlight those teams that are also involved in European tournaments.

Note that all covariates are logged, so that each coefficient in the regressions is to capture the punctual elasticity of the dependent variable with respect to the explanatory variables. Descriptive statistics of variables are found in table 2, and table 3 shows the correlations matrix of covariates.

Table 2. Descriptive statistics

	Obs.	Mean	Median	Std. Dev.	Min	Max
points_pct	180	0.455	0.430	0.138	0.193	0.895
relative_wages	180	1	0.628	0.868	0.200	3.661
aristocracy	180	46.656	49.500	26.739	1	84
age	180	26.609	26.560	1.231	23.340	30.110

foreigners_pct	180	0.461	0.463	0.159	0.038	0.846
experience	180	81.785	76.025	31.983	22.68	186.82
population	180	707,037	380,203	763,223	40,853	2,872,021
Gini	160	0.321	0.311	0.070	0.115	0.534
Shannon	160	0.928	0.935	0.033	0.811	0.988
Simpson	160	0.892	0.898	0.036	0.755	0.953

Table 3. Correlation Matrix of covariates

	relative_wages	*aristocracy*	*age*	*age_squared*	*foreigners_pct*	*Experience*	*population*	*Gini*	*Shannon*	*Simpson*
relative_wages	1	0.719	0.461	0.466	0.338	0.699	0.573	0.315	0.100	0.255
aristocracy		1	0.259	0.259	0.390	0.571	0.651	0.254	0.175	0.242
age			1	0.999	-0.012	0.632	0.264	-0.038	-0.051	0.117
age_squared				1	-0.012	0.639	0.226	-0.040	-0.054	0.115
foreigners_pct					1	0.050	0.409	0.179	0.044	0.138
experience						1	0.366	0.191	0.058	0.228
population							1	0.201	0.208	0.253
Gini								1	0.326	0.458
Shannon									1	0.702
Simpson										1

The econometric model and results

Our estimation strategy, for which the results are reported in table 4, consists of a baseline model introducing all the covariates (1 and 2), and alternatively the covariates associated to wage dispersion (models from 3 to 8). We then estimated the following equation:

$$\log(points\_pct_{i,t}) =$$
$$\alpha + \beta_1 \log(X_{i,t}) + \beta_2 \log(relative\_wages_{i,t}) + \beta_3 \log(aristocracy_{i,t}) +$$
$$\beta_4 \log\left(age_{i,t}\right) + \beta_5 [\log(age)]^2 + \beta_6 \log(population_{i,t}) +$$
$$\beta_7 \log\left(experience_{i,t}\right) + \beta_8 \log(foreigners\_pct_{i,t}) + \beta_9 (euro\_cup) + \mu_{i,t} + \varepsilon_{i,t},$$

where $X_{i,t}$ indicates alternatively the *Gini*, the *Shannon* and the *Simpson* indexes, α is the intercept term, $\mu_{i,t}$ and $\varepsilon_{i,t}$ are the between-entity and the within-entity errors respectively, and the β_i's are the coefficients associated to each covariate. We performed both random (RE) and fixed effects (FE) models with standard errors – adjusted for clusters in teams – taking into account potential heteroskedasticity across teams, computing the Hausman test in order to signal the model to be preferred. Results of the Hausman tests for the models (7 and 8), those where the coefficients associated to concentration/diversity measures, are statistically significant (chi squared value 13.41 with p-value of 0.161), indicate that the RE estimator has to be preferred to the FE; so that our comments concentrate on the RE model results.

Table 4. Panel data regression. Dependent variable: log *points_pct*

	1	2	3	4	5	6	7	8
	RE	FE	RE	FE	RE	FE	RE	FE
Constant	-29.999	-24.006	-40.978	-35.096	-36.841	-29.715	-41.530	-36.726
	(72.001)	(70.467)	(65.570)	(65.332)	(69.527)	(66.514)	(72.095)	(71.106)
log *Gini*			-0.126	-0.089				
			(0.094)	(0.128)				
log *Shannon*					-0.363	-0.033		
					(0.481)	(0.436)		
log *Simpson*							**-0.661****	**-0.539***
							(0.310)	**(0.274)**
log *relative_wages*	**0.270*****	**0.216****	**0.274*****	**0.199***	**0.268*****	**0.193***	**0.273*****	**0.203***
	(0.040)	**(0.103)**	**(0.041)**	**(0.111)**	**(0.040)**	**(0.112)**	**(0.042)**	**(0.109)**
log *aristocracy*	0.023	-0.033	0.019	0.012	0.026	0.012	0.028	0.026
	(0.023)	(0.229)	(0.023)	(0.243)	(0.023)	(0.245)	(0.042)	(0.243)
log *age*	17.914	20.812	24.417	27.231	21.823	23.629	24.670	27.364

	(43.571)	(44.460)	(39.804)	(41.963)	(42.176)	(42.708)	(43.759)	(45.315)
log *age_squared*	-2.759	-3.109	-3.737	-4.054	-3.318	-3.488	-3.753	-4.072
	(6.641)	(6.753)	(6.038)	(6.378)	(6.395)	(6.486)	(6.639)	(6.879)
log *population*	-0.013	-0.877	-0.009	-0.878	-0.006	-0.840	-0.012	-0.773
	(0.023)	(0.653)	(0.026)	(0.693)	(0.027)	(0.641)	(0.025)	(0.673)
log_*experience*	0.060	-0.030	0.057	-0.051	0.036	-0.061	0.049	-0.044
	(0.056)	(0.080)	(0.062)	(0.091)	(0.058)	(0.084)	(0.058)	(0.087)
log_*foreigners_pct*	0.037	0.032	0.027	0.002	0.015	-0.004	0.025	0.003
	(0.048)	(0.071)	(0.052)	(0.092)	(0.051)	(0.089)	(0.048)	(0.087)
euro_cup	0.054	0.019	0.064	0.022	0.051	0.019	0.051	0.016
	(0.042)	(0.042)	(0.049)	(0.048)	(0.044)	(0.045)	(0.045)	(0.046)
Observations	180	180	160	160	160	160	160	160
Cross sections	32	32	32	32	32	32	32	32
Periods min/max	1/9	1/9	1/8	1/8	1/8	1/8	1/8	1/8
R_squared overall	0.578	0.258	0.573	0.276	0.565	0.275	0.569	0.262
R_squared between	0.812	0.383	0.815	0.398	0.799	0.400	0.805	0.383
R_squared within	0.053	0.075	0.049	0.075	0.046	0.069	0.055	0.077
sigma_u	0.074	1.143	0.059	1.118	0.078	1.081	0.072	0.981
sigma_e	0.194	0.194	0.197	0.197	0.198	0.198	0.197	0.197
Rho	0.128	0.972	0.082	0.970	0.134	0.968	0.117	0.961
Wald chi_squared	239.45		200.07		178.96		161.62	
p-value	0.000		0.000		0.000		0.000	
F Stat		1.98		1.76		1.94		2.67
p-value		0.082		0.118		0.083		0.020
Level of statistical significance: ***1%, **5%, *10%. Random Effect: Swamy-Arora estimator; standard errors adjusted for 32 clusters in teams (in parenthesis). Fixed Effect: standard errors adjusted for 32 clusters in teams.								

In general, the Wald test strongly rejects the hypothesis that all coefficients in the model are zero. The overall R-squared in the RE models ranges between 0.565 and 0.578; our model is able to explain almost 80% of the variability of the dependent variable “between” teams, while “within” explained variability is low, and ranged between 0.046 and 0.055.

Considering the variables associated to the concentration and/or diversity of salary distribution, we note that the *Gini* index is not significant, and the same is true for *Shannon*. Only the *Simpson* index is statistically significant at 5% and presents a negative sign. The result suggests that as diversity increases the sport performance decreases. In particular, variables are expressed in logs, so that each associated coefficient measures the elasticity of the dependent variable with respect to the covariates, we note that a one per cent increase in the diversity index reduces the points percentage by 0.66 per cent. For example, if the points percentage and Simpson index are both at their average value, respectively 0.455 and 0.892, an increase of the Simpson index to 0.901 (namely by 1 per cent) will reduce the points percentage to 0.452.

This result contradicts the outcome of Bucciol et al. (2014) where the diversity of wages is measured at the roster level, and its relationship with performance is statistically significant and positive. Moreover, from the results of table 4 it emerges that *relative_wages* is the only variable statistically significant in all estimations. The magnitude and significance of the associated coefficients are independent from the model specification. In particular, the relationship between points percentage and relative wages is inelastic, and ranges around 0.27. This means that, on average, an increase of 1% in relative wages increases the percentage of points achieved by about 0.27%. The size of the coefficients also suggests that the negative impact of concentration seems to balance the beneficial impact of a higher level of talent. Yet the coefficients of all the other covariates do not reach the threshold for statistical significance. The latter evidence strongly confirms in particular that the main driver of sport performance is the availability of talent.

Conclusions

Economists have long recognized and studied the relationship between wages disparity and individual efforts when working in teams. This topic attracts even stronger interest if we consider the implication of these issues on firms' performance. In this study we investigate the effect of diversity on the performance of football teams in the Italian *Serie A* between the seasons 2007/2008 and 2015/2016. In line with a growing literature, we show that diversity in payroll is associated with lower teams' performance. In particular, we found that a one per cent increase in the Simpson diversity index reduces the points percentage by 0.66 per cent. In addition, real wages is the only statistically significant variable in all the different specifications. The empirical estimation shows that on average, an increase of 1% of relative wages increases the percentage of points achieved by about 0.27%.

In sum, alongside the main result, another relevant outcome we would claim for this paper is that different measures of diversity may lead to different results. In fact, two measures appear not to be significantly associated with the dependent variable, so leaving an open question on the appropriate quantitative measure to capture this aspect. Secondly, our findings confirm that sport success is heavily dependent on the wages of the players. In fact, the real wages appear to be statistically insignificant. In brief, sport performance depends heavily on relative talent and on within-team diversity in payoffs of players.

References

Akerlof, G.A., and Yellen, J.L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, 105(2), 255-283.

Almanacco del Calcio, Panini, Modena. Various years.

Annala, C.N., and Winfree, J. (2011). Salary distribution and team performance in Major League Baseball. *Sport Management Review*, 14(2), 167-175.

Avrutin, B.M., and Sommers, P.M. (2007). Work incentives and salary distribution in major league baseball. *Atlantic Economic Journal*, 35(4), 509–510.

Baroncelli A. and Caruso R. (2011). The Organization and Economics of Italian Top Football, in Gammelsaeter H., Senaux B., (eds.), *The Organisation and Governance of Top Football across Europe*, Routledge, London.

Berri, D.J., and Jewell, R.T. (2004). Wage inequality and firm performance: Professional basketball's natural experiment. *Atlantic Economic Journal*, 32(2), 130-139.

Berri, D.J., and Schmidt, M.B. (2010). *Stumbling on wins: two economists expose the pitfalls on the road to victory in professional sports*. Upper Saddle River, NJ, FT Press.

Bloom, M. (1999). The performance effect of pay dispersion on individuals and organizations. *Academy of Management Journal*, 42(1), 25-40.

Bose, A., Pal, D., and Sappington, D. (2010). Equal pay for unequal work: limiting sabotage in teams. *Journals of Economics & Management Strategy*, 19(1), 25-53.

Breunig, R., Garrett-Rumba, B., Jardin, M., and Rocaboy, Y. (2014). Wage dispersion and team performance: a theoretical model and evidence from baseball. *Applied Economics*, 46(3), 271-281.

Bryson, A., Rossi G. and R. Simmons (2014). The migrant wage premium in professional football: a superstar effect?. *Kyklos*, 67(1), 12-28.

Bucciol, A., Foss, N.J. and Piovesan, M. (2014). Pay Dispersion and Performance in Teams. *Plos One*, 9(11), pp. 1-16.

Burger, J.D. and Walters, S.J.K. (2003). Market size, pay and performance. A general model and application to Major League Baseball. *Journal of Sports Economics*, 4(2), 108-125.

Castellanos García, P., Dopico Castro, J.A., and Sánchez Santos, J.M. (2007). The economic geography of football success: empirical evidence from European cities. *Rivista di Diritto ed Economia dello Sport*, 3(2), 67-88.

Coates, D., Frick, B., and Jewell, T. (2016). Superstar salaries and soccer success. *Journal of Sports Economics*, 17(7), 716-735.

Cubel M., Sanchez-Pages S. (2012). The effect of within-group inequality in a conflict against a unitary threat, *Peace Economics, Peace Science and Public Policy*, 18(3), DOI: 10.1515/peps-2012-0015,

De Brock, L., Hendricks, W., and Koenker, R. (2004). Pay and performance – The impact of salary distribution on firm-level outcomes in baseball. *Journal of Sport Economics*, 5(3), 243-261.

Depken, C.A. (2000). Wage disparity and team productivity: Evidence from major league baseball. *Economics Letters*, 67(1), 87-92.

Di Betta, P. and Amenta, C. (2010). A die-hard aristocracy: competitive balance in Italian soccer, 1929-2009. *Rivista di Diritto ed Economia dello Sport*, 6(2), 13-39.

Di Domizio, M. (2008). Localizzazione geografica e performance sportiva: un'analisi empirica sul campionato di calcio di Serie A. *Rivista di Diritto ed Economia dello Sport*, 4(3), 105-127.

El-Hodiri, M. and Quirk, J. (1971). An economic model of a professional sports league. *Journal of Political Economy*, 79(6), 1302-1319.

Fort, R. and Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, 33(3), 1265-1299.

Franck, E., and Nüesch, S. (2010). The effect of talent disparity on team productivity in soccer. *Journal of Economic Psychology*, 31(2), 218-229.

Franck, E., and Nüesch, S. (2011). The effect of wage dispersion on team outcome and the way the outcome is produced. *Applied Economics*, 43(23), 3037-3049.

Frick B., Prinz J. and Winkelman K. (2003). Pay inequalities and team performance: Empirical evidence from the North American Major Leagues. *International Journal of Manpower*, 24(4), 472-488.

Frick, B. (2007). The football players labour market: empirical evidence from the major European leagues. *Scottish Journal of Political Economy*, 54(3), 422-446.

Frick, B. (2013). Team wage bills and sporting performance: evidence from (major and minor) European football leagues. In Rodríguez, P., S. Késenne and J. García (eds), *The Econometrics of Sport*, Edward Elgar, Cheltenham (UK).

García del Barrio, P., and Szymanski, S. (2009). Goal! Profit maximization versus win maximization in soccer. *Review of Industrial Organization*, 34(1), 45-68.

Hall, S., Szymanski, S., and Zimbalist, A. S. (2002). Testing causality between team performance and payroll. The cases of Major League Baseball and English soccer. *Journal of Sports Economics*, 3(2), 149-168.

Hoffmann, R., Lee, C.G., and Ramasamy, B. (2002). The socio-economic determinants of international soccer performance. *Journal of Applied Economics*, 5(2), 253-272.

Jewell, R.T., and Molina, D.J. (2004). Productivity and salary distribution: The case of US major league baseball. *Scottish Journal of Political Economy*, 51(1), 127-142.

Kahane, L.H. (2012). Salary dispersion and team production: Evidence from the national hockey league. In Shmanske, S., and Kahane, L.H. (Eds.), *The Oxford Handbook of Sports Economics Volume 2: Economics Through Sports*. New York, NY, Oxford University Press.

Kahn, L.M. (2000). The sport business as a labor market laboratory. *Journal of Economic Perspectives*, 14(3), 75-94.

Katayama, H. and Nuch, H. (2011). A game-level analysis of salary dispersion and team performance in the national basketball association. *Applied Economics*, 43(10), 1193-1207.

Késenne, S. (2000). Revenue sharing and competitive balance in professional team sports. *Journal of Sports Economics*, 1(1), 56-65.

Lazear E.P. and Rosen S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89(5), 841-864.

Lazear, E.P. (1989). Pay equity and industrial politics. *Journal of Political Economy*, 97(2), 561-580.

Levine, D.I. (1991). Cohesiveness, productivity, and wage dispersion. *Journal of Economic Behavior & Organization*, 15(2), 237-255.

Milgrom, P.R. (1988). Employment contracts, influence activities, and efficient organization design. *Journal of Political Economy*, 96(1), 42-60.

Mondello, M. and Maxcy, J. (2009). The impact of salary dispersion and performance bonuses in NFL organizations. *Management Decision*, 47(1), 110-123.

Patil, G.P. and Taillie, C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association*, 77(379), 548-561.

Ramaswamy, R., and Rowthorn, R.E. (1991). Efficiency wage and wage dispersion. *Economica*, 58(232), 501-541.

Richards, D.G., and Guell, R.C. (1998). Baseball success and the structure of salaries. *Applied Economics Letters*, 5(5), 291-296.

Rodríguez, P., Késenne, S. and García J. (2013). *The Econometrics of Sport*, Edward Elgar, Cheltenham (UK).

Scully, G. (1974). Pay and performance in Major League Baseball. *American Economic Review*, 64(6), 915-930.

Simmons, R. and Forrest, D. (2004). Buying success: team performance and wage bills in U.S. and European sports leagues. In Fort, R. and Fizel, J. (Eds.), *International Sports Economics Comparisons*. Westport (CT), Praeger.

Simmons, R. and Berri, D.J. (2011). Mixing the princes and the paupers: pay and performance in the National Basketball Association. *Labour Economics*, 18(3), 381-388.

Sommers, P.M. (1998). Work incentives and salary distributions in the national hockey league. *Atlantic Economic Journal*, 26(1), 119-119.

Stogdill, R.M. (1972). Group productivity, drive, and cohesiveness. *Organizational behavior and human performance*, 8(1), 26-43.

Szymanski, S. (2004). La relazione competitive tra posizione competitiva e posizione reddituale: quali sono le squadre migliori? In Lago, U., Baroncelli, A. and Szymanski, S. (Eds.), *Il business del calcio. Successi sportivi e rovesci finanziari*. Milano, EGEA.

Szymanski, S. (2013). Wages, transfers and variation of team performance in the English Premier League. In Rodríguez, P., Késenne, S. and García, J. (2013), *The Econometrics of Sport*, Edward Elgar, Cheltenham (UK), 53-62.

Szymanski, S. and Késenne, S. (2004). Competitive balance and gate revenue sharing in team sports. *Journal of Industrial Economics*, 52(1), 165-177.

Szymanski, S. and Smith, R. (1997). The English football industry: profit, performance and industrial structure. *International Review of Applied Economics*, 11(1), 135-153.

Tao, Y.L., Chuang, H.L. and Lin, E.S. (2016). Compensation and Performance in Major League Baseball: Evidence from Salary Dispersion and Team Performance. *International Review of Economics & Finance*, 43, 151-159.

Vrooman, J. (1995). A general theory of professional sports league. *Southern Economic Journal*, 61, 971-990.

Wiseman, F. and Chatteriee, S. (2003). Team payroll and team performance in major league baseball: 1985-2002. *Economics Bulletin*, 1(2), 1-10.

Yamamura, E. (2015). Wage disparity and team performance in the process of industry development: evidence from Japanese Professional Football League. *Journal of Sports Economics*, 16(2), 214-223.

---

[i] See among others Bose et al. (2010), Milgrom (1988).

[ii] On the differences between the three measures please see Patil and Taillie (1982).

[iii] See among others Fort and Quirk (1995), Vrooman (1995), Késenne (2000), Szymanski and Késenne (2004), Szymanski and Smith (1997), Hall et al. (2002), Burger and Walters (2003), Frick (2007), Berri and Schmidt (2010), Frick (2013), Rodríguez et al. (2013), and Szymanski (2013).

[iv] For details on Italian professional football see Baroncelli and Caruso (2011).

[v] In fact, some teams have been fined: Bologna (3 points in the season 2010/11), Atalanta (6 points and 2 points, respectively, in 2011/12 and 2012/13), Sampdoria (1 point in 2012/13), Siena (6 points in 2012/13), and Torino (1 point in 2012/13).

[vi] Note that the reported R-squared in the figures comes from an OLS estimation with constant term.

[vii] On the role of population size in determining sport success see, among others, Di Domizio (2008), Castellanos García et al. (2007), and Hoffmann et al. (2002).