TI 2026-089/V
Tinbergen Institute Discussion Paper

Age and Productivity in High-Skilled Work: Goal-scoring in the English Premier League

Jan van Ours [1]

1 Erasmus University Rotterdam

Age and Productivity in High-Skilled Work: Goal-scoring in the English Premier League

Jan C. van Ours*

October 31, 2026

Abstract

Professional football has many dimensions of productivity. This paper focuses on goal scoring and analyzes the relationship between goal scoring and the age of professional football players active in the English Premier League over an eight-season period. There are two dimensions of productivity in goal scoring. First, goal preparation in terms of creating scoring opportunities as measured by expected goals. Second, the conversion of expected goals into actual goals. The first dimension is the result of a combination of teamwork and individual skills, while the second dimension represents a purely individual skill. Age may affect both dimensions of goal scoring differently. The main finding of the analysis is that age has a non-linear effect on the productivity of goal scoring. For the youngest players the effect is positive, for the oldest players the effect is negative. In between age has no effect. The positive productivity effect at a young age comes from expected goals and could be related to a learning process. The negative productivity effect among older players could be caused by a deterioration of individual skills to convert expected goals into actual goals.

Keywords: Performance, pressure, heterogeneity, skills, football

JEL-codes: C93, D01, J81, Z20

Conflict of interest: The author declares no conflict of interest of any kind.

The data used in the analysis come from various public sources and will be made publicly available through zenodo.org.

*Erasmus School of Economics, Tinbergen Institute (Rotterdam) and ECASE (Erasmus Center for Applied Sports Economics), the Netherlands; CEPR (London). Address: Burg. Oudlaan 50, 3062 PA Rotterdam, the Netherlands; email: vanours@ese.eur.nl

1 Introduction

With aging populations and increasing statutory retirement ages, it is important to understand the relationship between labor productivity and age, particularly at the upper end of the age distribution. It is not easy to study productivity effects of aging, as productivity is rarely measured at the individual level; in regular work, productivity is often a group phenomenon. Data from professional sports can be helpful because the individual productivity of athletes is measured accurately and regularly.

This paper analyzes data from eight seasons of the English Premier League, focusing on the relationship between age and goal scoring. Scoring goals is an important dimension of a football player's productivity, but it is only one dimension. Providing assists, interrupting the opponent's play, and preventing the opponent from scoring are other important dimensions of productivity. After all, individual skills are highly valued and rewarded, but often teamwork determines who wins the game.

The main focus of this paper is goal scoring as a measure of productivity. Goal scoring consists of two dimensions. First, creating scoring opportunities as measured by expected goals. Second, finishing, that is, turning a scoring opportunity into a goal. Preparation is a combination of group productivity and individual productivity. Some players are capable of creating scoring opportunities entirely by themselves, starting with a dribble. However, usually the player who gets into a scoring position does so with the help of a teammate. A pass can turn into an assist if a goal is scored. A pass can look completely futile if the intended recipient runs in the wrong direction. Finishing, on the other hand, is an individual skill. Some players are simply better than others at converting a given scoring opportunity into an actual goal. In this paper, the focus is on the age effects on expected goals, converting expected goals into actual goals and actual goals. The number of expected goals in itself has two dimensions: quantity and quality. The number of shots is a quantity indicator, the expected goals per shot are an indicator of the quality of the shots. Furthermore, as part of the team production also the number of expected assists may be important. All indicators may be age dependent with the age dependence being different for different performance indicators.

The paper is set up as follows. Section 2 gives a brief overview of previous studies on the relationship between age and individual productivity. Section 3 presents a descriptive analysis of the relationship between age and productivity at the player-match level. The analysis is based on 32,770 observed shots of 500 football players. Section 4 presents the results of the empirical analysis of age and productivity, in which match characteristics are also taken into account. Section 5 concludes.

2 Previous Studies

Skirbekk (2004) provides a general overview of studies on the relationship between age and individual productivity, distinguishing between studies based on supervisors' evaluations, piece-rate studies, and studies based on employer–employee data. Each type of study has its own shortcomings. When based on supervisors' evaluations, the estimated relationship may be biased, since managers may reward older employees for their loyalty and past achievements rather than for their current productivity.

Using information on the quantity and quality of piece-rate work is more objective, but this type of work is not very common. Studies based on employer–employee data potentially cover a wide range of activities, but often one needs to rely on composite indicators of the age of the workforce, and reverse causality is a potential problem: expanding firms hire new workers who are on average younger than the current workforce, while contracting firms dismiss in particular younger, low-tenured workers. This creates a negative correlation between age and productivity that is caused by changes in the composition of the workforce.

Matched worker–firm studies often find that productivity does not change until higher ages. Van Ours and Stoeldraijer (2011), for example, using Dutch employer–employee data, conclude that the age profile of productivity is rather flat until age 50, after which productivity declines. Börsch-Supan and Weiss (2016), studying the relationship between age and productivity in work teams of a large German car manufacturer, find that productivity does not decline, at least up to age 60.

The relationship between age and the labor productivity of individual workers is difficult to establish because productivity is hard to measure at the level of individual workers. Sports data can help address this issue, as done in previous studies. Van Ours (2009), for example, analyzing data on amateur running, finds that performance declines after age 40 due to a decline in physical productivity. Bradbury (2009) finds for hitters and pitchers in US Major League Baseball that they reach peak performance around age 29. Castellucci et al. (2011) find that the productivity of F1 drivers peaks around age 30. **?** find hump-shaped relationships between age and performance in various sports. The age of peak performance is found to be 26.0 for track and field, 21.0 years for swimming, and 31.4 years for chess. Bertoni et al. (2015) concludes that the productivity of chess players, as measured by their Elo rating, peaks at a very early age, 21. Kalwij and De Jaegher (2023) find a hump-shaped relationship between age and performance in chess with a peak in the early thirties. In physical sports, performance often reaches a peak at an age at which the athletes' physical condition is at its maximum (Allen and Hopkins, 2015).

Data from professional football are also used to study the effects of age on performance. Torgler and Schmidt (2007), for example, find a nonlinear age effect on goal scoring in the top league of German professional football, with a turning point at age 31. They conclude that greater experience is overshadowed by lower physical performance. Dendir (2016), using performance ratings of players in the top four European professional football leagues, finds that productivity peaks between the ages of 25 and 27. Jamil and Kerruish (2020) use a wide variety of individual indicators, such as passes, interceptions, and dribbles, to analyze the productivity of players in the English Premier League. They conclude that forwards and wingers reach their peak performance before age 25, while age does not seem to affect the productivity of goalkeepers, defenders, or midfielders. Scarfe et al. (2024) use data from American professional football, concluding that productivity, as measured by minutes played and player ratings, peaks at age 20.

Recent examples in which professional football data are used to study the effects of age on productivity are Principe and van Ours (2026) and Tovar (2026). Principe and van Ours (2026) study players from the top Italian football league using a general rating of individual football players, concluding that there is an inverse U-shaped relationship between productivity and age, with productivity peaking at age 30. Tovar (2026) uses a wide variety of performance indicators, studying three leagues and finding peak performances at age 27–28 in the English Premier League and at about age 30 in the Spanish La Liga and the Colombian football league.

In regular work, it is not clear that productivity declines beyond a certain age. In professional football, this is not an issue: physical demands are high and productivity necessarily declines at higher ages. Finding an inverse U-shaped relationship between age and productivity is no surprise, as players retire from professional football long before they reach the usual retirement age for regular jobs. The end of a football player's career is usually determined by physical deterioration. Moreover, many football players have a second career in scouting, as transfer agents, managers, and so on.

3 Empirical Analysis

3.1 Data

The empirical analysis is based on information about the goal scoring performance of professional football players in eight seasons of the English Premier League, from 2017/18 to 2024/25. The data on the dependent variables, that is information about shots and related performance indicators are from Understat.com (more details of the data sources and definitions are provided in Appendix A).

The focus of the analysis is on individual players who are in the sample at least 25 times, that is, they took at least 25 shots aiming to score a goal. Table 1 provides descriptives of the data at player-match level. As indicated in the introduction a variety of performance indicators is used. A shot is defined as a deliberate attempt by a player to score, made with any legal part of the body, regardless of whether it is on target, blocked, saved or scored. Shots include headers and direct free kicks. As shown in Table 1 on average had 1.9 shots per match ranging from one to 12. An expected goal (xG) measures the quality of chances created. For each shot, it is the probability that the shot results in a goal. On average every shot created 0.1 expected goal. Per player the number of expected goals created was 0.2 per match. Table 1 shows that the players on average scored 0.2 goals (G) per match (excluding own goals and goals resulting from a penalty kick). The range was from zero to four goals per match. As was to be expected the difference between goals and expected goals was about zero with a range from -2.1 to +2.9. Finally, the upper part of Table 1 provides information about expected assists. An expected assist (xA) measures the likelihood that a completed pass will become a goal assist, based on the probability that the resulting shot is converted into a goal. The expected assist value is equal to the value of the expected goal of the shot resulting from the pass and is credited to the passer. The average expected assist per player was equal to 0.12.

Table 1: **Descriptives Player-Match Level Analysis**

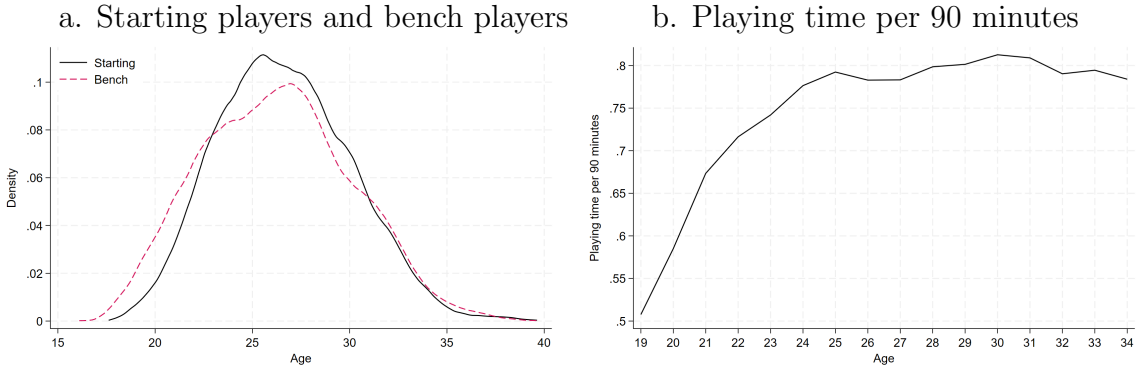
Variable	Mean	Std. dev.	Min	Max
Shots	1.943	1.259	1	12
xG/shot	0.107	0.125	0.005	0.975
xG	0.219	0.286	0.005	2.940
G-xG	-0.015	0.386	-2.115	2.896
G	0.203	0.454	0	4
xA	0.122	0.217	0	2.350
Playing time	0.853	0.267	0.011	1
Home-BCD	0.076	0.265	0	1
Home-Attendance	0.448	0.497	0	1
Δ Elo/1000	0.040	0.202	-0.617	0.617
Age/100	0.266	0.003	0.161	0.396

Note: Based on 32,770 observations of 500 players with at least 25 observations; playing time in minutes/90

As shown in the bottom part of Table 1, the average playing time per player was equal to 0.85 (of a full match). Of the players in the sample 7.6% of the observations are from a home match played behind closed doors while 44.8% is from a home match played with a stadium crowd present. The difference in Elo-ratings indicates the difference in strength between the two teams and is from the perspective of the player. The difference is slightly positive indicating that there are more observations of players from better teams. The

average age of the players in the sample is 26.6 years, with a range from 16.1 years to 39.6 years of age. Panel a of Figure 1 shows the age distribution of the football players

Figure 1: **Age Distribution and Playing Time**

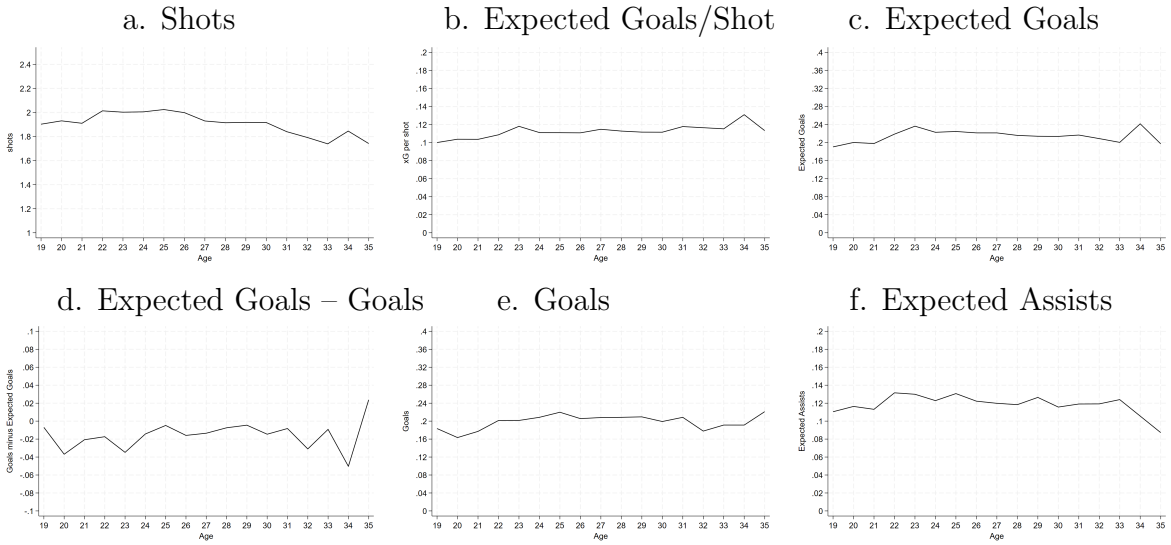


Note: All players who participated in at least one match

distinguishing between players who started in the match and players who started as a reserve waiting on the bench for their turn. On average, players starting on the bench are younger than players who started a match.

Panel b of Figure 1 shows that younger players have less playing time than older players. At age 19, players played about 50% of the 90 minutes in a regular match. Playing time increases with age until at age 24 on average players played close to 80% of the playing time of a full match.

Figure 2: **Average Productivity by Age**



Note: Appendix Figure C.1 distinguishes average productivity by age and position

Figure 2 provides information about the unconditional relationships between each of the productivity indicators and the age of the players. The relationships vary by

indicator but many have a positive relationship with age for young players and a negative relationship for older players. The peak of the productivity is early for shots and much later for expected goals per shot.

3.2 Model

Age may have a non-linear, inverse U-shaped effect on productivity. To investigate this, age is included as a linear and a quadratic term. The baseline estimates are based on the following specification:

$$Y_{im} = \alpha_i + \beta T_{im} + \gamma X_m + \delta_1 Age_{im} + \delta_2 Age_{im}^2 + \varepsilon_{im} \tag{1}$$

where Y_{im} represents six dependent variables related to individual player i in match m: total number of shots, expected goals per shot (xG/shot), expected goals (xG), goals (G), the difference between the two ($G - xG$) and expected assists (xA). Furthermore, T_{im} is the playing time of player i in match m, X_m is a vector of match characteristics, Age_{im} is a continuous variable representing the age of player i in match m; α_i is a vector of player fixed effects, β, δ_1 and δ_2 are parameters, γ is a vector of parameters and ε_{im} is an error term. The match characteristics included in the analysis are from the perspective of the player: whether the match was played at home behind closed doors or with a stadium audience. Furthermore, equation (1) includes the difference in strength between the two teams as measured by the difference in Elo-rating at the start of the match. Since the productivity information is about individual players standard errors are clustered at player level.

3.3 Baseline parameter estimates

The parameter estimates are shown in Table 2. The first column shows the estimates for the number of shots. As was to be expected shots increase with playing time. Players in home matches have more shots especially in the presence of a stadium crowd. Players in stronger teams provide more shots while age has an inverse u-shape non-linear effect with a peak at age 25.6. Players in home matches with a stadium crowd produce more expected goals per shot, an indicator of the quality of the shot. Players in better teams also produce higher quality shots but age has no significant effects. Ignoring the significance the maximum is at age 43.2, but this is estimated very imprecise.

From the third column it appears that expected goals scoring increases with playing time. Furthermore, playing a home match in front of a stadium crowd significantly increases the number of expected goals with 0.029, an increase of 13%. More expected

goals are created if the difference in strength between two teams increases. Age has a non-linear effect; expected goals significant increase with age and significantly decrease with age-squared. The maximum number of expected goals is at age 34.1. The fourth column shows that the transition of expected goals to actual goals is not influenced by the playing time while there is no home advantage. In fact, playing at home in the presence of a stadium crowd has a significant negative effect on the transition of expected goals to actual goals. The difference in Elo-ratings has no significant effects. Age has a non-linear effect. The productivity in terms of the transition of expected goals into actual goals has a peak at age 23.9. The fifth column of Table 2 shows similar parameter estimates for the actual number of goals as for the expected number of goals. Playing time has positive effect while the home advantage for expected goals is similar to the home advantage for actual goals. Playing at home in front of a stadium crowd makes players more productive in terms of expected goals as well as actual goals. The effect of differences in strength between two teams is significant also for actual goals. The peak age for actual goal scoring is at 27.6. Finally, the parameter estimates for expected assists in the sixth column show that these are very similar to those for the expected goals. Here the peak in the effect is at age 32.

Table 2: **Baseline Parameter Estimates**

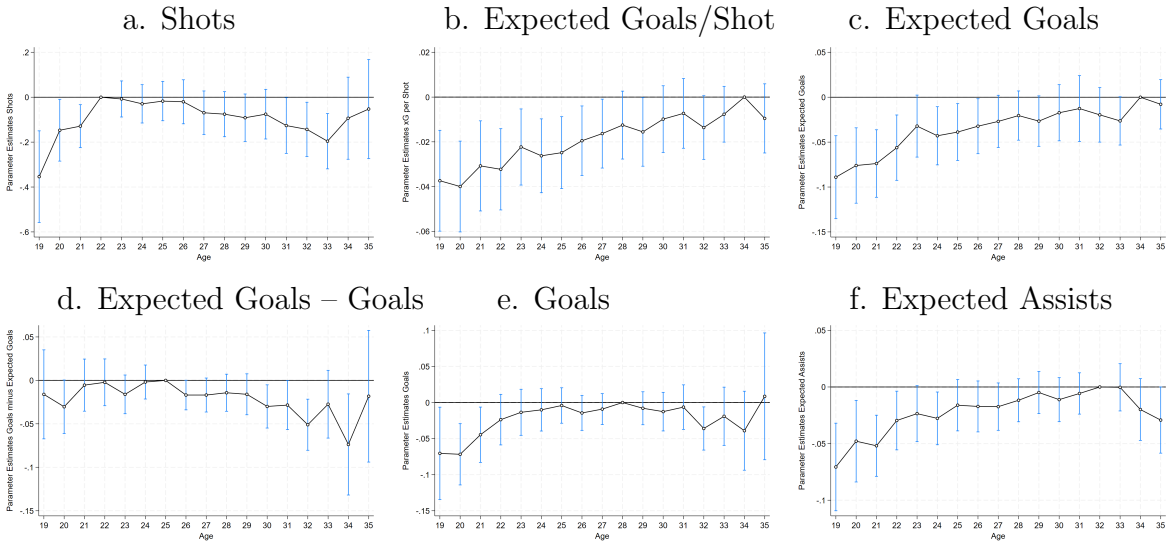
	Shots	xG/shot	xG	G-xG	G	xA
	(1)	(2)	(3)	(4)	(5)	(6)
Playing time	1.302***	−0.006*	0.150***	0.002	0.152***	0.134***
	(0.037)	(0.003)	(0.008)	(0.008)	(0.010)	(0.006)
Home-BCD	0.053**	−0.004	0.002	−0.007	−0.005	0.004
	(0.025)	(0.002)	(0.005)	(0.008)	(0.009)	(0.004)
Home-Attendance	0.200***	0.005***	0.029***	−0.009**	0.020***	0.024***
	(0.015)	(0.001)	(0.003)	(0.004)	(0.005)	(0.002)
Δ Elo/1000	0.807***	0.021***	0.134***	0.010	0.144***	0.109***
	(0.052)	(0.004)	(0.011)	(0.014)	(0.016)	(0.009)
Age/100	9.615**	0.588	2.038**	2.526**	4.564***	2.039***
	(4.402)	(0.432)	(0.891)	(1.143)	(1.271)	(0.774)
$(\text{Age}/100)^2$	−18.743**	−0.681	−2.992*	−5.283**	−8.276***	−3.185**
	(7.911)	(0.779)	(1.579)	(2.059)	(2.271)	(1.395)
Maximum at age	25.6***	43.2***	34.1***	23.9***	27.6***	32.0***
	(1.6)	(18.3)	(3.7)	(1.9)	(1.0)	(2.5)

Based on 32,770 observations of 500 players with at least 25 observations; playing time in minutes/90; standard errors clustered at player level in parentheses; standard errors maximum age calculated using the Delta-method; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.1 in Appendix B shows parameter estimates in which age and age-squared are replaced by age fixed effects. The non-age parameter estimates are hardly affected by this. The estimates for the age fixed effects are presented in Figure 3. To illustrate whether or not the indicators have a peak age differing from other ages the fixed age effects are normalized to zero at the peak age. As shown, initially at lower ages players

increase performance with age for shots, expected goals per shot, for expected goals, for goals and for expected assists. For the transition from expected goals to actual goals there is no such increase while at the high end of the age distribution performance decreases. This decrease in performance at higher ages is also present for the other indicators but not always substantially.

Figure 3: **Parameter Estimates Age Fixed Effects**



Note: At the peak age the effects are normalized to zero.
Based on parameter estimates presented in Table B.1

3.4 Age intervals

A quadratic specification is a functional form that imposes symmetry around an extreme value, in this case a maximum. As shown in Figure 3 such a relationship is not very likely here. As an alternative, to study how age affects productivity separate estimates are done splitting up the sample in three groups according to percentiles of the age distribution: below the 25^{th} percentile, from the 25^{th} to the 75^{th} percentile and above the 75^{th} percentile. Over these intervals, age is included as a linear term.

Table 3 shows the parameter estimates. Apart from the age effects many of the parameter estimates show a similar pattern across the three age groups. Playing time has a positive effect of the number of shots, expected goals, actual goals and expected assists but not on the quality of the shots nor on the productivity of transferring expected goals into actual goals. In other words, when players have more time on the pitch they shoot more, score more expected and actual goals and provide more expected assists but they are not more efficient. Playing a home match in an empty stadium has hardly any effect on any of the productivity measures. This is different for home matches played in

the presence of stadium crowds. If this is the case, shots, the quality of shots as well as expected goals and expected assists increase. For the youngest age group there is a negative effect on the transfer of expected goals to goals of home matches with a stadium crowd. This causes the number of actual goals not to be affected by home matches with stadium crowds. Perhaps young players feel too much pressure and choke because of that. For the other age groups the productivity of transferring expected goals into actual goals is not influenced by playing a home match and therefore there is a positive home advantage for actual scoring in the presence of a stadium crowd.

Table 3: **Parameter Estimates Three Age Intervals**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
a. Age<24.04						
Playing time	1.384***	−0.014**	0.142***	−0.006	0.135***	0.132***
	(0.068)	(0.007)	(0.014)	(0.015)	(0.018)	(0.010)
Home-BCD	0.041	−0.000	0.006	0.007	0.013	−0.002
	(0.060)	(0.006)	(0.011)	(0.017)	(0.019)	(0.008)
Home-Attendance	0.245***	0.006*	0.034***	−0.024***	0.010	0.029***
	(0.027)	(0.003)	(0.006)	(0.008)	(0.009)	(0.006)
Δ Elo/1000	0.845***	0.017	0.138***	−0.023	0.115***	0.102***
	(0.091)	(0.010)	(0.024)	(0.032)	(0.034)	(0.017)
Age/100	3.599**	0.303*	1.001***	0.077	1.079**	0.776**
	(1.698)	(0.154)	(0.353)	(0.481)	(0.490)	(0.319)
b. 24.04≤Age≤28.96						
Playing time	1.275***	−0.002	0.157***	0.002	0.159***	0.131***
	(0.047)	(0.005)	(0.011)	(0.013)	(0.014)	(0.008)
Home-BCD	0.051	−0.004	0.004	−0.019	−0.015	0.003
	(0.036)	(0.003)	(0.008)	(0.012)	(0.014)	(0.006)
Home-Attendance	0.206***	0.003*	0.030***	−0.006	0.023***	0.022***
	(0.020)	(0.002)	(0.004)	(0.007)	(0.008)	(0.003)
Δ Elo/1000	0.807***	0.022***	0.132***	0.017	0.149***	0.125***
	(0.077)	(0.006)	(0.016)	(0.020)	(0.024)	(0.012)
Age/100	−2.127*	0.271**	0.349	−0.449	−0.101	0.371*
	(1.255)	(0.106)	(0.248)	(0.321)	(0.372)	(0.214)
c. Age>28.96						
Playing time	1.229***	−0.010	0.143***	−0.010	0.133***	0.129***
	(0.065)	(0.007)	(0.015)	(0.017)	(0.021)	(0.011)
Home-BCD	0.074	−0.004	−0.004	0.006	0.002	0.015*
	(0.052)	(0.004)	(0.010)	(0.021)	(0.020)	(0.009)
Home-Attendance	0.152***	0.006**	0.023***	0.000	0.023**	0.023***
	(0.029)	(0.002)	(0.006)	(0.008)	(0.009)	(0.004)
Δ Elo/1000	0.789***	0.009	0.108***	0.009	0.117***	0.078***
	(0.094)	(0.009)	(0.021)	(0.030)	(0.034)	(0.019)
Age/100	−1.071	0.074	−0.012	−0.950**	−0.962**	−0.050
	(1.428)	(0.122)	(0.322)	(0.370)	(0.443)	(0.187)

Players with at least 25 observations overall; panel a 8,169 observations of 242 players; panel b 16,391 observations of 396 players; panel c 8,210 observations of 227 players; standard errors clustered at the player level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Differences in strength between two teams as measured by differences in Elo-ratings have positive effects on almost indicators except for the transition of expected to actual goals. Note that this is conditional on the player fixed effects.

The main differences in parameter estimates between the three age groups are related to the age effects. By and large, the differences are as follows. For the youngest group age has a positive effect, for the middle group age has no effect while for the oldest players age has a negative effect. This is in line with an inverse U-shaped age pattern albeit with a plateau for the players with an intermediate age. For the youngest players one year increase in age represents 0.01 increase in goal scoring. From an average of 0.2 goals per match per player this implies a five percent increase. At the high end of the age distribution one year older implies a five percent drop in goal scoring. The components of goal scoring are also very age dependent. The positive age effect for the youngest age group on goal scoring is predominantly caused by players who are getting better at creating expected goals. This could be both because they learn where to find better scoring opportunities or because they learn how their team mates play and what is expected from them. Presumably, the effect is more an effect of experience rather than age. Among the older players age has a negative effect on goal scoring which is predominantly caused by a negative age effect on the transition of expected goals to actual goals.

The age effect of goal scoring consists of two dimensions. First, the productivity related to the combination of team work and individual skills in the creation of scoring opportunities and the individual productivity of transferring these opportunities into goals. The parameter estimates in Table 3 suggest that the positive age effects at a young age are related to the player learning to benefit from team work activities while the negative age effects at an older age are related to the loss in individual productivity of transferring scoring opportunities into goals.

In the middle age range the effects are specific to the performance indicator. There is, for example, a significant negative age effects on the number of shots while the age effect on the expected goals per shot is positive.

4 Sensitivity Analysis

4.1 Additional estimates

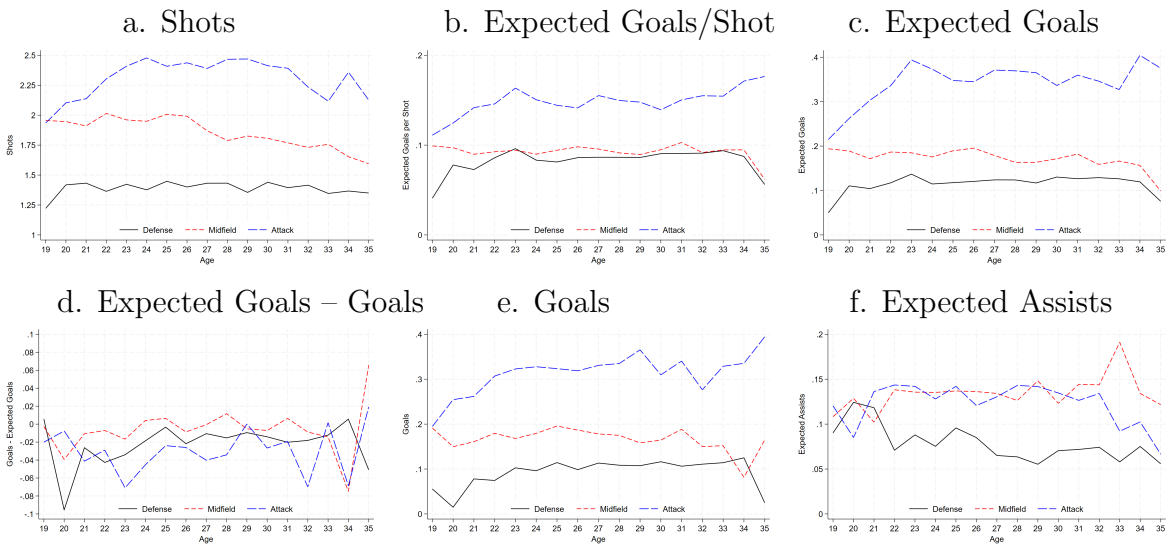
To illustrate the sensitivity of the main findings a robustness analysis was done of which the details are shown in Appendix B. First, there is the issue to what extent performance of players is affected if they change clubs. It may be that their performance increases or

decreases if they move to a better or worse playing club. If this would be the case then part of the age effect may be related to job mobility. To investigate this, in addition to player fixed effects club fixed effects are introduced. The parameter estimates are shown in Table B.2. Many of the parameter estimates are very similar to those presented in Table 2 including the ages where the various performance indicators have their maximum values.

Second, as shown in Section 3 the playing time varies by age with young players playing substantially less than older players. It could be that part of the age effects are due to differences in playing time. To investigate whether this is indeed an issue, estimates were done without playing time as explanatory variable. The results are shown in Table B.3. The parameter estimates are very much the same as those presented in Table 2. There are some differences with respect to the age at which an performance indicator reaches it's maximum. This is three to four years older for expected goals and expected assists.

Finally, age-related performance by player position is investigated (see for details Table B.4). Figure 4 presents overviews of the relationship between age and the performance indicators distinguished by position of the players.

Figure 4: **Productivity by Position and Age**



Note: Match averages; goal keepers excluded

Panel a of Figure 4 shows the relationships for the expected goals. Players in attacking positions have a substantially higher number of expected goals per scoring attempt. There is an increase from 0.25 at age 20 to 0.30 at age 23. After that the average stays at 0.30. For players in midfield or defense there is no clear age gradient with an average of about 0.15 for midfield players and 0.10 for players in defense. Panel b of Figure 4 shows that

players in defense have lower transfer rates from expected goals to actual goals but the differences are small. Because of this, panel c showing the actual number of goals per attempt is very similar to panel a.

Table B.5 shows separate parameter estimates of the performance indicators for players in the three positions according to equation (1). As was to be expected, playing time has the largest effects for attacking players and the smallest effects for defense players for shots, expected goals and actual goals. Similar orderings occur for players who are participating in a home match with a stadium crowd present and for the differences in strength between two teams. In terms of the age profiles the number of shots peaks late for defenders and earlier for midfielders while for attacking players there is a flat age profile. The same ordering is present for transferring expected goals into actual goals and actual goals. For expected goals the peaks occur at a similar age irrespective of the player position. By and large, there are clear difference in determinants of performance by player position but the main results are very much the same as for the baseline estimates.

4.2 Selective attrition?

In the main analysis player fixed effects are included to account for unobserved differences in ability. The analysis focuses on the relationship between performance and age conditional on the player fixed effects. However, conditioning on player fixed effects does not necessarily mean that the relationship between age and performance is estimated unbiased. Here, two potential types of selectivity are investigated. The first is selection on shooting which may occur during a season. The second is selection on performance which may occur across seasons.

To investigate the first type of selectivity all player-match appearances are included in a sensitivity analysis. The parameter estimates including all players – also those with zero shots – are shown in Table B.6. The magnitude of some parameter estimates differs somewhat from those presented in Table 2 but the ages for which the maximum is obtained for the all performance variables are very much the same.

The second type of selectivity may occur at the high end of the age distribution. There could be selective attrition. If players with a larger age-related drop in performance are more likely to dropout the sample the age-related decline is underestimated.

To investigate the potential relevance the following equation is estimated:

$$D_{i,s+1} = \alpha_i^D + \beta^D \Delta xG_{i,s} + \varepsilon_{i,s+1}^D \tag{2}$$

where $D_{i,s+1}$ is an indicator of the dropout of player i in season $s+1$, $\Delta xG_{i,s}$ indicates the change in the number of expected goals scored by player i from season $s-1$ to season

s; α_i^D represents a vector of player fixed effects, β^D is the parameter of main interest and ε is an error term. If $\beta^D < 0$, a player is more likely to drop out the sample if his goal scoring declined over the last two seasons.

Table 4: Parameter estimates dropout of sample

Variable	Logit model		Linear probability model	
Δ xG	-0.053	(0.140)	-0.056	(0.151)
Δ G-xG	0.026	(0.107)	0.029	(0.121)
Δ G	-0.079	(0.103)	-0.087	(0.113)

Note: Based on 763 observations of 212 players; logit model: marginal effects; standard errors clustered at the player level in parentheses

Since dropping out has a value of zero or one, equation 2 is estimated using a logit model. By way of sensitivity analysis equation 2 is also estimated using a linear probability model. All parameter estimates in Table 4 are insignificantly different from zero. Neither the change in the expected goals nor the change in the difference between actual goals and expected goals or the difference in the number of actual goals has an effect on the probability to drop out the sample. Dropping out of the sample, i.e. no longer playing in the Premier League, does not seem to be related to recent changes in performance. All in all, selective attrition from the panel does not seem to be an issue.

5 Conclusions

The relationship between age and the labor productivity of individual workers is difficult to establish for regular jobs, if only because productivity is hard to measure at the level of individual workers. This situation is radically different for sports data, where the individual performance of athletes is measured accurately, frequently, and often over a long period of time. This paper uses goal scoring as an indicator of individual productivity and analyzes the relationship between goal scoring and age among professional football players active in the English Premier League over an eight-season period. The main finding is that there is a nonlinear relationship between age and productivity, with a plateau in the age range from 24 to 29. This relationship is observed for actual goals but also for expected goals and for the conversion of expected goals into actual goals.

The age effect of goal scoring has two dimensions. First, the productivity related to the combination of teamwork and individual skills in creating scoring opportunities. Second, the individual productivity of transferring goal-scoring opportunities into goals. The empirical analysis presented suggests that the positive age effects at a young age are related to players learning to benefit from teamwork, while the negative age effects

at an older age are related to the loss of individual productivity in converting scoring opportunities into goals. For the youngest players, age has positive effects; for the middle group, age has no effect; and for the oldest players, age has negative effects. The positive age effect for the youngest age group on goal scoring is predominantly caused by players becoming better at creating expected goals. Expected goals are assigned to individual players, but they are in fact the outcome of the interaction between team productivity and individual skills. The positive age effect on expected goals could come from a learning process through which players learn where to find better scoring opportunities or because they learn how their teammates play and what is expected of them. This suggests that the positive effect on productivity is related to experience rather than an age effect. Among older players, age has a negative effect on goal scoring, which is predominantly caused by a negative age effect on the purely individual productivity involved in converting expected goals into actual goals.

As with other economic analysis based on sports data a relevant question is to what extent the main findings have external value, i.e. in regular work activities. For the current analysis of the relationship between age and goal scoring as a measure of productivity there is. As with regular jobs, as workers grow older their contribution to the production process changes. Whereas at young age workers have to learn of the job at a higher age they can provide support to younger players. For young players it is individual productivity that matters; for older players there are positive productivity spillovers to younger players.

References

Allen, S. V. and W. G. Hopkins (2015). Age of peak competitive performance of elite athletes: A systematic review. *Sports Medicine 45*, 1431–1441.

Berthelot, G., S. Len, P. Hellard, M. Tafflet, M. Guillaume, J.-C. Vollmer, B. Gager, L. Quinquis, A. Marc, and J.-F. Toussaint (2012). Exponential growth combined with exponential decline explains lifetime performance evolution in individual and human species. *Age 34*(4), 1001–1009.

Bertoni, M., G. Brunello, and L. Rocco (2015). Selection and the age-productivity profile. Evidence from chess players. *Journal of Economic Behavior & Organization 110*, 45–58.

Börsch-Supan, A. and M. Weiss (2016). Productivity and age: Evidence from work teams at the assembly line. *The Journal of the Economics of Ageing 7*, 30–42.

Bradbury, J. C. (2009). Peak athletic performance and ageing: Evidence from baseball. *Journal of Sports Sciences 27*(6), 599–610.

Castellucci, F., M. Padula, and G. Pica (2011). The age-productivity gradient: Evidence from a sample of F1 drivers. *Labour Economics 18*(4), 464–473.

Dendir, S. (2016). When do soccer players peak? A note. *Journal of Sports Analytics 2*(2), 89–105.

Elo, A. E. (1978). *The Rating of Chessplayers, Past and Present*. London Batsford.

Jamil, M. and S. Kerruish (2020). At what age are English Premier League players at their most productive? A case study investigating the peak performance years of elite professional footballers. *International Journal of Performance Analysis in Sport 20*(6), 1120–1133.

Kalwij, A. and K. De Jaegher (2023). The age-performance relationship for a cognitive-intensive task: Empirical evidence from chess grandmasters. *Sports Economics Review 2*, 100010.

Principe, F. and J. C. van Ours (2026). Labor market dynamics in a highly competitive industry. *Labour Economics 98*, 102845.

Scarfe, R., C. Singleton, A. Sunmoni, and P. Telemo (2024). The age-wage-productivity puzzle: Evidence from the careers of top earners. *Economic Inquiry 62*(2), 584–606.

Skirbekk, V. (2004). Age and individual productivity: A literature survey. *Vienna Yearbook of Population Research 2*, 133–153.

Torgler, B. and S. L. Schmidt (2007). What shapes player performance in soccer? Empirical findings from a panel analysis. *Applied Economics 39*(18), 2355–2369.

Tovar, J. (2026). The role of age in labor productivity; how age matters in professional football. *Journal of Sports Economics 27*(4), 426–445.

Van Ours, J. C. (2009). Will you still need me when I'm 64? *De Economist 157*, 441–460.

Van Ours, J. C. and L. Stoeldraijer (2011). Age, wage and productivity in Dutch manufacturing. *De Economist 159*, 113–137.

Appendix A: Details on the Data

A1: Variables

1. Shot: Any deliberate attempt by a player to score, made with any legal part of the body, regardless of whether it is on target, blocked, saved or scored. This includes headers and direct free kicks.

2. G = Goal(s) scored excluding own goals and goals resulting from a penalty kick.

3. xG = Expected goal: Statistical measure of the quality of chances created. For each shot, it is the probability that the shot could have resulted in a goal.

4. xA = Expected assist: Measure of the quality of a pass that leads directly to a shot. The xA of a pass equals the xG of the resulting shot and is credited to the passer, regardless of whether the shot is converted into a goal.

5. Home-BCD = Home team played behind closed doors.

6. Home-Attendance = Home team played in the presence of a stadium crowd.

7. Δ Elo-rating = Difference in Elo-ratings between two teams at the start of a match. Elo-ratings were originally used to indicate the relative strength of chess players. (Elo, 1978) but nowadays they are used in many sports to indicate the strength of an athlete or team.

A2: Data sources

- Shots, goals, expected goals, expected assists, playing time: Understat.com. Expected goals are estimated with a neural network trained on more than 100,000 shots, using over 10 characteristics of each shot, such as its location, the body part used and the type of play it came from.

- Elo-rating: Elofootball.com

- Attendance: Football-data.co.uk

A3: Player-season observations

In the sample of 500 players with at least 25 observations, there are 8 players who are observed for only one season while there are 42 who are observed during all eight seasons.

Seasons	Players	Seasons	Players
1	8	5	85
2	66	6	63
3	108	7	51
4	77	8	42
		Total	500

Appendix B: Additional Parameter Estimates

B1: Age Fixed Effects

Table B.1: **Parameter Estimates with Age Fixed Effects**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
Playing time	1.293***	−0.007**	0.149***	0.003	0.152***	0.134***
	(0.037)	(0.003)	(0.008)	(0.009)	(0.010)	(0.006)
Home-BCD	0.052**	−0.004	0.002	−0.008	−0.006	0.005
	(0.025)	(0.002)	(0.005)	(0.008)	(0.009)	(0.004)
Home-Attendance	0.200***	0.005***	0.030***	−0.009*	0.021***	0.024***
	(0.014)	(0.001)	(0.003)	(0.005)	(0.005)	(0.002)
Δ Elo/1000	0.814***	0.021***	0.134***	0.009	0.144***	0.109***
	(0.052)	(0.004)	(0.011)	(0.014)	(0.016)	(0.009)
Age 19	−0.354***	−0.037***	−0.089***	−0.016	−0.070**	−0.071***
	(0.104)	(0.011)	(0.023)	(0.026)	(0.033)	(0.020)
Age 20	−0.147**	−0.040***	−0.076***	−0.030*	−0.072***	−0.048***
	(0.070)	(0.010)	(0.021)	(0.016)	(0.022)	(0.018)
Age 21	−0.129***	−0.031***	−0.074***	−0.005	−0.045**	−0.052***
	(0.049)	(0.010)	(0.019)	(0.015)	(0.020)	(0.014)
Age 22		−0.032***	−0.056***	−0.002	−0.024	−0.030**
		(0.009)	(0.019)	(0.014)	(0.018)	(0.013)
Age 23	−0.008	−0.022**	−0.032*	−0.016	−0.014	−0.024*
	(0.041)	(0.009)	(0.018)	(0.011)	(0.016)	(0.013)
Age 24	−0.029	−0.026***	−0.043***	−0.002	−0.010	−0.028**
	(0.043)	(0.008)	(0.016)	(0.010)	(0.015)	(0.012)
Age 25	−0.017	−0.025***	−0.039**		−0.004	−0.016
	(0.045)	(0.008)	(0.016)		(0.012)	(0.012)
Age 26	−0.020	−0.020**	−0.032**	−0.017*	−0.014	−0.017
	(0.050)	(0.008)	(0.016)	(0.009)	(0.012)	(0.011)
Age 27	−0.069	−0.016**	−0.027*	−0.017*	−0.009	−0.017
	(0.049)	(0.008)	(0.015)	(0.010)	(0.011)	(0.011)
Age 28	−0.075	−0.013	−0.020	−0.014		−0.012
	(0.051)	(0.008)	(0.014)	(0.011)		(0.010)
Age 29	−0.091*	−0.016**	−0.027*	−0.016	−0.008	−0.005
	(0.054)	(0.008)	(0.014)	(0.012)	(0.012)	(0.009)
Age 30	−0.076	−0.010	−0.017	−0.030**	−0.013	−0.011
	(0.056)	(0.008)	(0.016)	(0.013)	(0.014)	(0.010)
Age 31	−0.126**	−0.007	−0.013	−0.028**	−0.006	−0.006
	(0.063)	(0.008)	(0.019)	(0.014)	(0.016)	(0.009)
Age 32	−0.143**	−0.014*	−0.020	−0.051***	−0.036**	
	(0.061)	(0.007)	(0.015)	(0.015)	(0.015)	
Age 33	−0.196***	−0.008	−0.026*	−0.027	−0.019	−0.000
	(0.063)	(0.006)	(0.014)	(0.020)	(0.021)	(0.011)
Age 34	−0.094			−0.074**	−0.039	−0.020
	(0.093)			(0.030)	(0.028)	(0.014)
Age 35	−0.053	−0.010	−0.008	−0.018	0.009	−0.029**
	(0.112)	(0.008)	(0.014)	(0.039)	(0.045)	(0.015)

Based on 32,452 observations of 498 players with at least 25 observations; playing time in minutes/90; age effects maximum is omitted category; standard errors clustered at the player level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B2: Club Fixed Effects

Table B.2: **Baseline Parameter Estimates with Club Fixed Effects**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
Playing time	1.289***	−0.006*	0.150***	0.001	0.151***	0.133***
	(0.037)	(0.003)	(0.008)	(0.008)	(0.010)	(0.006)
Home-BCD	0.052**	−0.004*	0.001	−0.007	−0.005	0.005
	(0.025)	(0.002)	(0.005)	(0.008)	(0.009)	(0.004)
Home-Attendance	0.200***	0.005***	0.029***	−0.009**	0.020***	0.024***
	(0.015)	(0.001)	(0.003)	(0.004)	(0.005)	(0.002)
Δ Elo/1000	0.868***	0.016***	0.129***	0.014	0.144***	0.111***
	(0.051)	(0.004)	(0.011)	(0.014)	(0.017)	(0.010)
Age/100	11.517***	0.495	1.962**	2.571**	4.533***	2.000**
	(4.260)	(0.410)	(0.866)	(1.110)	(1.223)	(0.775)
$(\text{Age}/100)^2$	−21.903***	−0.548	−2.939*	−5.452***	−8.391***	−3.107**
	(7.680)	(0.745)	(1.545)	(1.975)	(2.159)	(1.408)
Maximum at age	26.3***	45.1	33.4***	23.6***	27.0***	32.2***
	(1.3)	(24.8)	(3.6)	(2.0)	(1.1)	(2.8)

Based on 32,770 observations of 500 players with at least 25 observations; playing time in minutes/90; standard errors clustered at the player level in parentheses; standard errors of maximum age calculated using the Delta-method; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B3: Without Playing Time as Explanatory Variable

Table B.3: **Baseline Parameter Estimates without Playing Time**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
Home-BCD	0.065**	−0.004	0.004	−0.007	−0.003	0.006
	(0.028)	(0.002)	(0.005)	(0.008)	(0.009)	(0.004)
Home-Attendance	0.200***	0.005***	0.029***	−0.009**	0.020***	0.024***
	(0.015)	(0.001)	(0.003)	(0.004)	(0.005)	(0.002)
Δ Elo/1000	0.742***	0.021***	0.127***	0.010	0.137***	0.102***
	(0.055)	(0.004)	(0.011)	(0.014)	(0.016)	(0.009)
Age/100	19.928***	0.539	3.226***	2.541**	5.767***	3.102***
	(5.313)	(0.431)	(0.932)	(1.144)	(1.348)	(0.798)
$(\text{Age}/100)^2$	−40.041***	−0.578	−5.446***	−5.315**	−10.761***	−5.381***
	(9.564)	(0.776)	(1.660)	(2.060)	(2.411)	(1.444)
Maximum at age	24.9***	46.5***	29.6***	23.9***	26.8***	28.8***
	(1.0)	(25.8)	(1.2)	(1.9)	(0.8)	(1.0)

Based on 32,770 observations of 500 players with at least 25 observations; playing time in minutes/90; standard errors clustered at the player level in parentheses; standard errors of maximum age calculated using the Delta-method; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B4: Productivity by Position

Goal scoring is dependent on the position of the player; more likely for players in attacking position than for players in midfield or defense positions. Table B.4 shows the way positions are coded into three types.

Table B.4: Position codes and grouping

Code	Description	Position group
GK	Goalkeeper	Goalkeeper
DL / DC / DR	Left-back / centre-back / right-back	Defence
DML / DMC / DMR	Defensive midfielder left / centre / right	Midfield
ML / MC / MR	Midfielder left / centre / right	Midfield
AML / AMC / AMR	Attacking midfielder left / centre / right (incl. wingers in a 4-3-3 or 4-2-3-1)	Midfield
FWL / FW / FWR	Forward left / centre / right	Attack

Source: Understat.com. Codes refer to the formation slot in which a player started a match, as recorded by Understat (L = Left, C = Center, R = Right). Substitutes are assigned the slot they most frequently started in during the same season, or in other seasons if they did not start that season.

Table B.5: **Parameter Estimates by Player Position**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
a. Defense						
Playing time	0.424***	−0.019*	0.008	0.031	0.039*	0.093***
	(0.058)	(0.010)	(0.013)	(0.021)	(0.023)	(0.015)
Home-BCD	0.039	0.001	0.004	0.009	0.013	−0.006
	(0.033)	(0.006)	(0.008)	(0.012)	(0.015)	(0.006)
Home-Attendance	0.115***	0.003	0.011**	−0.005	0.006	0.018***
	(0.017)	(0.004)	(0.005)	(0.008)	(0.009)	(0.004)
Δ Elo/1000	0.315***	0.003	0.026**	0.018	0.045*	0.046***
	(0.059)	(0.010)	(0.013)	(0.026)	(0.025)	(0.016)
Age/100	9.792	0.821	1.867*	3.048**	4.915***	1.350
	(7.190)	(0.817)	(1.087)	(1.364)	(1.668)	(1.587)
$(\text{Age}/100)^2$	−16.125	−1.201	−2.884	−5.245**	−8.129***	−2.465
	(12.353)	(1.439)	(1.967)	(2.424)	(3.058)	(2.759)
Maximum at age	30.4***	34.1***	32.4***	29.1***	30.2***	27.4***
	(2.1)	(8.2)	(4.2)	(2.1)	(1.8)	(3.3)
b. Midfield						
Playing time	1.155***	−0.012***	0.096***	0.010	0.106***	0.146***
	(0.045)	(0.004)	(0.008)	(0.011)	(0.012)	(0.008)
Home-BCD	0.034	−0.006**	−0.003	−0.007	−0.010	0.004
	(0.032)	(0.003)	(0.006)	(0.011)	(0.012)	(0.006)
Home-Attendance	0.174***	0.006***	0.022***	−0.008	0.015**	0.026***
	(0.019)	(0.002)	(0.004)	(0.006)	(0.007)	(0.003)
Δ Elo/1000	0.768***	0.024***	0.113***	−0.001	0.112***	0.137***
	(0.065)	(0.006)	(0.012)	(0.016)	(0.019)	(0.013)
Age/100	10.472**	0.804	2.408**	3.242**	5.651***	2.203**
	(5.231)	(0.524)	(0.933)	(1.467)	(1.510)	(1.076)
$(\text{Age}/100)^2$	−21.809**	−1.120	−3.841**	−6.597**	−10.438***	−3.219
	(9.399)	(0.985)	(1.708)	(2.701)	(2.772)	(1.954)
Maximum at age	24.0***	35.9***	31.4***	24.6***	27.1***	34.2***
	(2.1)	(8.9)	(2.6)	(1.6)	(0.9)	(4.6)
c. Attack						
Playing time	1.682***	0.003	0.254***	−0.017	0.237***	0.124***
	(0.051)	(0.006)	(0.012)	(0.016)	(0.019)	(0.008)
Home-BCD	0.122**	−0.005	0.010	−0.020	−0.010	0.014
	(0.056)	(0.005)	(0.014)	(0.021)	(0.023)	(0.010)
Home-Attendance	0.305***	0.005**	0.054***	−0.013	0.041***	0.024***
	(0.033)	(0.003)	(0.007)	(0.010)	(0.011)	(0.005)
Δ Elo/1000	1.203***	0.026***	0.233***	0.022	0.255***	0.092***
	(0.106)	(0.008)	(0.025)	(0.035)	(0.039)	(0.016)
Age/100	1.302	0.923	2.053	0.025	2.077	1.391
	(10.728)	(1.120)	(2.938)	(2.907)	(3.675)	(1.325)
$(\text{Age}/100)^2$	−5.965	−1.239	−3.213	−1.872	−5.085	−2.281
	(19.571)	(1.903)	(5.230)	(5.250)	(6.519)	(2.352)
Maximum at age	10.9	37.3***	31.9***	0.7	20.4**	30.4***
	(54.5)	(13.0)	(8.4)	(75.8)	(10.7)	(5.2)

Players with at least 25 observations overall; panel a 6,422 observations of 183 players; panel b 16,796 observations of 395 players; panel c 9,552 observations of 252 players; standard errors clustered at the player level in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B5: Including Players with Zero Shots

Table B.6: **Baseline Parameter Estimates; Including Players with Zero Shots**

	Shots (1)	xG/shot (2)	xG (3)	G-xG (4)	G (5)	xA (6)
Playing time	1.264***	−0.007**	0.140***	−0.010***	0.130***	0.104***
	(0.033)	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)
Home-BCD	0.049***	−0.004*	0.005*	−0.005	−0.000	0.000
	(0.016)	(0.002)	(0.003)	(0.004)	(0.004)	(0.002)
Home-Attendance	0.173***	0.005***	0.021***	−0.005***	0.016***	0.016***
	(0.009)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Δ Elo/1000	0.784***	0.019***	0.100***	0.000	0.100***	0.078***
	(0.034)	(0.004)	(0.006)	(0.006)	(0.008)	(0.005)
Age/100	8.250**	0.631	1.569***	0.947*	2.516***	1.675***
	(3.593)	(0.416)	(0.528)	(0.509)	(0.706)	(0.435)
$(\text{Age}/100)^2$	−15.734**	−0.765	−2.535***	−1.982**	−4.517***	−2.721***
	(6.264)	(0.751)	(0.909)	(0.896)	(1.252)	(0.761)
Maximum at age	26.2***	41.2	30.3***	23.9***	27.8***	30.8***
	(1.5)	(13.9)	(1.5)	(2.5)	(1.0)	(1.2)

Based on 80,050 observations of 928 players with at least 25 observations; xG/shot based on 38,276 observations of 880 players; playing time in minutes/90; standard errors clustered at the player level in parentheses; standard errors of maximum age calculated using the Delta-method; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$