THE AGE OF DATA:
DOES AGE INFLUENCE PRODUCTIVITY OF TRANSFER PLAYERS IN THE ENGLISH PREMIER LEAGUE?

A THESIS

Presented to

The Faculty of the Department of Economics and Business

The Colorado College

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Arts

By

Juan Mota

December 2024

The Age of Data:
Does Age Influence Productivity of Transfer Players in the English Premier League?

Juan Mota

December 2024

Mathematical Economics

Abstract

Each year football clubs are tasked with finding players on the market to improve their team. Teams are trying to maximize competitive performance while constrained by their respective budgets. Established football clubs use player statistics developed this century to better predict future performances of players. This study builds on existing literature by examining performances of attacking players following a transfer to the English Premier League (EPL). The main variable in the model to explain productivity in the year after a transfer to the EPL is non-penalty expected goals plus expected assisted goals per 90. Other variables including player value as measured by transfer fee or market value, and where a player was transferred from are included in the model. Results from Random Effects and annual Fixed Effects (FE) OLS and 2SLS models find that age is not statistically significant when determining player performance. However, both the Transfer Fee and Market value which are modelled as a function of the player's most recent performance before the transfer are significant in explaining productivity after the transfer to the EPL.

ON MY HONOR, I HAVE NEITHER GIVEN NOR RECEIVED
UNAUTHORIZED AID ON THIS THESIS

Signature

I. Introduction

The transfer market is an opportunity for football teams to strengthen their team if they believe their team needs to improve. Not every team views the transfer market the same due to economic differences across teams. The promotion and relegation system used by football leagues across Europe aims to promote a competitive league system where teams must be ambitious to avoid economic downfall. In reality, European football leagues are closed league systems where established teams stay at the top and the rest of the teams must face uncertainty regarding their future, according to Dherbecourt and Drut (2009). It is challenging for teams with a limited economic budget to compete for major trophies because they cannot sign the best and most reputable players.

There has been a rise in analytics in the football industry over the past two decades. New statistics such as expected goals were created to estimate player output that traditional statistics do not measure. According to Mead et al. (2023), expected goals are a superior predictor of future performances compared to traditional metrics. Football clubs looking to gain an advantage in the market have been using statistics like expected goals to sign players that are undervalued.

This study contributes to the literature by analyzing player performances following a transfer to the English Premier League (EPL) in between the years 2017 and 2023. Only attacking players are considered due to the availability of statistics measuring their productivity. Loan transfers are not included in this study. For each transfer, a player’s attacking output was recorded along with other variables describing details of the player and past performances. Specifically, a player’s age, player value as measured by market value or transfer fee, previous production, and where they transferred from were included in the model. A player’s attacking output was measured using two variables, non-penalty expected goals plus expected assisted goals per 90 and expected goals per 90. Due to the results from Mead et al. (2023), expected goals and assisted goals were used as the statistic to measure attacking output of a player. Fixed Effects Ordinary least squares (FEOLS) and Fixed Effects two-stage least squares (FE2SLS) regressions were performed for both dependent variables. Findings from the models do not find age to be a statistically significant variable for predicting either non-penalty expected goals plus expected assisted goals per 90 or expected goals per 90. This goes against the intuition that older players, up to a certain age, are the best performing players. Transfer values and market values for players are both positive and statistically significant explanatory variables for player output. The rest of the results and summary of the data will be discussed in the appropriate chapters.

II. Literature Review

The current literature has focused on the economic and competitive value athletes throughout different sports bring to their teams. In each sport there are different ways to measure the productivity of an athlete and whether their performances justify their salaries. Koop (2002) uses regression techniques to measure the offensive performances of baseball players in Major League Baseball (MLB). Using data from 1995 to 1999 and statistical methods explained in Koop (2002), the most efficient players on offense are ranked. Hadley et al. (2000) is the first paper to examine the performance of head coaches in the National Football League (NFL). A Poisson regression is used to find the most efficiently managed teams. One of the results was that an efficient head coach can add an additional three to four victories for their team. Higgins (2022) analyzes the value of a quarterback in the NFL. Based on their findings, NFL teams should evaluate quarterback performances to determine those that are underpaid and overpaid to get a better understanding of the market.

There has been an analytical movement in the European football industry. The introduction of statistics such as expected goals (xG) change the way football executives and fans view the game. Mead et al. (2023) find that expected goals are a superior predictor of future performances compared to traditional metrics. Club analysts and betting companies use these models to make decisions that will help them outperform expectations. Rathke (2017) examines the factors that most accurately predict expected goals. The distance from the shot taken to the goal and the angle of the shot in relation to the net are accurate predictors of the expected goals statistic. Understanding how expected goals work can help clubs predict a player's ability to score goals in the future. Mead et al. (2023) aims to improve expected goal models by determining more factors that influence the probability of a shot resulting in a goal. The final model includes two variables that were not included in past literature and proved to be significant when calculating goal probabilities. One of the variables is player ability, which is determined by how much a player is worth. The other significant variable is psychological effects during a match which is measured by the goal differential when the shot was taken. Adding these variables to a traditional expected goals model can improve the accuracy of the model and will be a better predictor of future performances.

There is literature on the economics of European football clubs and how that affects player recruitment. Expenditures associated with the players such as wages and transfer fees are a football club's main source of costs, according to Dobson and Goddard (2001). Economic inequality between football clubs has been growing rapidly over the past half century due to systemic changes that have allowed for clubs to spend more money. According to FIFA (2024), football clubs spent a record of USD 9.63 billion on international transfer fees in 2023 which surpassed the previous record of USD 7.35 billion set in 2019. It was also the first time more than 1,000 clubs spent money on international transfer fees. With so much money on the table, there has been more research and analysis on the transfer value of players. Poli et al. (2021) analyses over 2000 player transfers from clubs in the top five European leagues starting from 2012 and ending in 2021 to determine how the market decides the prices of players. Multiple linear regression is used and the fees paid by clubs for players are an independent variable. A key result from Poli et al. (2021) is that the remaining duration on player contracts plays a key role in determining how much a player is worth in the market. Clubs can use this methodology throughout many applications in the football industry including contract negotiations, financial planning and communication, transfer negotiations, etc.

Football clubs with large revenues can afford to spend more on players. English football clubs with a higher player-related expenditure will benefit from superior team performances on the field compared to teams that spend less on players, according to Hall et al. (2002). A team with more resources than the rest can "buy" themselves a trophy or championship without anybody stopping them. In 2011, the Union of European Football Associations (UEFA) implemented Financial Fair Play (FFP) regulations to ensure football clubs have stability and are not in danger of economic collapse. The regulations require clubs to balance their spending to be aligned with their revenues and to stop clubs from accumulating debt (UEFA, 2011). With strict regulations in place, clubs must be careful when spending a lot of money on players. There are now financial and sporting consequences in place when clubs continually overpay for players and as a result spend more than their revenue.

Promotion and relegation football leagues add a layer of complexity when determining player values and transfer strategies. The promotion and relegation system is used across football leagues around the world. The purpose is to encourage competitive growth because any team can do well enough to play in the top division. In theory, an open league system as a result of promotion and relegation rewards ambitious teams. Noll (2002) analyzes data from the English football league to determine the economics of promotion and relegation. An "unbalancing effect" is observed which can be explained by newly promoted teams not wanting to invest in their team due to uncertainty in their future in the league. Dherbecourt and Drut (2009) observe this phenomenon and conclude that European football leagues are remarkably similar to closed league systems rather than open league systems. It is challenging for a promoted team to challenge the wealthy, established teams because they do not have the financial capabilities of the teams at the top. This affects the business and sporting decisions football clubs make because their future is uncertain. Pantuso and Hvattum (2020) look to maximize team performances while limited on a budget using a chance-constraint model. The model uses data from the English Premier League, and it can be used to help make decisions in the transfer market. Established and wealthy teams can spend more on players even if it is not the smartest financial decision which is different for promoted teams because they lack the budget to heavily invest on one player. I will be examining player transfers to the Premier League from 2017 to 2023.

III. Theory

Football clubs aim to maximize wins while being constrained by their respective budgets. Clubs can gain an advantage over their competitors by predicting the future performances of their players and analyzing transfer risk better than their rivals. To determine the performance of a player in their first season in the English Premier League, statistics measuring player output will serve as the dependent variables for two models. The two measures of player productivity variables will be non-penalty expected goals plus expected assisted goals per 90

(*NPXG_XAG_PER90*) and expected goals per 90 (*XG*90). A model of a player's non-penalty expected goals plus expected assisted goals per 90 is specified in Equation 1.

NPXG_XAG_PER90 = f(AGE, AGE_SQUARED, REALFEE_PERMIL, REALVALUE_PERMIL, LASTYEARGOALS, LASTYEARASSISTS, EPL_DUMMY, ENGLAND_DUMMY, EUROPE_DUMMY, YEARS) **(1)**

NPXG_XAG_PER90 represents the non-penalty expected goals plus expected assisted goals per ninety of a player. A player's non-penalty expected goals is the number of goals a player should have scored on average given the shots they have taken and excluding penalties. Each shot taken by a player is given an expected goal value based on variables such as assist type, shot angle and distance from goal, whether it was a headed shot and whether it was defined as a big chance, according to Opta. A player's assisted expected goals value is the expected goal value from a shot that was assisted from a pass. These two values are summed and adjusted per 90 minutes to standardize the minutes played for each player. *AGE* refers to the age of the player at the time of the transfer. *AGE_SQUARED* is the squared value of a player's age at the time of the transfer. *REALFEE_PERMIL* is the amount of Euros (€) paid by a football club to acquire the player, adjusted for inflation of the Euro. The value is scaled to per million to help with the interpretation of the regression. *REALVALUE_PERMIL* represents the market value in Euros of the player at the time of the transfer, adjusted for inflation of the Euro and scaled to a per million bases. The Harmonised Index of Consumer Prices (HICP) of the Euro area consisting of 19 countries from 2015 to 2022 is the price index used to convert nominal prices to real prices. *LASTYEARGOALS* is the number of goals scored by a player in the season before their transfer.

LASTYEARASSISTS represents the number of assists scored by the player in the season before their transfer. To account for players who played for more than one club in the season before their transfer, the statistics at the club they played for the most will be considered. *EPL_DUMMY* is a binary dummy variable with a one representing a player that played in the Premier League in the season before their transfer and zero otherwise. *ENGLAND_DUMMY* is a dummy variable with a one representing a player that played in any league in England excluding the Premier League and zero otherwise. *EUROPE_DUMMY* is a dummy variable with a one representing a player that played in any domestic league in Europe excluding England and zero otherwise. *YEARS* is a series of year specific dummy variables from 2018 to 2023 where a one represents the year in which the transfer took place and 0 otherwise. These dummy variables measure the impact of each year with respect to the year 2017.

Equation 2 specifies a model of a player's expected goals per 90. A player's expected goals per 90 is the number of goals a player should have scored on average. Penalty shot attempts are included and each value is standardized on a per 90 basis. The same explanatory variables will be used.

XG90 = f(AGE, AGE_SQUARED, REALFEE_PERMIL, REALVALUE_PERMIL, LASTYEARGOALS, LASTYEARASSISTS, EPL_DUMMY, ENGLAND_DUMMY, EUROPE_DUMMY, YEARS) **(2)**

The question to address is whether age plays a factor in a player's output following their first season after a transfer. A club that spends a large amount of money on a player expects them to perform well and to provide significant output on the field. A player that does well in a season could generate interest from other clubs which is why past player production was included.

IV. Data and Methods

The data set includes 284 transfers classified as attacking players transferring to the English Premier League (EPL) between the years 2017-2023. A total of 942 transfers were conducted during this period, including all player positions. A player's position, market, value, transfer fee, and age were taken from https://www.transfermarkt.us/. To meet the criteria of an attacking player, the player must be listed in one of four positions: left wing (LW), attacking midfielder (AM), right wing (RW), or center forward (CF). To determine the market value of a player, Transfermarkt uses various pricing models. Some of the variables Transfermarkt use in their models are Age, injury susceptibility, marketing value, performance potential, etc. For a list containing all the variables used, see Appendix A.

The non-penalty expected goals plus expected assisted goals per 90 and expected goals per 90 statistics for each player were scraped from https://fbref.com/en/. Each player's goal and assist tallies from the season before their transfer were provided by FBref. There were 333 transfers that met the criteria for an attacking player and the year of their transfer, but 49 have been removed due to missing data. Examples of missing data include missing market values, unknown transfer fees, and lack of statistics. In this dataset, the lack of statistics for some players is due to those players not getting any game time in the Premier League. When determining how many goals and assists a player got the season before their transfer, there were often players that played for more than one club. To account for this, the statistics for a player's time at a club in which they played for the most during that year are considered. Combining goal and assist tallies from different teams and leagues will influence the model by adding another variable that is not accounted for initially. Table 1 highlights descriptive statistics for variables included in the model.

The average non-penalty expected goals plus expected assisted goals per 90 is about .40 while the mean expected goals per 90 is about .28. The average age of the attacking players transferring in is 24. The average market value adjusted for inflation is about 17 million Euros while the average transfer fee paid is over 19 million Euros. Most players in this dataset were playing in leagues outside of England, but still situated in Europe. The most transfers happened in 2022 which was over double compared to the lowest amount which happened in 2020.

V. Results

I use ordinary least squares (OLS) and two-stage least squares (2SLS) with Random Effects (RE) and annual Fixed Effects (FE) to regress the dependent variables, non-penalty expected goals plus expected assisted goals per 90 and expected goals per 90, against the independent variables from equations (1) and (2). Due to the high collinearity between *realvalue_permil* and *realfee_permil*, regressions with both variables were not conducted. Instead, two sets of regressions were performed for each dependent variable, one containing the market value and the other containing the transfer fee. The instrumental variables used for the 2SLS estimates are *lastyeargoals, lastyearassists, age,* and *age_squared.* The regression results when using *npxg_xag_per90* as the dependent variable appear in Tables 2a and 2b. Table 2a uses the real transfer fee per million Euros (*realfee_permil*) and Table 2b uses the real value per million Euros (*realvalue_permil*) as alternative measures of player value. All the estimates were tested for heteroskedasticity, and all the models were adjusted for it. For results testing heteroskedasticity, refer to Appendix B.

This data demonstrates transfer fee and market value are the most significant predictors of *npxg_xag_per90*. For every 1 million Euros paid for a transfer fee, a club can expect an increase of about .00338 non-penalty expected goals plus expected assisted goals per 90, when using the random effects OLS model. This is equivalent to an increase of .0676 for an additional 20 million Euros added to a transfer fee. An EPL club looking to strengthen their attack through the transfer market will need to pay a significant amount if the goal is to become better than their rivals.

Age is not a statistically significant variable when determining attacking output. The coefficient for age is negative which goes against intuition. However, there has been a trend of EPL clubs buying players at younger ages, according to Worville (2021). The data supports the notion that players who are skillful enough to play in the top division should be considered in the transfer market. A skillful player, regardless of age, can generate attacking football. A football club with ambition to win trophies while constrained by a budget can look at any player regardless of age.

None of the years are statistically significant when running the fixed effects models and are all similar in magnitude. The dummy variables measuring where a player transferred from are not significant either, but the magnitude of the values follow a trend. The coefficients for *europe_dummy* are always smaller than the coefficient values for *epl_dummy and england_dummy*. Players transferring in from other European leagues are expected to generate less expected goals and expected assisted goals per game compared to players transferring in from other English clubs. This could be explained by the Union of European Football Association's (UEFA) association club coefficient rankings. This ranking is based on club results over the previous five seasons in European competitions. England is number one in the coefficient table which means English clubs have performed the best against other European clubs in European competitions UEFA (2024). Although the estimates are not significant at the .05 level, players transferring in from other European leagues generate less attacking output compared to players transferring in from English clubs.

Tables 3a and 3b present the regression results when *xg90* is used as the dependent variable. Table 3a uses the transfer fee per million Euros (*realfee_permil*) and Table 3b uses the real value per million Euros (*realvalue_permil*) as alternative measures of player value. Heteroskedasticity was tested and robust standard errors were used as needed. Appendix B contains heteroskedasticity test results when using *xg90* as the dependent variable.

Regardless of the choice between transfer fee or market value to be included in the model, none of the other variables were statistically significant at the .05 level. The estimates for age are negative and the trend for the league dummy variables can be seen in these models as well.

VI. Conclusion

This study was organized to analyze English Premier League attacking player transfers from 2017 to 2023 to examine if age was a significant factor in a player's attacking output the following year. The results of this study concluded with age not being statistically significant. The estimates for market value and transfer fee are statistically significant and both positive. A higher market value and transfer fee associated to a player are variables that predict an increase in non-penalty expected goals plus expected assisted goals per 90 and expected goals per 90. This result indicates that EPL clubs, on average, are paying more for more productive players. Indicator variables determining where a player was transferring from are not statistically significant at the .05 level, but the regression coefficients indicated that players transferring from other European leagues are less productive (output wise) than players transferring from other English clubs.

One limitation of this study is the number of observations and variables used in the model. Including transfers from other leagues and increasing the number of explanatory variables would likely provide a more accurate model. Expanding the model to include variables related to career appearances, international appearances, and the team they play for are ways to improve this model. Instead of taking season statistics for players, a different approach can compile player statistics from individual games. The model can test for inconsistencies in performances by taking the variance of player statistics during a stretch of games. Implementing these changes and many more can tell us additional information about the transfer market in football.

References

Alabi, M., & Urquhart, A. (2024). The financial impact of financial fair play regulation: Evidence from the English premier league. *International Review of Financial Analysis*, *92*, 103088. https://doi.org/10.1016/j.irfa.2024.103088

Dobson, S., & Goddard, J. A. (2011). *The Economics of Football* (pp. 22–100). Cambridge University Press.

FBref. (n.d.). *Football Statistics and History*. FBref.com. https://fbref.com/en/

FIFA Global Transfer Report 2023. (2024). In *InsideFIFA*. https://digitalhub.fifa.com/m/114622e4e17cf6a8/original/FIFA-Global-Transfer-Report-2023.pdf

Hadley, L., Poitras, M., Ruggiero, J., & Knowles, S. (2000). Performance Evaluation of National Football League Teams. *Managerial and Decision Economics*, *21*(2), 63–70. http://www.jstor.org/stable/3108334

Hall, S., Szymanski, S., & 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. https://doi.org/10.1177/152700250200300204

HICP - annual data (average index and rate of change). (2024). Europa.eu. https://ec.europa.eu/eurostat/databrowser/view/prc_hicp_aind/default/table?lang=en&category=prc.prc_hicp

Higgins, J. P. (2022). Evaluating the Value of an NFL Quarterback. *Soar.suny.edu*. http://hdl.handle.net/20.500.12648/12033

Jean-Baptiste Dherbecourt, & Bastien Drut. (2009). Who will go down this year ? The Determinants of Promotion and Relegation in European Soccer Leagues. *University of Paris West - Nanterre La Défense, EconomiX, EconomiX Working Papers*. https://www.researchgate.net/publication/46469693_Who_will_go_down_this_year_The_Determinants_of_Promotion_and_Relegation_in_European_Soccer_Leagues

Koop, G. (2001). Comparing the Performance of Baseball Players: A Multiple Output Approach. *Edinburgh School of Economics: Discussion Paper Series, Number 72*. http://www.econ.ed.ac.uk/papers/id72_esedps.pdf

Mead, J., O'Hare, A., & McMenemy, P. (2023). Expected goals in football: Improving model performance and demonstrating value. *PLOS ONE*, *18*(4), e0282295. https://doi.org/10.1371/journal.pone.0282295

Noll, R. G. (2002). The Economics of Promotion and Relegation in Sports Leagues: The Case of English Football. *Journal of Sports Economics*, *3*(2), 169–203. https://doi.org/10.1177/152700250200300205

Opta. (2022). *Opta Event Definitions*. Stats Perform. https://www.statsperform.com/opta-event-definitions/

Pantuso, G., & Hvattum, L. M. (2020). Maximizing performance with an eye on the finances: a chance-constrained model for football transfer market decisions. *TOP*, *29*(1), 583–611. https://doi.org/10.1007/s11750-020-00584-9

Poli, R., Besson, R., & Ravenel, L. (2021). Econometric Approach to Assessing the Transfer Fees and Values of Professional Football Players. *Economies*, *10*(1), 4. https://doi.org/10.3390/economies10010004

Rathke, A. (2017). An Examination of Expected Goals and Shot Efficiency in Soccer. *Journal of Human Sport and Exercise*, *12*(Proc2). https://doi.org/10.14198/jhse.2017.12.proc2.05

Transfermarkt. (2021, May 13). *Transfermarkt Market Value explained - How is it determined?* Transfermarkt.co.in. https://www.transfermarkt.co.in/transfermarkt-market-value-explained-how-is-it-determined-/view/news/385100

Transfermarkt.us. (n.d.). Transfermarkt.us. https://www.transfermarkt.us/

UEFA. (2011, January 11). *Financial fair play ensures football's stability*. UEFA.com; UEFA. https://www.uefa.com/news-media/news/01ed-0f86158ca3d7-227e20a44734-1000--financial-fair-play-ensures-football-s-stability/

UEFA. (2015, June 30). *Financial fair play: all you need to know*. UEFA.com. https://www.uefa.com/news-media/news/0253-0d7f34cc6783-5ebf120a4764-1000--financial-fair-play-all-you-need-to-know/

UEFA. (2024, December). *UEFA rankings*. UEFA.com; UEFA. https://www.uefa.com/nationalassociations/uefarankings/country/?year=2025

Worville, T. (2021, September 1). Transfer window analysed: Less spent, young players targeted and free agents have defined key moves. *The New York Times*. https://www.nytimes.com/athletic/2802812/2021/09/02/transfer-window-analysed-less-spent-young-players-targeted-and-free-agents-have-defined-key-moves/

Appendix A

Transfermarkt Market Value Variables

Most key factors:

Future prospects, age, performance at club and national team, level and status of the league, reputation/prestige, development potential, league-specific features, marketing value, number and reputation of interested clubs, performance potential, experience level, injury susceptibility, different financial conditions of clubs and leagues, general demand and "trends" on the market, general development of transfer fees, external factors such as the coronavirus pandemic and its consequences.

Individual transfer modalities:

Transfers via an option to buy/obligation to buy, loan fee, only part of transfer rights acquired, exit clause, buyback option, player swap deal, contract length, resale participation, bonus payments, improvement of financial balance.

Situational conditions:

Pressure situations such as competitive, success or financial, pressure, etc., will/desire/interests of the player, club does not sell to highest bidder, player goes on strike or similar, high salary, club wants to sell player.

Appendix B

Heteroskedasticity Test Results

Breusch-Pagan test for heteroskedasticity:

H_0: Constant Variance

H_1: Non-Constant Variance

Dependent variable: npxg_xag_per90

Independent variables: age, age_squared, realfee_permil, epl_dummy, england_dummy, europe_dummy

OLS RE: chi2(1) = 84.36, Prob > chi2 = 0.0000

2SLS RE: F(5, 278) = 29.92, Prob > F = 0.0000

Dependent variable: npxg_xag_per90

Independent variables: age, age_squared, realfee_permil, epl_dummy, england_dummy, europe_dummy, d_18, d_19, d_20, d_21, d_22, d_23

OLS FE: chi2(1) = 147.07, Prob > chi2 = 0.0000

2SLS FE: F(11, 272) = 19.22, Prob > F = 0.0000

Dependent variable: npxg_xag_per90

Independent variables: age, age_squared, realvalue_permil, epl_dummy, england_dummy, europe_dummy

OLS RE: chi2(1) = 78.13, Prob > chi2 = 0.0000

2SLS RE: F(5, 278) = 39.38, Prob > F = 0.0000

Dependent variable: npxg_xag_per90

Independent variables: age, age_squared, realvalue_permil, epl_dummy, england_dummy, europe_dummy, d_18, d_19, d_20, d_21, d_22, d_23

OLS FE: chi2(1) = 142.32, Prob > chi2 = 0.0000

2SLS FE: F(11, 272) = 21.99, Prob > F = 0.0000

Dependent variable: xg90

Independent variables: age, age_squared, realfee_permil, epl_dummy, england_dummy, europe_dummy

OLS RE: chi2(1) = 204.70, Prob > chi2 = 0.0000

2SLS RE: F(5, 278) = 64.34, Prob > F = 0.0000

Dependent variable: xg90

Independent variables: age, age_squared, realfee_permil, epl_dummy, england_dummy, europe_dummy, d_18, d_19, d_20, d_21, d_22, d_23

OLS FE: chi2(1) = 284.35, Prob > chi2 = 0.0000

2SLS FE: F(11, 272) = 31.76, Prob > F = 0.0000

Dependent variable: xg90

Independent variables: age, age_squared, realvalue_permil, epl_dummy, england_dummy, europe_dummy

OLS RE: chi2(1) = 193.80, Prob > chi2 = 0.0000

2SLS RE: F(5, 278) = 111.31, Prob > F = 0.0000

Dependent variable: xg90

Independent variables: age, age_squared, realvalue_permil, epl_dummy, england_dummy, europe_dummy, d_18, d_19, d_20, d_21, d_22, d_23

OLS FE: chi2(1) = 270.73, Prob > chi2 = 0.0000

2SLS FE: F(11, 272) = 42.54, Prob > F = 0.0000

Table 1 Descriptive Statistics of Variables

Variable	**Mean**	**SD**	**Min**	**Max**
npxg_xag_per90	.3989789	.2622181	0	3.23
xg90	.2797183	.257211	0	3.23
age	24.33451	3.834575	18	38
age_squared	606.8204	197.4705	324	1444
realvalue_permil	17.01544	17.27193	.1931248	128.4027
realfee_permil	19.38159	19.02703	0	109.0184
lastyeargoals	8.088028	6.840639	0	31
lastyearassists	3.911972	3.354655	0	17
epl_dummy	.2640845	.4416227	0	1
england_dummy	.2007042	.4012342	0	1
europe_dummy	.4859155	.5006839	0	1
d_18	.1197183	.3252049	0	1
d_19	.1478873	.3556149	0	1
d_20	.0950704	.2938299	0	1
d_21	.1232394	.3292921	0	1
d_22	.193662	.3958646	0	1
d_23	.1549296	.3624763	0	1
N	284			

TABLE 2a Dep. Variable = npxg_xag_per90 n=284

Variable	OLS RE	OLS FE	2SLS RE	2SLS FE
age	-0.0468	-0.0450	-0.0407	-0.0435
	(-1.06)	(-1.02)	(-0.92)	(-0.95)
age_squared	0.000912	0.000907	0.000775	0.000872
	(1.11)	(1.09)	(0.93)	(1.00)
realfee_permil	0.00338***	0.00340***	0.00271	0.00324*
	(5.33)	(5.15)	(1.74)	(2.06)
epl_dummy	-0.111	-0.123	-0.102	-0.121
	(-0.60)	(-0.64)	(-0.58)	(-0.66)
england_dummy	-0.115	-0.117	-0.116	-0.117
	(-0.60)	(-0.60)	(-0.61)	(-0.61)
europe_dummy	-0.154	-0.152	-0.146	-0.150
	(-0.80)	(-0.79)	(-0.80)	(-0.81)
d_18		-0.0101		-0.0109
		(-0.22)		(-0.24)
d_19		-0.0231		-0.0235
		(-0.56)		(-0.58)
d_20		-0.0247		-0.0243
		(-0.56)		(-0.56)
d_21		-0.0178		-0.0174
		(-0.30)		(-0.29)
d_22		-0.0440		-0.0437
		(-0.92)		(-0.94)
d_23		0.0694		0.0696
		(1.17)		(1.19)
Constant	1.047	1.015	0.987	0.999
	(1.50)	(1.46)	(1.45)	(1.45)
R^2	0.0687	0.0865	0.0667	0.0864

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 2b Dep. Variable = npxg_xag_per90 n=284

Variable	OLS RE	OLS FE	2SLS RE	2SLS FE
age	-0.0426	-0.0426	-0.0345	-0.0385
	(-0.98)	(-0.99)	(-0.79)	(-0.85)
age_squared	0.000772	0.000800	0.000608	0.000717
	(0.96)	(0.99)	(0.74)	(0.84)
realvalue_per mil	0.00359***	0.00371***	0.00252	0.00323
	(5.51)	(5.21)	(1.54)	(1.96)
epl_dummy	-0.104	-0.117	-0.0928	-0.112
	(-0.56)	(-0.61)	(-0.52)	(-0.60)
england_dum my	-0.106	-0.107	-0.109	-0.108
	(-0.54)	(-0.54)	(-0.57)	(-0.56)
europe_dumm y	-0.154	-0.153	-0.143	-0.148
	(-0.80)	(-0.80)	(-0.77)	(-0.79)
d_18		-0.0172		-0.0185
		(-0.38)		(-0.40)
d_19		-0.0366		-0.0362
		(-0.88)		(-0.89)
d_20		-0.0429		-0.0394
		(-0.97)		(-0.90)
d_21		-0.0335		-0.0305
		(-0.55)		(-0.48)
d_22		-0.0576		-0.0548
		(-1.22)		(-1.17)
d_23		0.0571		0.0591
		(0.94)		(0.99)
Constant	1.029	1.031	0.944	0.985
	(1.49)	(1.50)	(1.39)	(1.42)
R^2	0.0661	0.0855	0.0619	0.0847

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 3a Dep. Variable = xg90 n=284

Variable	OLS RE	OLS FE	2SLS RE	2SLS FE
age	-0.0323	-0.0345	-0.0276	-0.0326
	(-0.72)	(-0.77)	(-0.61)	(-0.70)
age_squared	0.000743	0.000815	0.000640	0.000774
	(0.88)	(0.96)	(0.74)	(0.85)
realfee_permil	0.00276***	0.00281***	0.00225	0.00262
	(4.32)	(4.18)	(1.41)	(1.65)
epl_dummy	-0.192	-0.201	-0.186	-0.198
	(-1.01)	(-1.02)	(-1.02)	(-1.04)
england_dummy	-0.149	-0.146	-0.149	-0.146
	(-0.75)	(-0.73)	(-0.76)	(-0.75)
europe_dummy	-0.203	-0.201	-0.198	-0.199
	(-1.04)	(-1.02)	(-1.05)	(-1.04)
d_18		-0.0205		-0.0214
		(-0.51)		(-0.53)
d_19		-0.00854		-0.00909
		(-0.23)		(-0.25)
d_20		-0.0579		-0.0574
		(-1.46)		(-1.48)
d_21		-0.0175		-0.0171
		(-0.29)		(-0.29)
d_22		-0.0448		-0.0444
		(-1.01)		(-1.03)
d_23		0.0527		0.0528
		(0.92)		(0.94)
Constant	0.741	0.763	0.696	0.744
	(1.04)	(1.09)	(1.00)	(1.06)
R^2	0.0568	0.0729	0.0557	0.0728

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 3b Dep. Variable = xg90 n=284

Variable	**OLS RE**	**OLS FE**	**2SLS RE**	**2SLS FE**
age	-0.0292	-0.0331	-0.0208	-0.0272
	(-0.66)	(-0.76)	(-0.46)	(-0.58)
age_squared	0.000636	0.000740	0.000465	0.000620
	(0.77)	(0.89)	(0.54)	(0.70)
realvalue_permil	0.00298***	0.00314***	0.00186	0.00245
	(3.75)	(3.64)	(1.11)	(1.48)
epl_dummy	-0.187	-0.197	-0.175	-0.189
	(-0.99)	(-1.00)	(-0.96)	(-0.99)
england_dummy	-0.141	-0.138	-0.145	-0.140
	(-0.71)	(-0.68)	(-0.73)	(-0.71)
europe_dummy	-0.204	-0.203	-0.193	-0.196
	(-1.04)	(-1.03)	(-1.02)	(-1.02)
d_18		-0.0262		-0.0281
		(-0.65)		(-0.70)
d_19		-0.0198		-0.0192
		(-0.54)		(-0.53)
d_20		-0.0735		-0.0685
		(-1.77)		(-1.65)
d_21		-0.0310		-0.0267
		(-0.51)		(-0.41)
d_22		-0.0565		-0.0525
		(-1.29)		(-1.19)
d_23		0.0422		0.0450
		(0.72)		(0.78)
Constant	0.730	0.783	0.641	0.717
	(1.04)	(1.13)	(0.92)	(1.01)
R^2	0.0562	0.0740	0.0514	0.0722

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$