Is there a nationality wage premium in European football?

Abstract

We investigate the presence of nationality salary premia in two top European football leagues (the Premier League and Serie A). We uncover a substantial pay premium for South American players (primarily driven by Argentina and Brazil) of between 11 and 15 per cent in magnitude. We investigate possible mechanisms, such as whether these salary effects are driven by new entrants to the league, and whether they are reflected in team attendances and team performance. Fans appear to respond to higher proportions of South American players in England, but not in Italy. We discuss the implications of these results and suggest why potential differences might exist across the leagues.

Word Count: 8226

Keywords

Football, Nationality, Salary

INTRODUCTION

There is a well-established literature examining the migrant wage premium in labour markets (Dustmann et al., 2016). The underpinning theoretical model suggests that workers will move if career earnings abroad exceed that of staying at home, having adjusted for the costs of migrating (Borjas, 2013). It might be, for example, that a particular skill is valued more highly in a foreign country, hence encouraging workers to migrate.

Empirical work on identifying such premia is challenging however, without access to highly sensitive data on the earnings and migration status of labour market participants. We turn to an industry where wages are public information, performance is well observed, and nationality is common knowledge. The industry is professional football. Faced with challenges of measurement of wages or performance, economists often turn to sports data, and with good reason (see Bar-Eli et al., 2020 for a discussion).

In sports labour markets, migration is commonplace, and leagues have increased their hiring of foreign-born talent in recent years. According to information on Transfermarkt.com, in the inaugural Premier League season (1992-93), the most represented nationality outside of the British Isles was Norway with eight players, followed by Australia with six. Shortly after the Bosman ruling, the EPL experienced an increase in the number of Northern European (especially Norwegian) players.[1] Ten years later in the 2002-03 season, France had 39 players and Norway had 13. Migration from non-EU countries has increased steadily since the 2001/02 season, particularly in the number of players from Western Africa. At present,

[1] The Bosman ruling of 1995 imposed by the European Court of Justice allowed players within the European Union to move freely to a new team at the end of their contract.

Brazil makes up the highest number of foreign players with 33, followed by France (30) and Spain (28).[2]

In this paper, we seek to examine a) whether a pay premium exists for players of different continental origins, and b) if there are any identifiable reasons behind these pay premia. Our setting is the most elite football leagues of England (the Premier League) and Italy (Serie A). We find strong evidence of a South American pay premium in both leagues. In England, this can largely be explained by South American players appearing to be more popular among fans as shown by higher attendances. This is not the case in Italy. We shed light on these findings and conjecture why estimates might differ between the 2 countries.

It is worth noting that we do not ask whether discrimination is because of race. We believe this is a different research question, and one that has been examined in recent literature. Racist behaviour by fan groups is unfortunately a common offence in football, and Principe and Van Ours (2022) show that there is a racial bias by journalists in Italy when assigning player ratings for Serie A matches – black players at the lower end of the salary distribution receive systematically lower ratings than their white counterparts. During their paper however, they also show that there is no effect of a black player dummy on log salary. We do not pursue replication of this finding here.

The remainder of the paper proceeds as follows. Section 2 presents our literature review while Section 3 covers data and methodology. The results are reported and discussed in Section 4. Section 5 concludes.

[2] This trend is not just limited to football. The NBA (National Basketball Association) has also increased its hiring of foreign-born players over the last 30 years, while more recently, the NFL (National Football League) now has an international player pathway aimed at increasing non-United States and non-Canadian representation in the league.

LITERATURE REVIEW

A considerable literature has developed attempting to explain wage premia / penalties for workers of different nationalities. This is true not only for sports labour markets, but also more broadly, for example the examination of the wages of migrant workers (e.g. Cengiz & Tekgüc, 2022).

Theoretically, according to Wilson and Ying (2003), sports teams would be willing to pay a premium for domestic talent under 3 scenarios, which align with Becker's (1971) three types of taste-based discrimination. The first is employer discrimination; that teams are willing to indulge their nationalistic preferences. The second is that players prefer playing with domestic talent (co-worker discrimination), while third is that consumers prefer watching domestic talent (customer discrimination). Wilson and Ying (2003) demonstrate little evidence in support of customer and co-worker discrimination, and as such, an under-representation of players of certain nationalities must originate from the preferences of team owners.

Directly testing such a claim is, however, difficult without data on individual player wages and performance. Indeed, early examinations of wage premia were hindered by these problems. Lucifora and Simmons (2003) was a rare exception, though their study was attempting to uncover superstar effects, rather than evidence of salary premia by nationality.

As such, the gold-standard for many years was a market-test type approach developed by Szymanski (2000). The approach involved regressing total club wage bills on team performance and proportions of black players on the team's roster. This work found evidence of discrimination of black players, in the sense that teams composed of undervalued black players would be capable of achieving a greater level of success with the same wage bill as a team composed of overvalued talent. More recent work by Deschamps and De Sousa (2021)

finds that this difference remained for only one identifiable group of players (black non-EU players) following the 1995 Bosman Ruling. This loosened mobility restrictions within the EU and curbed the monopsony power of clubs; non-EU black players who faced restrictions on immigration.

As data on player performance and individual wages improved, so too has the number of papers examining the determinants of sports salaries. Closely related to our work is that of Bryson et al. (2014). Their work finds that domestic players playing in Italy's Serie A experience a salary penalty compared to their migrant counterparts, particularly so at the upper portions of the salary distribution. Other closely related papers that also demonstrate salary premia by nationality include Frick (2006) who notes salary premia for Eastern European, Western European and South American players (of 15%, 30% and 50% respectively) over a comparable German player in the German Bundesliga. Garcia-del-Barrio and Pujol (2007) also reports a salary premium for non-Spanish European playing in Spain, while non-EU players receive a salary penalty.[3] More recently, Scarfe et al. (2021), using a similar modelling technique to our own, find that superstar players in the highest level of North American football (or soccer – Major League Soccer, MLS) are paid more because of the additional gate revenues they can produce, rather than via their effects on team performance. They do not however, consider the effect of a player's nationality. We ask if a similar reasoning exists for the nationality of players.

Pedace (2008), using the "market-test" approach described above, finds that South American players are typically overpaid in the English Premier League.[4] Finally, in the context of MLS, Medcalfe and Smith (2018) find that domestic born players earn less than

[3] Also see work by Bell et al. (2024) who find that English players tend to be valued higher and paid more.
[4] Pedace goes on to explain that this may actually be a perfectly reasonable decision by team owners, on the basis that average attendances are higher when teams give South American players more appearances.

equally productive foreign-born players. Talent from Brazil, Argentina and Colombia are higher earners (among other nationalities). Furthermore, they find evidence that foreign born players positively affect attendance.

In European football, trades of player registrations are for cash via the transfer market. Depken and Globan (2021) demonstrate a transfer fee premium for players transferred into the English Premier League relative to players transferred into other European leagues, with this premium rising over time. To the extent that fees and player salaries are correlated, this result suggests a salary premium for migrant players entering the English Premier League over the Italian Serie A. Our focus in this paper, though, is the source of any wage premium by continental origin.

DATA & METHODOLOGY

Our data consist of 1,587 players (or 3,944 player-season observations), over a seven-season period from 2013/14 to 2019/20 for the top divisions of England (the Premier League) and Italy (Serie A). These leagues are made up of 20 teams per season, and teams will typically have a squad of approximately 25 players which they can choose from. We omit goalkeepers from our analysis, as they have qualitatively different roles and performance metrics (Berri et al., 2023). We also apply a minute's filter as a requirement for a player to be included in the sample. This filter is set to 450 minutes in a given season (equivalent to 5 completed games) to exclude any reserve or backup players. Each player is then matched with a suite of performance statistics and experience variables from whoscored.com. Recognising the possible problem of reverse-causality between current performance and pay, we adopt the standard approach of lagging our performance covariates by one season.

Player salary data is sourced from www.capology.com. The salary data is a player's base salary, in that it contains no performance related bonuses, or no additional income from

sponsorships and endorsements. While the website does contain salary information from other major leagues, only the Premier League and Serie A have a high enough number of verified salaries to be confident in using them. A verified salary is one which has been confirmed either by the club and/or the player's agent. To check the accuracy of the salary data, we also cross-checked Serie A salaries with the salary data reported by the Italian sports magazine *Gazzetta dello Sport*. There was an exact match between the two sources. Table 1 shows the descriptive statistics for the (raw) salary figures, while the accompanying figures 1 and 2 present the Kernel salary density plots, split by league and by position.

To uncover any potential continental salary premium, we control for a set of confounding factors. Our control variables, listed in Table 2, comprise age (and its square), appearances for the national team (both at senior and Under 20 level), appearances in top five European leagues, appearances in other leagues, appearances in UEFA club competitions, and a dummy variable indicating whether a player was previously contracted to a club in a top five European league. This latter variable will help us distinguish between those who arrive directly to the Premier League or Serie A versus those who arrive from another elite European league. This may be somewhat indicative of a player's bargaining power. If they have experience in another major European league, a team might be willing to pay more for their services, while a player might have higher salary expectations. Bargaining power has been shown to be important in the context of footballer salaries (see Carrieri *et al.*, 2018). It is worth noting that all appearance measures are career totals prior to the current season of recorded salary. After estimating the salary models, we also estimate models for team attendances and team points per game.

/*TABLE 1 HERE (Descriptive Statistics for Salary)*/

/*FIGURES 1 and 2 HERE*/

/*TABLE 2 HERE (Descriptive Statistics (for players with minutes>450))*/

Average salaries are higher in the Premier League than in Serie A, most likely due to higher value of broadcast rights in the former (Depken and Globan, 2021). Salaries are more skewed in Serie A. The long right tail of the Italian league distribution is essentially due to the contract of Cristiano Ronaldo whilst at Juventus.[5]

Our methodology begins by specifying an OLS regression as follows,

$$\log(salary)_{ijt} = \alpha + X'_{ij(t-1)}\beta + \sum_N \gamma N_i + \phi_t + \eta_j + u_{ijt} \qquad (1)$$

For player *i*, playing for team *j*, in season *t*. The vector X includes a series of (lagged) performance and human capital controls. Rather than deflating salary by a consumer price index, we allow any salary inflation to be captured by the season fixed effects, ϕ_t (similar to Berri et al., 2022). Team fixed effects are denoted by η_j. Since different clubs will have different ability to pay, these fixed effects are necessary to include.[6] Our variable of interest is then N_i, a series of dummy variables identifying player i's continent of origin (detailed below). Due to institutional differences and possible differences in returns to specific player attributes between the two leagues, we estimate a separate model for each league.[7] To complete (1), u_{ijt} is a random error term.

Our nationality dummies are *Africa, Asia, North America, South America, Oceania*, along with a dummy variable *Home*. *Home* is a dummy variable equal to one if an English player is playing in the Premier League, and equal to one if an Italian player is playing in

[5] See robustness checks.
[6] Other approaches in the literature to proxy for ability to pay include using lagged team performance (Montanari et al., 2008) and lagged team attendance (Berri et al., 2023). In a perfectly competitive model, these team fixed effects would be unimportant. In each model which includes team fixed effects, we report the F-statistic justifying their inclusion.
[7] This approach is also justified econometrically. We ran a pooled leagues model which included covariates interacted with a league dummy, and an F-test on these interactions reveals they are jointly significant (F=7.94, p=0.000).

Serie A. Hence, the omitted category is non-home European players.[8] Further to these broad categories, we also disaggregate South America into Brazil, Argentina and Uruguay versus the rest of South America. These countries are traditionally the highest quality footballing countries in South America and so talent might plausibly be different (or at least perceived to be) to the rest of the continent. In fact, three of the greatest football players of all time – Diego Maradona, Pele, and Lionel Messi – were born in Argentina or Brazil. It is possible that decision-makers in Europe perceive talent from these two nations as being better than talent from other nations.

Of course, it is very difficult to claim all possible controls have been included in a model to replicate the decision of club offering a contract to a player. If these unobservable factors are indeed a driver of salary premium by continent of origin, our estimated nationality coefficients will be biased. We address this possibility in the robustness checks by employing the methodology of Cinelli and Hazlett (2020). While these checks should not be considered a silver bullet to make causal claims, they do provide a benchmark of how much confounding would be required to overturn a research conclusion.

As demonstrated by the distributional plots of salaries in figures 1 and 2, salary reveals excess skewness, as is true of athlete salaries across many sports leagues. It is worth noting of course, that none of the assumptions about the desirable properties of OLS (unbiasedness, consistency, BLUE etc.) rely on the normality of the outcome variable. Nevertheless, we might still be a) interested in effects of covariates across the distribution of salary, and b) concerned about the presence of large outliers. To address this concern, we considered two alternative estimators. We first employed Unconditional Quantile Regression (UQR), as developed by Firpo et al. (2009). This technique has been applied widely, including in the

[8] Since there was no player from Oceania playing in Italy in our sample period, this coefficient cannot be estimated for Serie A.

analysis of sports wages (Deutchser et al., 2017; Carrieri et al., 2018). However, our UQR results failed to reveal any systematic or consistent results across quantiles, including coefficients for control variables. Moreover, there were no clear patterns as to the significance of effects of continents across the salary distribution.

Our lack of confidence in the UQR results led us to apply a Huber robust regression estimator to deal with outliers at the top of the salary distribution.[9] The procedure performs an initial screening of the data based on Cook's distance above one. Once any gross outliers are removed, iterations are performed using Huber weights and biweights sequentially, until convergence is achieved.

Huber regression uses a loss function that combines the properties of Mean Squared Error (for OLS) and Mean Absolute Error. This loss function then reduces the contribution of outliers to the squared error loss function.The resulting standard errors are robust to heteroskedasticity. This method addresses criticisms of non-standard distributions of the dependent variable (see Elliott and Simmons, 2008, for an application to the UK film industry). We also carry out a robustness check where we simply remove the largest outlier – Cristiano Ronaldo - from the sample.

The implicit causal mechanisms underlying our work are twofold. First, basic theory suggests that in a competitive labour market with few restrictions on mobility (following the Bosman ruling of 1995) and fully observable player performance, pay should approximate the value of marginal product. Although we do not directly test that conjecture, our models suggest that observed player performance from the previous season is a predictor of current salary. That notion is consistent with pay equalling the value of marginal product. Our model is also consistent with a positive correlation between team payrolls, relative to league

[9] The results of the UQR models are available on request.

average, and team performances, as initially proposed by Szymanski and Smith (1997), and now well established within the sports economics literature.

The second mechanism is that fan demand, in the form of attendance, responds to team success and/or team ability (rather than vice versa). Some of the recent literature from the COVID-19 pandemic induced behind-closed-doors (or limited attendance) matches would offer some support to the claim. For example, Fischer and Haucup (2021) find no significant changes to home advantage in the 2nd and 3rd tiers of German football when games were played with no crowds present. Moreover, Wunderlich *et al*. (2021) present evidence across 10 major European football leagues, while Bryson *et al.* (2021) report estimates from 23 leagues, and neither find any significant changes to home advantage/final match scorelines. This would not support the view that larger crowds lead to improved performances.

RESULTS & DISCUSSION

Baseline Models

We begin by presenting an OLS model of log gross pay. Results for both leagues modelled separately are shown in Table 3.

/*TABLE 3 HERE (OLS salary models)*/

Results from the OLS regressions reveal that, as expected, players are rewarded for shots on target, goals, pass completion percentage and the number of completed passes per game. These simple performance metrics are significant predictors of player pay. Experience, in the form of appearances for a national team, UEFA club competitions and in top 5 leagues, is also rewarded in salary. Age has the typical positive but diminishing effect on salary, with the turning point occurring in the intuitively reasonable range of 28-30 years old (with a slightly older peak in Italy). Forwards are paid more than midfielders (omitted category), who are in turn paid more than defenders (significantly for Italy), consistent with Frick (2007).

Turning to the effects of the continent dummies, players originating from South America receive a significant salary premium in each league. Point estimates are 14.3% for England and 12.4% for Italy. Note that we cannot rule out that these coefficients are significantly different from on another (p=0.65). It is notable that when the South America continent dummy is interacted with previous experience in a big five league, this term delivers a smaller and less precisely estimated coefficient (p-value of 0.052 for England, and 0.065 for Italy), while the continent dummy itself has an insignificant coefficient. This suggests that the South American wage premium is being driven by players moving between the top five European leagues, rather than those who are direct entrants from South America.

South America aside, no other continent dummies reveal a significant effect at the conventional 5% level. Butler and Coates (2022) find a similar salary premium for South American players in MLS. However, their work differs from ours in that they also present positive salary effects for players from Africa and Europe. It is worth noting however, that MLS is a very different institutional setting to European football, with 'star' players being able to command significant premia with a 'designated player' status.

In contrast to Butler and Coates, we find the Africa continent dummy to have an insignificant coefficient, though it is negative (yet imprecisely estimated) for England only, when interacted with Previous Big 5. There is a suspicion here of a wage penalty for African players in the English Premier League (but not Serie A). One could speculate about various mechanisms driving this. We suspect African players are riskier prospects due to uncertainty over previous performances in lower calibre domestic leagues. Also, scouting and recruitment networks and better established in South America relative to Africa. African players or clubs may not have comparable access to elite representation (agents etc.), which may impact their bargaining power or negotiation of higher salaries. Finally, African players

may hold a stronger preference to move, accepting lower salaries if a transfer facilities greater exposure and a long-term pathway to better future contracts in Europe.

A further interesting result is the effect of the 'Home' dummy. This measures the effect of playing in the same country as their nationality. The effect is not significant for either Italy or England. This contrasts with Bryson et al. (2014) who uncover a negative effect on salary for Italian players playing in Italy. They explain this by arguing that Italian players prefer to stay in their domestic league, and hence clubs can offer a lower salary. It is possible that over time this effect or preference has disappeared since this work. Bryson et al. (2014) cover a period up to 2007, yet Adessa et al. (2022) show that since this period, there has been a dramatic shift in the demographic makeup of teams in Serie A. Since 2012, the fraction of non-Italian players has been greater than the fraction of Italian players in Serie A. It is possible that this shift has contributed to our results contrasting with Bryson et al. (2014). We also seek to offer other explanations in the following sections.

/*TABLE 4 HERE (Huber robust regressions)*/

Table 4 presents the models similar to those in Table 3 but using Huber robust regression instead of OLS. Among the control variables, coefficients that were significant at 5% or greater in OLS estimation remain so in Huber regression. Goals scored now enter significantly for England. Our focus – the South America continent dummy variable - also retains significance, with point estimates of 12.1% and 10.8% for England and Italy respectively, while columns (2) and (4) again suggest this premium is driven by players moving between major European leagues.

Robustness checks

Alternative Continent Definitions

We begin our robustness checks by testing if the effects of the South American salary premia are consistent across the entire continent, or if they are limited to particular nations which might be considered to have stronger footballing pedigree (Bellos, 2002). Table 5 introduces new dummy variables, indicating whether the player is from Brazil and Argentina (*Brazil & Argentina*), or Brazil Argentina and Uruguay (*Brazil, Argentina & Uruguay*). The remainder of the continent dummies and the variables which capture performance and experience and not shown for brevity, but they perform identically to those in Table 3.

/*TABLE 5 HERE (South American continent checks)*/

From Table 5, it appears the South American salary premium behaves quite differently across England and Italy. In Italy, the South American pay premium originates from Brazil, Argentina and Uruguay. Given that these countries are traditionally the strongest footballing nations of South America, it is plausible that players from these countries can command higher salaries. If historically, players from these countries are productive in Serie A (or perceived as being productive by decision makers), then this may act as a signal for current employers that current talent is also productive.[10] Of course, other South American countries do produce very successful talent, but there may be reasons why the talent is less successful at breaching the international marketplace, including a lack of investment in youth development facilities, political unrest, poverty etc. In England however, these nations hold no additional explanatory power over and above the South American [illegible]ct, Below, we

[10] Note that Berri et al. (2015) and Berri et al. (2022) demonstrate a preference for talent born in the USA in the Spanish and Chinese basketball leagues respectively. Like our findings for European football, this is further supporting evidence that decision makers favour players from nations where historically some of the best players are from.

examine whether these differences can be explained by their contributions to team attendances and/or team performances.

Further robustness checks

In this section we outline two robustness checks. The first is simple. This entails the removal of the superstar Portuguese forward Cristiano Ronaldo during his time at the Italian club Juventus from our sample. Ronaldo's salary is a clear outlier in the salary distributions. His gross salary was 57 million euros, while the next highest salary (in Serie A) was Gonzalo Higuaín, the Argentine forward who played for AC Milan and Juventus, earning 1.8 million euros. Results (available on request) show the removal of Cristiano Ronaldo from the analysis makes virtually no difference to results.

The second check concerns the robustness of our continent dummies in the wage equations. While we do include a wide set of performance and experience controls to account for heterogeneity between players, one still might be concerned about the role of unobserved confounders biasing these coefficients. In particular, there may be some unobserved variable (or variables) correlated both with a player's continent of origin and their wage. Since the South American continent dummy has been the source of most of our discussion, this is the one we pay attention to in the sensitivity testing. It is worth discussing the potential sources of this confounding – unobserved performance traits, bargaining power, player popularity, are all possibilities.

We carry out sensitivity analysis as described in Cinelli and Hazlett (2020). [11] We find a robustness value of around 9% (for both the Premier League and Serie A models). That is, if unobserved confounding accounts for 9% of the residual variance in *S.America* and *log(salary)*, the estimated effect of *S.America* would be reduced zero. Confounding could

[11] This is implemented in *Stata* using the command *sensemakr* (Cinelli *et al.*, 2020)

also bring the coefficient on *S.America* to a range where it is no longer statistically different from zero (at a 5% level). In such a case, the equivalent figure would be a little under 5%. Cinetti and Hazlett also suggest benchmarking these measures against an observed covariate, typically the "strongest" covariate (in our case *shots on target*). For a variable as strong as *shots on target*, confounding explains between 1.2%-1.4% (Premier League and Serie A respectively) of residual in *log(salary)* and less than 0.1% in the residual variation in *S.America*. Figures 3A and 3B show that even a confounding variable 20 times as strong as *shots on target* would not be sufficient to bring the estimated coefficient on *S.America* to zero.

Attendance and Points per game models

In line with marginal productivity theory, the salary premia by nationality are justified if such players contribute positively to team performance (marginal product) and/or team attendances (marginal revenue, in that teams can sell wins to fans), or both. We test these possibilities by regressing team points per game and (log) team attendances on the proportions of players from each continent on the roster of a team in a given season.

Following Szymanski and Smith (1997), team revenues, and team attendances underlying this, should be a function of team points. In our team attendance models we also include the predicted log salary estimated using models in Table 3 (but excluding the nationality dummies). For the attendance models we include points per game as an additional regressor. Table 6 shows the estimates from these team-level models for England, and Table 7 shows the models for Italy. All models include team and season fixed effects.

In England, fans appear to respond to higher proportions of South American players on their team, as demonstrated by the positive and significant effect on the South American variable. As with the salary estimates, the majority of this effect is driven by higher numbers

of Brazilian, Argentinian and Uruguayan players. In some sense, this justifies their higher salaries if teams are more able to sell wins to fans with higher proportions of South American players on their roster. The same cannot be said about their impact on team performance. In short, fans and teams think they are good, but performances would not support this view.

In Italy, while South American players received a salary premium, there is no evidence that this is driven by fan preferences or on-pitch performances. In Serie A, South American players appear to be overpaid. The insignificance of the remainder of the continent variables appears to square with the salary estimates from Table 3.

As for why the results between Italy and England differ, we offer the following thoughts according to the differences in institutional settings between the two leagues. According to UK immigration laws, for footballers to be granted a work visa, they must have been named in 75% of national team squads over the last two years, before being allowed to be transferred to a Premier League team. Therefore, these players are already likely to be the star players, commanding very big salaries, and attracting intense fan interest. Hence, they have strong outside options, and this bargaining power might improve their hand in salary negotiations. The same cannot be said of Italian immigration laws which mean it is much easier for South American players to migrate to Italy. Many South American players have dual nationality and so already have an EU passport when moving to Italy. To illustrate this point, of the current (or recent) Italian national team(s), Rafael Toloi, Emerson Palmieri and Jorginho were all born in Brazil, while Mateo Retegui was born in Argentina. Historically, players have moved to Italian football even at fairly low pay to aid their career progression, using lower status Serie A teams as stepping stones.

The inconsistency between impacts of proportions of South Americans on attendance and team points, and corresponding zero effects for Italy, remain puzzling and offer an

opportunity for further work. It could be that the English Premier League teams on average attract better, and more appealing, South American players than Italian clubs can. The popular stereotype of South American players is that they deliver skill, style, creativity and flair not exhibited by European players (Bellos, 2002). However, this supposed skill set does not appear to translate into improved team performance in our results in Table 6. We should caution that these estimates are based on small sample sizes (only 140 observations per league). Also, we cannot rule of productivity spillovers from South Americans onto domestic players. Identification of this channel would require more granular, match level data which is beyond the scope of this paper.

/*TABLE 6 HERE (Attendance and Points per game models, England)*/

/*TABLE 7 HERE (Attendance and Points per game models, Italy)*/

CONCLUSIONS

This paper has looked to quantify the existence of salary premia by continent across the English Premier League and the Italian Serie A. OLS estimates uncover a nationality pay premium for South American players in both leagues of between 11% and 15%, with higher premia in England. These premia are driven primarily by players from Brazil, Argentina, and Uruguay. Unlike past research, we do not observe a salary penalty for domestic players.

In attempting to explain these salary premia, we estimated models for team attendances and points per game. English teams have some justification for paying a premium to South American players, but African players appear to be underpaid in England given their contributions to team performances. In Italy, it is not possible to explain the pay premium to South American players either via performance or attendances. Further work might seek to examine this, or investigate this issue across other sports leagues.

It is interesting to offer further conjecture on why Brazilian, Argentinian and Uruguayan players are paid more than players from other South American countries. These thoughts could possibly offer new empirical questions going forward. Numerous (potentially interactive) mechanisms are plausible. First, buyers may deem these nationalities as a credible quality signal, given each nation's historic successes. There may also be positive feedback loops in scouting systems given past recruitment successes. Second, on the surface football appears to be cherished across the entire continent of South America but these nations could have superior football infrastructure and scouting systems in place. Player pay may be tied to the quality of the environment in which the player was produced. Third, it is possible that wider economic and political conditions across the continent have an impact. For example, child poverty and malnutrition could be a more pressing issue outside of these countries – these factors can impact the supply of talent. Different political systems (e.g. Chile) may also impact the possibility to play and develop within the sport.

Further work could proceed in two additional directions. It will be of interest to pin down the source of South American wage premia between players who arrive directly from South America versus those who migrate via various European football leagues, using these leagues as a stepping stone for eventual arrival into elite leagues. Finally, we have presented a static picture of the South American wage premium. Future work could usefully identify a) how this premium varies (if at all) over player careers, and b) how the premium might be affected over time by newly negotiated broadcast deals in line with the work of Depken and Globan (2021) on transfer fee premia. For now, our message to English and Italian team executives is a simple one. Be very careful when assessing South American players in recruitment strategies as you may end up paying more than you need to, based upon observed ability and performances.

References

Adessa, F., Pazzona, M. & Rossi, G. (2022). Migrant diversity and team performance in a high-skilled labour market. *Kyklos,* 75(3), 365-384

Bar-Eli, M., Krumer, A. & Morgulev, E. (2022). Ask not what economics can do for sports – Ask what sports can do for economics. *Journal of Behavioural and Experimental Economics*, 89, 101597

Becker, G. S. (1971). *The Economics of Discrimination*. University of Chicago press.

Bell, A., Brooks, C. & Brooks, R. (2024). Are English football players overvalued? *Applied Economics*, 56(21), 2568-2584

Bellos, A. (2002). *Fuebol: The Brazilian way of life*. Bloomsbury, London.

Berri, D., Burdekin, R. & Deutscher, C. (2022). Nationality Effects on the Allocation of Playing Time in the Chinese Basketball Association: Xenophilia or Xenophobia? *Journal of Sports Economics*, 23(2), 156-174

Berri, D., Butler, D., Rossi, G., Simmons, R. & Tordoff, C. (2023). Salary determination in professional football: empirical evidence from goalkeepers. *European Sports Management Quarterly,* 1-17

Berri, D., Deutscher, C. & Galletti, A. (2015). Born in the USA: National origin effects on the time allocation in US and Spanish professional basketball. *National Institute Economic Review*, 232(1), R41-R50

Berri, D., Farnell, A. & Simmons, R. (2022). The determinants of Black quarterback pay in the National Football League. *Managerial and Decision Economics*, 44(3), 1491-1503

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

Bryson, A., Dolton, P., Reade, J.J., Schreyer, D. & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters*, 198, 109664

Butler, D., & Coates, D. (2022). Position Premium in Major League Soccer. *International Journal of Sport Finance,* 17(4), 201-214

Borjas, G.J., 2013. Labor Economics, *6th edition,* McGraw-Hill

Carrieri, V., Principe, F. & Raitano, M. (2018). What makes you 'super-rich'? New evidence from an analysis of football players' wages. *Oxford Economic Papers*, 70(4), 950-973

Cengiz, D. & Tekgüc, H. (2022). Is it merely a labor supply shock? Impacts of Syrian migrants on local economies in Turkey. *ILR Review*, 75(3), 741-768

Cinelli, C. & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1), 39-67

Cinelli, C., Ferwerda, J. & Hazlett, C. (2020). sensemakr: Sensitivity analysis tools for OLS in R and Stata. Available at SSRN 3588978

Depken, C.A. & Globan, T. (2021). Football transfer fee premiums and Europe's big five. *Southern Economic Journal*, 87(3), 889-908

Deschamps, P. & De Sousa, J. (2021). Labor mobility and racial discrimination. *European Economic Review*, 135, 103738

Deutscher, C., Gürtler, O., Prinz, J. & Weimar, D. (2017). The payoff to consistency in performance. *Economic Inquiry*, 55(2), 1091-1103

Dustmann, C., Schönberg, U. & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4), 31-56

Elliott, C. & Simmons, R. (2008). Determinants of UK box office success: The impact of quality signals. *Review of Industrial Organization*, 33(2), 93-111

Firpo, S., Fortin, N.M. & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica,* 77(3), 953-973

Fischer, K. & Haucup, J. (2021). Does crowd support drive the home advantage in professional football? Evdence from German ghost games during the COVID-19 pandemic. *Journal of Sports Economics,* 22(8), 982-1008

Frick, B. (2006). Salary determination and the pay-performance relationship in professional soccer: Evidence from Germany. *Sports Economics After Fifty Years: Essays in Honour of Simon Rottenberg. Oviedo: Ediciones de la Universidad de Oviedo*, 125-146

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

Garcia-del-Barrio, P. & Pujol, F. (2007). Hidden monopsony rents in the winner-take-all markets – sport and economic contribution of Spanish soccer players. *Managerial and Decision Economics*, 28(1), 57-70

Lazear, E.P. (1981). Agency, earnings profiles, productivity, and hours restrictions. *The American Economic Review*, 71(4), 606-620

Lucifora, C. & Simmons, R. (2003). Superstar effects in sport: Evidence from Italian soccer. *Journal of Sports Economics*, 4(1), 35-55

Medcalfe, S., & Smith, R. (2018). The effect of foreign players on pay and performance in Major League Soccer. *International Journal of Sport Finance*, 13(4), 297-318.

Montanari, F., Silvestri, G. & Bof, F. (2008). Performance and individual characteristics as predictors of pay levels: The case of the Italian 'Serie A'. *European Sport Management Quarterly*, 8(1), 27-44

Pedace, R. (2008). Earnings, performance and nationality discrimination in a highly competitive labor market as an analysis of the English professional soccer league. *Journal of Sports Economics*, 9(2), 115-140

Principe, F. & Van Ours, J. (2022). Racial bias in newspaper ratings of professional football players. *European Economic Review*, 141, 103980

Scarfe, R., Singleton, C. & Telemo, P. (2021). Extreme wages, performance, and superstars in a market for footballers. *Industrial Relations: A journal of Economy and Society*, 60(1), 84-118

Szymanski, S. (2000). A market test for discrimination in English professional soccer leagues. *Journal of Political Economy*, 108(3), 590-603

Szymanski, S. & Smith, R. (1997). The English football industry: Profit, performance and industrial structure. *International Review of Applied Economics*. 11(3), 135-153

Wilson, D.P. & Ying, Y.H. (2003). Nationality preferences for labour in the international football industry. *Applied Economics*, 35(14), 1551-1559

Wunderlich, F., Weigelt, M., Rein R. & Memmert, D. (2021). How does spectator presence affect football? Home advantage remains in European top-clasd football matches played without spectators during the COVID-19 pandemic. *Plos one*, 16(3), e0248590

TABLES and FIGURES

Table 1: Descriptive Statistics for Salary

	Combined (€)	Premier League (€)	Serie A (€)
Mean	2,910,858	3,779,859	2,048,878
Standard Deviation	2,872,686	2,799,130	2,679,085
10th percentile	641,000	1,344,304	494,000
25th percentile	1,100,000	1,966,934	725,000
Median	2,128,490	2,906,800	1,253,000
75th percentile	3,704,000	4,675,840	2,545,000
90th percentile	6,081,401	7,224,750	4,630,000
95th percentile	8,091,720	9,816,301	6,203,000
Skewness	5.058	2.087	9.978
Kurtosis	72.314	9.122	188.183
N	3,944	1,964	2,304

*Note that formal tests of normality strongly reject the null hypothesis (p=0.000) of a normal distribution. This is true for salary distribution of both leagues considered either jointly or separately.

Table 2: Descriptive Statistics (for players with minutes>450)

Variable	Mean	Std. Dev.	Min	Max
(n=3,944)				
Dependent variables				
Gross salary (€m)	2.91	2.87	0.028	57.4
Log of salary	14.400	0.557	12.887	15.560
Log attendance*	10.192	0.538	8.341	11.230
Points per game*	1.380	0.481	0.42	2.68
Performance and Experience				
Shots on target	0.349	0.345	0	2.9
Minutes played	1918.906	796.656	451	3420
Goals	3.128	4.368	0	38
Assists	2.098	2.497	0	20
Pass completion %	79.983	6.854	49.4	95.8
Passes	34.259	14.657	4.4	102.6
Age	30.263	4.462	19	46
National team apps	18.704	26.435	0	158
U20 apps	4.347	6.255	0	40
Other apps	94.616	70.914	0	526
UEFA competition apps	19.647	24.109	0	170
Top 5 league apps	137.985	105.401	0	660
Previous Big 5	0.834			
Continent proportions				
Africa	0.089			
Asia	0.009			
Europe	0.753			
N. America	0.011			
Oceania	0.003			
S. America	0.135			
Home	0.380			
Position proportions				
DEF	0.397			
MID	0.334			
FWD	0.269			
League proportions				
Premier League	0.498			
Serie A	0.502			

*Note: log attendance and points per game are seasonal team averages, thus the sample size is n=280
Minutes played refers to the number of minutes played across a season. Goals and Assists are measured at the season level, while shots on target, and passing statistics are averages per game for a given season.

Table 3: OLS salary models

VARIABLES	EPL		SERIE A	
	Log(salary)			
Shots on target	0.302***	0.302***	0.337***	0.333***
	(0.062)	(0.062)	(0.066)	(0.065)
Minutes played	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Goals	0.006	0.005	0.014***	0.014***
	(0.004)	(0.005)	(0.005)	(0.005)
Assists	0.011**	0.010**	0.000	0.000
	(0.004)	(0.004)	(0.005)	(0.005)
Pass completion %	0.006***	0.006***	0.005**	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)
Passes per game	0.002**	0.002**	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.343***	0.343***	0.237***	0.234***
	(0.026)	(0.026)	(0.024)	(0.024)
Age squared	-0.006***	-0.006***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
National team apps	0.000	0.000	0.002***	0.002***
	(0.000)	(0.000)	(0.001)	(0.001)
Uefa competition apps	0.004***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Top 5 league apps	0.001***	0.001***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Defender	-0.004	-0.011	-0.061**	-0.064***
	(0.025)	(0.025)	(0.025)	(0.025)
Forward	0.093***	0.090***	0.161***	0.162***
	(0.034)	(0.034)	(0.035)	(0.035)
Africa	-0.005	-0.103*	0.006	-0.124
	(0.033)	(0.060)	(0.041)	(0.090)
Africa * PrevBig5		0.130**		0.154
		(0.066)		(0.097)
Asia	-0.002	-0.045	0.058	0.074
	(0.090)	(0.128)	(0.122)	(0.194)
Asia * PrevBig5		0.074		-0.028
		(0.173)		(0.246)
N. America	-0.014	-0.041	0.001	0.414
	(0.074)	(0.103)	(0.139)	(0.313)
N. America * PrevBig5		0.056		-0.508
		(0.145)		(0.347)
S. America	0.143***	0.058	0.124***	0.042
	(0.034)	(0.055)	(0.030)	(0.053)
S. America * PrevBig5		0.128*		0.104*
		(0.066)		(0.056)
Oceania	0.112	0.058		
	(0.116)	(0.126)		
Oceania * PrevBig5		0.321		
		(0.319)		
Home (England)	0.000	0.002		
	(0.023)	(0.023)		
Home (Italy)			0.019	0.021
			(0.027)	(0.027)

Constant	8.678***	8.682***	10.036***	10.086***
	(0.429)	(0.429)	(0.411)	(0.411)
Team & Season FE	Yes	Yes	Yes	Yes
F stat (p-val) Team FE	18.31 (0.00)	18.48 (0.00)	45.86 (0.00)	45.84 (0.00)
Observations	1,964	1,964	1,980	1,980
R-squared	0.664	0.666	0.767	0.768

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Huber robust regressions

VARIABLES	EPL		SERIE A	
	Log(salary)			
Shots on target	0.276***	0.279***	0.257***	0.255***
	(0.053)	(0.053)	(0.060)	(0.060)
Minutes played	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Goals	0.008**	0.007*	0.019***	0.019***
	(0.004)	(0.004)	(0.005)	(0.005)
Assists	0.008**	0.008**	0.002	0.003
	(0.004)	(0.004)	(0.005)	(0.005)
Pass completion %	0.006***	0.006***	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Passes per game	0.002**	0.002***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.226***	0.226***	0.206***	0.204***
	(0.023)	(0.023)	(0.022)	(0.022)
Age squared	-0.004***	-0.004***	-0.004***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
National team apps	0.001*	0.001*	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Uefa competition apps	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Top 5 league apps	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Defender	-0.025	-0.034	-0.045**	-0.046**
	(0.021)	(0.021)	(0.023)	(0.023)
Forward	0.080***	0.077***	0.160***	0.159***
	(0.029)	(0.029)	(0.032)	(0.032)
Africa	0.001	-0.181***	0.004	-0.121
	(0.029)	(0.052)	(0.037)	(0.083)
Africa * PrevBig5		0.234***		0.142
		(0.057)		(0.089)
Asia	-0.037	-0.070	0.008	-0.023
	(0.077)	(0.110)	(0.112)	(0.178)
Asia * PrevBig5		0.063		0.050
		(0.149)		(0.225)
N. America	-0.023	-0.067	0.048	0.358
	(0.064)	(0.088)	(0.127)	(0.287)
N. America * PrevBig5		0.121		-0.404
		(0.124)		(0.319)
S. America	0.121***	0.030	0.108***	0.024
	(0.029)	(0.047)	(0.028)	(0.049)
S. America * PrevBig5		0.134**		0.107**
		(0.056)		(0.052)
Oceania	0.188*	0.160		
	(0.100)	(0.108)		
Oceania * PrevBig5		0.108		
		(0.273)		
Home (England)	0.018	0.020		
	(0.020)	(0.020)		
Home (Italy)			0.025	0.027
			(0.025)	(0.025)

Constant	10.580***	10.575***	10.792***	10.834***
	(0.370)	(0.368)	(0.376)	(0.377)
Team & Season FE	Yes	Yes	Yes	Yes
F stat (p-val) Team FE	26.46 (0.00)	26.94 (0.00)	60.23 (0.00)	60.02 (0.00)
Observations	1,964	1,964	1,980	1,980
R-squared	0.713	0.716	0.795	0.796

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: South American continent checks

	EPL		SERIE A	
VARIABLES	Log(salary)			
Africa	-0.005	-0.006	0.003	0.005
	(0.033)	(0.033)	(0.041)	(0.041)
Asia	-0.001	-0.004	0.054	0.053
	(0.090)	(0.090)	(0.122)	(0.122)
N. America	-0.013	-0.014	-0.001	-0.004
	(0.074)	(0.074)	(0.139)	(0.139)
S. America	0.154**	0.105	0.054	0.035
	(0.060)	(0.068)	(0.045)	(0.053)
Brazil & Argentina	-0.015		0.108**	
	(0.070)		(0.051)	
Brazil, Argentina & Uruguay		0.049		0.115**
		(0.076)		(0.057)
Oceania	0.112	0.110		
	(0.116)	(0.116)		
Home (England)	0.000	0.000		
	(0.023)	(0.023)		
Home (Italy)			0.019	0.019
			(0.027)	(0.027)
Constant	8.675***	8.680***	10.047***	10.024***
	(0.430)	(0.429)	(0.410)	(0.410)
Season & Team FE	Yes	Yes	Yes	Yes
F stat (p-val) Team FE	18.15 (0.000)	18.09 (0.00)	45.90 (0.00)	45.94 (0.00)
Controls	Yes	Yes	Yes	Yes
Observations	1,964	1,964	1,980	1,980
R-squared	0.664	0.664	0.767	0.767

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Attendance and Points per game models (England)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Log(attendance)			Points per game		
Predicted wage	0.159**	0.164**	0.143*	0.089	0.088	0.023
	(0.080)	(0.080)	(0.078)	(0.212)	(0.212)	(0.210)
Points per game	0.026	0.031	0.034			
	(0.038)	(0.038)	(0.038)			
Home (England)	-0.143	-0.158	-0.204**	0.165	0.143	0.133
	(0.108)	(0.106)	(0.102)	(0.286)	(0.282)	(0.273)
Prop. Africa	0.022	0.002		0.639	0.580	
	(0.161)	(0.158)		(0.423)	(0.416)	
Prop. Asia	-0.195	0.124		0.702	0.553	
	(0.635)	(0.647)		(1.686)	(1.722)	
Prop. S.America	0.371**			0.041		
	(0.151)			(0.400)		
Prop. Braz, Arg, Uru		0.498**			-0.204	
		(0.190)			(0.506)	
Prop. N.America	0.314	0.381		0.386	0.390	
	(0.345)	(0.343)		(0.916)	(0.913)	
Prop. Oceania	0.162	-0.043		1.369	1.238	
	(0.721)	(0.703)		(1.910)	(1.869)	
Prop. Africa with PrevBig5			-0.214			0.275
			(0.158)			(0.423)
Prop. Asia with PrevBig5			-0.430			1.267
			(0.738)			(1.979)
Prop. S.America with PrevBig5			0.315*			-0.230
			(0.170)			(0.457)
Prop. N.America with PrevBig5			0.450			-0.353
			(0.496)			(1.333)
Prop. Oceania with PrevBig5			0.591			1.498
			(1.128)			(3.027)
Constant	8.524***	8.446***	8.782***	0.504	0.542	1.544
	(1.200)	(1.196)	(1.171)	(3.186)	(3.185)	(3.145)
Season & Team FE	Yes	Yes	Yes	Yes	Yes	Yes
F stat (p-val) Team FE	50.37 (0.00)	50.07 (0.00)	45.65 (0.00)	3.73 (0.00)	3.67 (0.00)	3.64 (0.00)
Observations	140	140	140	140	140	140
R-squared	0.967	0.968	0.968	0.807	0.807	0.803

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Attendance and Points per game models (Italy)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Log(attendance)			Points per game		
Predicted wage	-0.012	-0.004	-0.008	0.152	0.165	0.152
	(0.116)	(0.115)	(0.112)	(0.179)	(0.178)	(0.177)
Points per game	0.148**	0.149**	0.152**			
	(0.066)	(0.065)	(0.064)			
Home (Italy)	0.151	0.190	0.189	-0.133	-0.088	-0.043
	(0.160)	(0.164)	(0.138)	(0.248)	(0.254)	(0.218)
Prop. Africa	-0.185	-0.166		0.258	0.291	
	(0.274)	(0.272)		(0.424)	(0.421)	
Prop. Asia	0.241	0.318		-1.080	-0.957	
	(1.055)	(1.049)		(1.630)	(1.624)	
Prop. S.America	-0.045			-0.228		
	(0.228)			(0.352)		
Prop. Braz, Arg, Uru		0.066			-0.114	
		(0.264)			(0.409)	
Prop. N.America	-1.076	-1.056		-1.419	-1.393	
	(1.201)	(1.201)		(1.854)	(1.857)	
Prop. Africa with PrevBig5			-0.340			0.233
			(0.298)			(0.471)
Prop. Asia with PrevBig5			-1.877			0.004
			(1.787)			(2.824)
Prop. S.America with PrevBig5			0.158			-0.051
			(0.239)			(0.378)
Prop. N.America with PrevBig5			-1.480			-0.073
			(1.253)			(1.981)
Constant	10.475***	10.320***	10.437***	-0.355	-0.591	-0.467
	(1.679)	(1.668)	(1.610)	(2.598)	(2.586)	(2.545)
Season & Team FE	Yes	Yes	Yes	Yes	Yes	Yes
F stat (p-val) Team FE	12.65 (0.00)	12.97 (0.00)	12.93 (0.00)	3.14 (0.00)	3.11 (0.00)	3.03 (0.00)
Observations	140	140	140	140	140	140
R-squared	0.932	0.932	0.933	0.811	0.810	0.808

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1: Salary Distribution by league

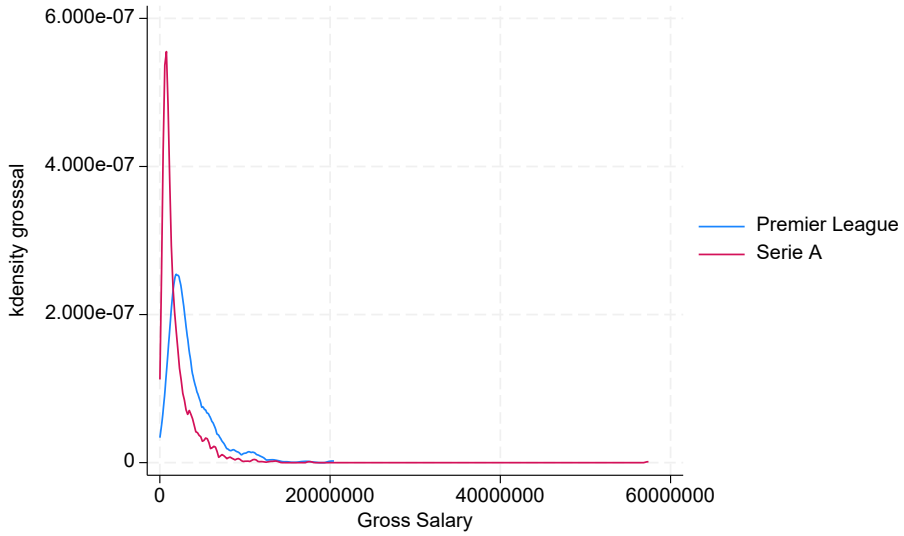
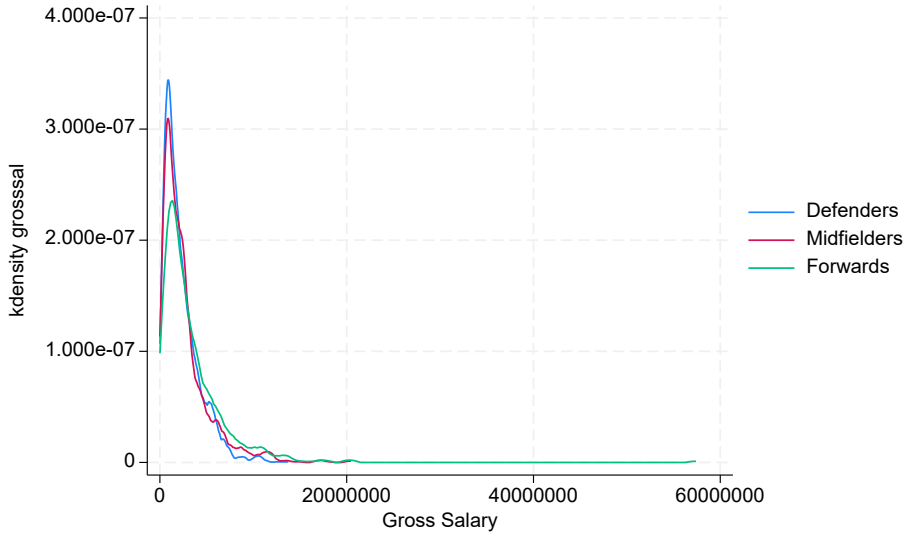
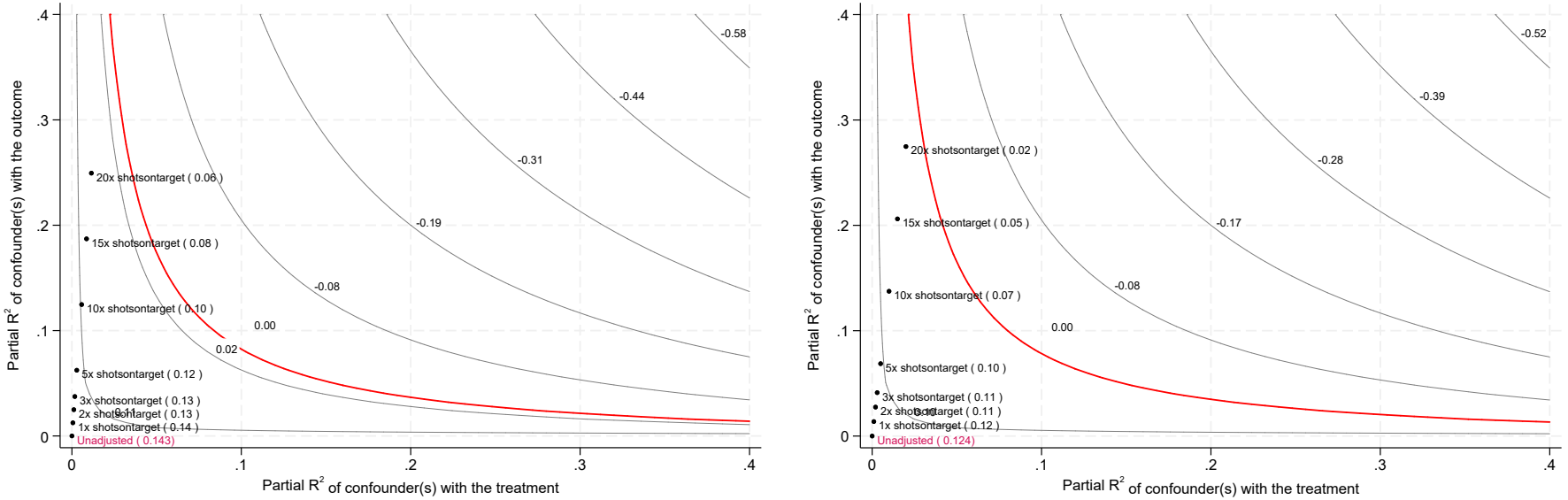


Figure 2: Salary Distribution by position



Figures 3A and 3B: Sensitivity analysis

3A: Premier League **3B: Serie A**



Note: Sensitivity of point estimates of *S.America* coefficient based on methodology of Cinelli and Hazlett (2020). The vertical axis shows the hypothetical partial R-squared of unobservable confounding with the outcome, and the horizontal axis shows the hypothetical partial R-squared of the treatment (*S.America*) with unobservable confounding. The contours show the point estimates that would have been obtained for *S.America* from a full model including the unobserved confounders under different hypothetical strengths. The plotted points show that the effect is robust to confounding that is one, two, three, five, ten, and up to twenty times as strong as the observed effect of *shots on target*. Figure 3A is based on the model in Table 3, column 1, and Figure 3B is based on the model in Table 3, column 3.