

Do high wage footballers play for high wage teams?

Rachel Scarfe

Carl Singleton

Paul Telemo*

December 2018

[Link to latest version](#)

Abstract

Intuition and sports knowledge suggest the best professional footballers play for the best teams, i.e. positive assortative matching between employer and employee on productivity. We use wage data for all players and teams in Major League Soccer between 2007 and 2017 and find that estimated player and team fixed wage effects are negatively correlated. This is a puzzle, which could be explained if players match to teams according to some compensating wage differential, for example from a desire to play for successful teams. The estimated wage premiums of teams are highly and negatively correlated with their success in the league (productivity).

Keywords: firm-specific wages, AKM wage equation, matching, superstar pay

JEL codes: J31, J49, Z22

*rachel.scarfe@ed.ac.uk & paul.telemo@ed.ac.uk: School of Economics, University of Edinburgh, UK.
c.a.singleton@reading.ac.uk: Department of Economics, University of Reading, UK.

We are grateful for helpful comments from J. James Reade and Daniel Schaefer, and for research assistance from Leon Singleton. The data used in this study are publicly available from Major League Soccer (MLS) LLC and the MLS Player's Association.

Singleton thanks the Economic and Social Research Council (UK) for funding support under Grant No. ES/J500136/1.

1 Introduction

The question of whether high productivity workers are matched to high productivity firms is of considerable interest, not least because it sheds light on the matching process taking place between workers and firms (e.g. [Becker, 1973](#); [Shimer and Smith, 2000](#)). To address this, studies have used longitudinal linked employer-employee data to estimate AKM wage equations, which admit both worker and firm fixed effects following [Abowd et al. \(1999\)](#): the sign and magnitude of the correlation between these two sets of estimated effects may indicate whether a labour market features assortative matching of worker and firm productivity types (e.g. [Andrews et al., 2012](#)).

In this note, we estimate an AKM-type wage equation using data on the earnings of professional football players in US Major League Soccer (MLS) between 2007 and 2017. The market for football talent is particular, with high levels and variance of earnings. It is generally closed off from the rest of the labour market, and has its own rules and regulations.¹ However, there are advantages of studying this market. First, we observe the salaries of almost all players rather than a limited sample. Second, our data includes worker and firm productivity measures (e.g. minutes played and league performance).

We find that the correlation between estimated player (worker) and team (firm) fixed wage effects is small and negative. High wage footballers do not appear to play for high wage teams. Yet intuitively, and based on sports knowledge, we would expect that the best footballers do in fact play for the best teams.

There are potential ways to reconcile these facts. The first is negative ‘limited mobility bias’ in the estimated correlation of the player and team fixed effects ([Andrews et al., 2008](#)). However, the data we use are approximately a universe of players and teams, and we show this bias is not an issue. Second, it may be difficult to identify assortative matching using wage data alone (e.g. [Eeckhout and Kircher, 2011](#)). There are specific reasons why this might be the case in a football talent market. First, salary regulation could affect hiring decisions and the distribution of surplus when a player and team match. Second, footballers want to play for successful (productive) teams. Players who switch to more successful teams may be willing to accept relatively low wages. There is evidence to suggest this as a plausible explanation: the correlation between the estimated team fixed effects and a measure of a team success is large and negative in MLS.

2 Data

Player wage data and team affiliations come from the MLS Players’ Association (MLSPA), the collective bargaining representative of players during the negotiation of league-wide salary rules. The data refer to midway through the MLS season in August of each year, after the secondary transfer window when players can be signed from abroad. The wage measure is the guaranteed annualized compensation or salary, henceforth referred to as wages. This includes payments for signing with a team or related to

¹See for example on the MLS case [Kuethe and Motamed \(2010\)](#). The latest MLS talent market regulations: <https://www.mlssoccer.com/.../mls-roster-rules-and-regulations>.

marketing, but does not include performance-related pay.² Figure 1 tracks percentiles of player log wages between 2007 and 2017. This was a period of rapid pay growth in MLS, with nominal wages more than doubling across the distribution, and an increase in the median of 160%. This coincided with the growing popularity of the league, higher revenues and more foreign players choosing to play in the US (Brownlee and Lorgnier, 2017). The league expanded in this period, from 13 teams in 2007 to 22 in 2017.³

There are detailed regulations concerning pay in MLS. Like other US sports leagues there is a salary cap, but this only covers the first 18-20 of up to 30 players on a team's squad. Teams can hire so-called designated players, whose wages are effectively outside the salary cap. In most cases these are the highest paid players in a team. Despite these and other regulations, in reality it appears that teams are operating to a budget rather than an effective league-wide cap. For example, Figure 2 shows the guaranteed salary totals for teams in the 2017 season. There is substantial variation across teams, ranging from \$5million for the lowest budget teams to over \$20million for the highest.

We obtain information on player age, the number of games started and minutes played during a season, designated player status and team-level performance data from the MLS official website.⁴ A handful of observations are dropped, where we could not match a record in the salary data. We also exclude any observations where the player did not play for their team in that season, thus ignoring any players who were either continuously injured or making up weight in a team's squad. We are left with 4,213 player-year observations. The minimum, median and maximum ages of players in the analysis are 15, 25 and 42, respectively. Almost 7% of player-year observations are designated players. The median number of teams played for over player-year observations is 2. Just under 30% of players represented played for more than one team in 2007-2017.

3 Method

Let there be $i = 1, \dots, P$ players, $j = 1, \dots, J$ teams, $t = 1, \dots, T$ years, T_i years per player and $N = \sum_i T_i$ total observations. The AKM-type wage equation is given by:

$$w_{it} = \alpha_i + \phi_{J(it)} + \mathbf{x}'_{it} \boldsymbol{\beta} + \varepsilon_{it} , \quad (1)$$

where w_{it} is the log wage and α_i is the player fixed effect. Team fixed effects with mean zero are given by $\phi_{J(it)}$, where $J(it) = j$ indicates the team employing player i in year t . \mathbf{x}_{it} is a vector of time-varying observable characteristics. The remaining heterogeneity is in the residual, ε_{it} . Equation (1) can be estimated using least squares with a strict exogeneity assumption: $E[\varepsilon_{it} | \mathbf{x}_{it}, \alpha_i, \phi_{J(it)}] = 0$. The

²According to the league's annual bargaining agreements, team-level performance payments are a tiny fraction of overall salary.

³The analysis covers 23 teams, since one was dissolved in 2014; 2007: 13, 2008: 14, 2009: 15, 2010: 16, 2011: 18, 2012-2014: 19, 2015-2016: 20, 2017: 22.

⁴<https://www.mlssoccer.com/stats>

estimated team fixed effects will be biased if players change teams due to some component of ε_{it} , such as transitory shocks to team-wide or player-team match-specific wage effects.⁵

The player fixed effects are transferable, affecting wages wherever and whenever a player features in MLS. The team fixed effects are wage premiums received by all players within a team, identified by team switching. Therefore, estimates of these premiums are only comparable within connected sets of players and teams (Abowd et al., 2002). As we observe the salaries of almost all players during their time in MLS, in our main results all players and the 23 teams are connected in a single set.

We consider the correlation between the estimated player and team fixed effects from Equation (1), $Corr_{it}(\hat{\alpha}_i, \hat{\phi}_{J(it)})$. This measure is negatively biased because the fixed effects are estimated with errors (Andrews et al., 2008). This ‘limited mobility bias’ is large in small sample settings and related to the amount of observed mobility. Andrews et al. (2008, 2012) predict and demonstrate empirically that the estimated value of $Corr_{it}(\hat{\alpha}_i, \hat{\phi}_{J(it)})$ will be increasing and concave in the number of observed movers per firm, asymptoting towards its true value.

4 Results

Column (1) of Table 1 shows our preferred estimates of Equation (1). \mathbf{x}_{it} contains the following covariates: a dummy indicating designated players; player age minus 15, this value squared and cubed, to address football-life-cycle wage growth; and year effects, accounting for average annual wage growth.

The fit of the wage equation is high with an $R^2 = 0.91$, which increases to 0.95 when we estimate a version with player-team fixed effects rather than separate player and team effects, suggesting the assumed additive separability in Equation (1) is a workable assumption. The designated player premium is large and significant at 65 log points, identified from players who change status. The coefficients on the age terms imply that on average a player’s wages grow until he is 27 and decline steeply thereafter.

The estimated team fixed effects, $\hat{\phi}_j$, are shown in Figure 3. Relative to the average player-year value, these team wage premiums range from -23 log points to 18 log points. The correlation between the worker and team fixed effects is -0.10, suggesting that high wage footballers do not play for high wage teams. One potential explanation for this result, which would be consistent with the intuition that high productivity players in fact do play for high productivity teams, is the existence of some non-pecuniary attraction to playing for high productivity teams. In this case, the ‘best’ players might be willing to accept a lower team wage premium in order to play for the ‘best’ teams. To check for evidence of this, we look at the correlation over teams between $\hat{\phi}_j$ and the average over seasons of the league points achieved per game.⁶ This correlation is -0.52, suggesting high wage premium teams have had less success in the league.

⁵See Card et al. (2018) for a general discussion of this assumption in reality and practice.

⁶Teams earn 3 points for a win, 1 for a draw and 0 for a loss. Cumulative points determine final league standings and qualification for playoffs.

Despite using approximately the universe of players in this period, we nonetheless check that limited mobility is not biasing this estimated correlation downwards. To do this we drop at random 10%, 20% . . . 90% of the observations and re-estimate Equation (1) for the remaining largest connected sets of players and teams. In Figure 4 we plot the correlation from each of these smaller samples against the number of observed player moves per team, M/J . As per Andrews et al. (2012), in smaller samples the negative bias is substantial, and decreasing with M/J , asymptoting to the figure of approximately -0.1 we found in the 100% sample.

For robustness we estimate three alternative versions of Equation (1). In Column 2 of Table 1, we show results using all player-year observations rather than only those who played during a season. Column 3 drops designated players from the sample used in the preferred results. Similarly, column 4 drops any players who featured in MLS for less than 3 years of the sample period, to focus a subset with higher mobility. Column 5 uses as the dependent variable the log of wages per minute played during the season. Across all these specifications the main result on $Corr_{it}(\hat{\alpha}_i, \hat{\phi}_{J(it)})$ remains robust.

5 Conclusion

High wage professional footballers do not play for high wage teams, at least as far as MLS is concerned, despite common knowledge that the best players play for the best teams. This is a particular labour market, and this result may not even generalize to other football talent markets with fewer salary regulations, such as in Europe. However, the result provides an example which suggests wage data alone are insufficient to identify positive assortative productivity matching where intuitively it should exist.

References

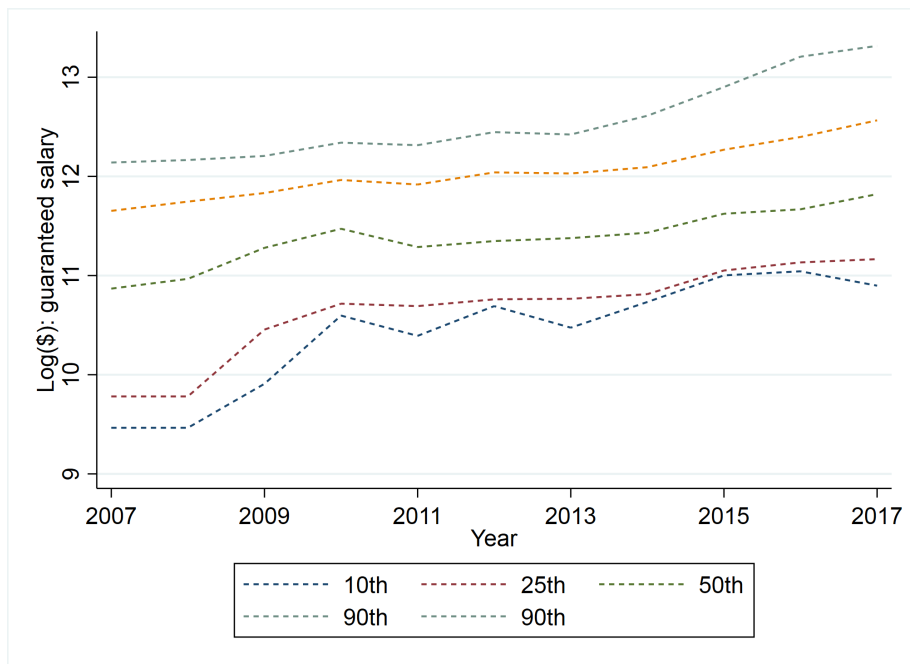
- Abowd, J., R. H. Creecy, and F. Kramarz.** 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Longitudinal Employer-Household Dynamics Technical Papers, Center for Economic Studies, U.S. Census Bureau.
- Abowd, J., F. Kramarz, and D. Margolis.** 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–334.
- Andrews, M. J., L. Gill, T. Schank, and R. Upward.** 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?." *Journal of the Royal Statistical Society Series A*, 171(3): 673–697.
- Andrews, M., L. Gill, T. Schank, and R. Upward.** 2012. "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias." *Economics Letters*, 117(3): 824–827.
- Becker, G.** 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–46.
- Brownlee, E., and N. Lorgnier.** 2017. "A case study of Major League Soccer." *Routledge Handbook of Football Marketing*, p. 73.
- Card, D., A. R. Cardoso, J. Heining, and P. Kline.** 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.
- Eeckhout, J., and P. Kircher.** 2011. "Identifying Sorting – In Theory." *Review of Economic Studies*, 78(3): 872–906.
- Kuethe, T., and M. Motamed.** 2010. "Returns to Stardom: Evidence From U.S. Major League Soccer." *Journal of Sports Economics*, 11(5): 567–579.
- Shimer, R., and L. Smith.** 2000. "Assortative Matching and Search." *Econometrica*, 68(2): 343–370.

TABLE 1: Regression model estimation results for log player wages, 2007-2017

	(1)	(2)	(3)	(4)	(5)
Designated player	0.652 (0.109)	0.658 (0.108)		0.655 (0.112)	0.536 (0.108)
Age (years)	-0.070 (0.053)	-0.124 (0.041)	-0.073 (0.052)	0.093 (0.059)	-0.411 (0.084)
Age squared	0.012 (0.005)	0.016 (0.004)	0.012 (0.005)	0.013 (0.005)	0.030 (0.007)
Age cubed	-0.0006 (0.0001)	-0.006 (0.0001)	-0.0005 (0.0001)	-0.0005 (0.0001)	-0.0008 (0.0001)
Player, team & year effects	Yes	Yes	Yes	Yes	Yes
R^2	0.913	0.922	0.862	0.849	0.685
RMSE	0.349	0.342	0.341	0.361	0.732
St. dev. of log wages - $Std(w_{it})$	0.959	0.981	0.743	0.929	1.057
St. dev. player effects - $Std_{it}(\hat{\alpha}_i)$	0.893	0.909	0.765	0.723	0.965
St. dev. team effects - $Std_{it}(\hat{\phi}_{J(it)})$	0.092	0.091	0.089	0.097	0.146
St. dev. observables - $Std(\mathbf{x}'_{it}\hat{\beta})$	0.601	0.591	0.551	0.673	0.728
Correlation - $Corr_{it}(\hat{\alpha}_i, \hat{\phi}_{J(it)})$	-0.102	-0.097	-0.084	-0.110	-0.121
N : player-years	4,213	5,195	3,906	3,138	4,213
P : players	1,410	1,820	1,320	606	1,140
J : teams	23	23	23	23	23
M/J : mobility	27	30	26	25	27

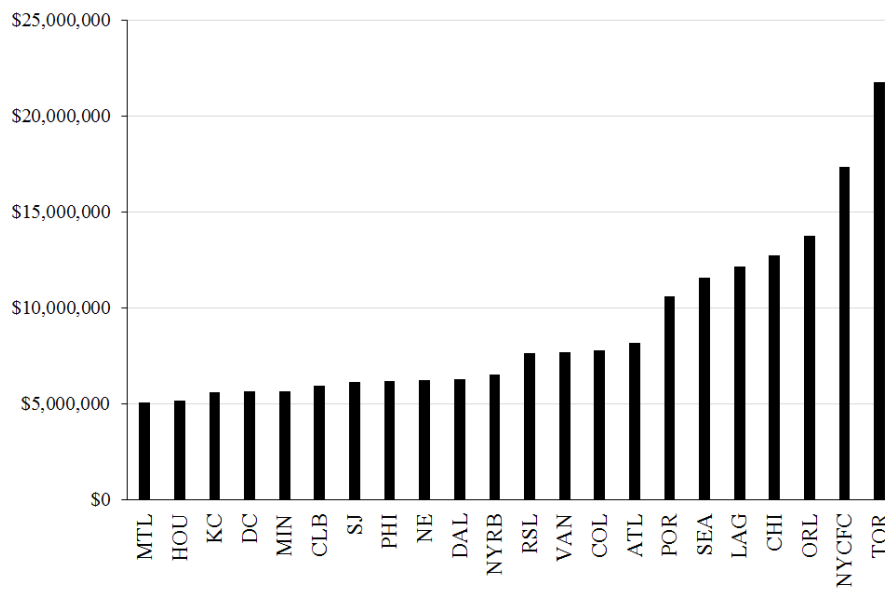
Notes.- author calculations using MLS and MLSPA data. Values in parentheses give standard errors robust to player-level clusters.

FIGURE 1: Percentiles of MLS player log (guaranteed) wages, 2007-2017



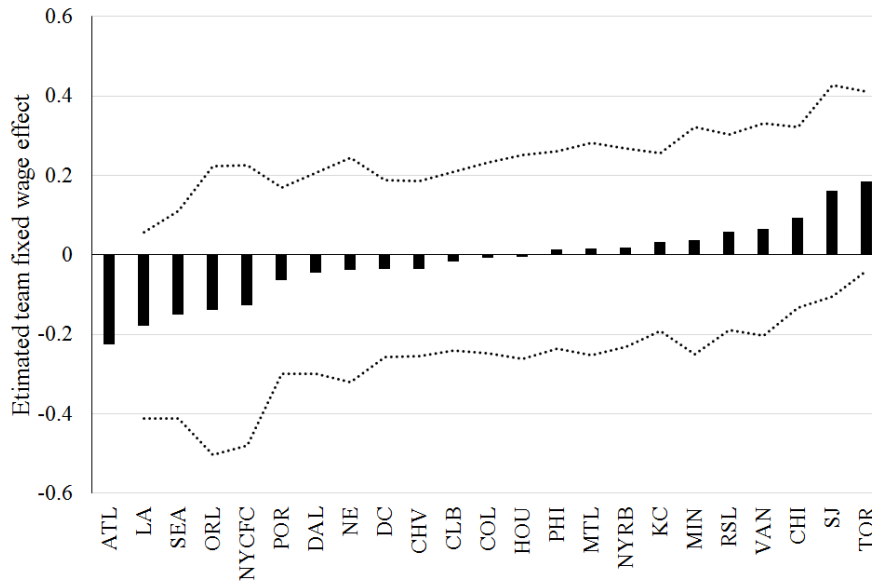
Notes.- author calculations using MLSPA data.

FIGURE 2: Team guaranteed salary totals for the 2017 MLS season



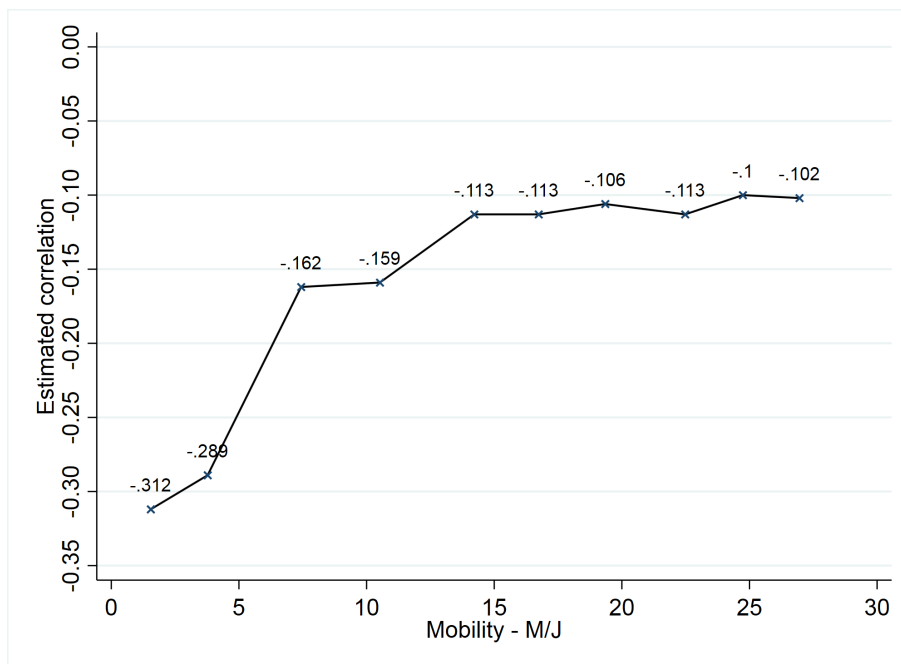
Source.- MLSPA

FIGURE 3: Estimated relative team log wage premiums, $\hat{\phi}_j$



Notes.- author calculations using MLS and MLSPA data. Estimates relate to Column 1 of Table 1. By construction the average wage premium over player-year observations is zero. The dashed lines show 95% confidence intervals.

FIGURE 4: Limited mobility bias and the estimated correlation between player and team fixed wage effects, $Corr_{it}(\hat{\alpha}_i, \hat{\phi}_{J(it)})$



Notes.- author calculations using MLS and MLSPA data. Estimates are calculated by drawing 10%, 20% etc. samples of player-year observations from those who started at least one game during the year. M/J gives the average number of player moves per team in the resulting dataset's largest connected set of players and teams.