Working Papers in Management
Birkbeck, Department of Management

BWPMA 1501

Relative Age Effect on Labor Market Outcomes for High Skilled Workers – Evidence from Soccer

Luca Fumarco
Department of Economics
Linnaeus University

Giambattista Rossi
Department of Management
Birkbeck, University of London

March 2015

ISSN 2052 – 0581

Birkbeck University of London, Malet Street, London, WC1E 7HX

Relative Age Effect on Labor Market Outcomes for High Skilled Workers – Evidence from Soccer

by

Luca Fumarco[†] and Giambattista Rossi[*]

Abstract. In sport and education contexts, children are divided into age-groups which are arbitrary constructions based on the admission dates. This age-group system is thought to determine differences in maturity between pupils within the same group, that is, relative age (RA). In turn, these within-age-group maturity differences produce performance gaps, that is, relative age effects (RAE), which might persist and affect the labor market outcome. I analyze the RAE on labor market outcomes using a unique dataset providing information on a particular group of high skilled workers: soccer players in the Italian major soccer league. In line with previous studies, evidence on the existence of RAE in terms of representativeness is found, meaning that players born relatively early in the age-group are over-represented, while players born relatively late are under-represented, even accounting for specific population trends. Moreover, players born relatively late in the age-group receive lower gross wages than players born relatively early. This wage gap seems to increase with age and in the quantile of the wage distribution.

JEL-Classification: J24, J31, J71, L83, M53
Keywords: Relative age, labor markets in sports

[†] Linnaeus University, e-mail: luca.furmaco@lnu.se
* Birkbeck University of London, email: g.rossi@bbk.ac.uk
Acknowledgements. I would like to thank: my supervisors Dan-Olof Rooth and Magnus Carlsson for guidance and many useful suggestions; Simone Scarpa, Peter Karlsson, Håkan Locking and PhD fellows from the Department of Economics and Statistics of the Linnaeus University, for valuable feedback; the participants to the AIEL conference, 2014, in Pisa; the participants to the SWEGPEC PhD workshop, 2014, in Jönköping.

1 Introduction

There is large empirical evidence that children born late in the education and sport admission year are systematically disadvantaged throughout childhood up to the late teens. Scholars from different disciplines justify this evidence with the existence of the so called relative age effect. This concept has recently gained popularity even outside academy (e.g., Gladwell, 2008; Levitt & Dubner, 2010).

The relative age effect is given by similar complex mechanisms in education and sports. In both contexts, age-groups are formed using arbitrary admission dates which determine some children to be older than others within the same age-group. This chronological difference, called relative age (henceforth RA), is responsible for early differences in maturity (e.g., Bedard & Dhuey, 2006; Musch & Hay, 1999), which cause a performance gap, that is, the relative age effect (henceforth RAE),[1] and affect children's achievements. Because of its nature, this effect is expected to dissipate with age and eventually to disappear. However, it might persist, and even widen, because of some characteristics of the human capital accumulation process, which lead to "path dependence" (Bryson et al., 2014, p.12). It means that children born early in the admission year are more likely to be perceived as talented (e.g., Allen & Barnsley, 1993), and thus they are given more chances to develop their skills (e.g., teachers and parents motivate them more, or children could be provided with superior educational quality).

Although there is large consensus about the negative RAE on relatively young children's achievements, no equivalent consensus exists on the RAE on labor market

[1] Consider the case where all children who turn 6 in a given calendar year are expected to start the first grade of primary school in that year (i.e., the admission date is the 1st of January; note that the beginning of the school year is irrelevant). In the same class, there might be children who turn 6 in January and children who turn 6 in December; relatively old pupils born in January are 17% older than relatively young pupils born in December. This chronological difference is the RA, which causes differences in terms of maturity, leading to a performance gap; this performance difference is the RAE.

outcomes (Ponzo & Scoppa, 2014). Whether there is such a long-run effect is a compelling economic question (e.g., Allen & Barnsley, 1993; Bedard & Dhuey, 2006).

One of the possible reasons for the lack of consensus is the presence of two important confounders which affects scholars' analyses: "season-of-birth effects" and heterogeneous ages within age-groups. The "season-of-birth effects" are confounding factors because they are unrelated to within-age-group maturity differences, and are due to climatic, environmental, sociocultural and biological factors (Musch & Grondin, 2001). The season of birth explains the performance gaps between children born in the *same calendar year* with the *position of their birthdates within the calendar year*; whereas the RAE explains the performance gaps between children born in the *same admission year* with the *maturity gap caused by the relative position of their birthdates within the admission year*. On one hand, when the beginning of the admission year coincides with a period of the calendar year which conveys advantages due to seasonal effects, the estimate of the RAE might likely be upward biased.[2] On the other hand, the estimate could be downward biased if later months of the selection year coincide with a period of the calendar year which conveys advantages to children born within that period.[3, 4] Also the presence of heterogeneous ages within age-groups may bias the results from analyses on RAE. Consider the education context, when children born late in the admission year, that is, relatively young children, are held back one year, that is, they either repeat a grade or they enter primary school one year later, they end up in an age-group where typical children are younger, and, thus, they become relatively old

[2] Consider the case when the school admission year coincides with that of sport, and the researcher was interested only in the RAE from either education or sport, not in their combined effect. The estimates would be biased (Musch & Hay, 1999; Helsen et al., 2012).

[3] If the admission date was shifted by a few months, e.g., shift the admission date in the example in Footnote 2 by 6 months, the estimated RAE would be downward biased. Also, if households with high socio-economic status tend to give birth in months that do not coincide with the beginning of the admission year, as in the US (Bound & Jaeger, 2001) and in Sweden (Carlsson et al., forthcoming), the estimate of RAE from education would be downward biased.

[4] The seminal paper by Angrist and Kruger (1992) might be interpreted as a particular case where the estimate of the RAE could be downward biased because of the school leaving age. The authors find that pupils born at the beginning of the admission year attain less schooling than their younger peers, since they are legally allowed to leave school before to graduate.

children in their new age-group (Bedard & Duhey, 2006). In this case the estimate of the RAE might be downward biased.[5] Moreover, as Bedard and Duhey (2006) suggest, in countries where pre-school institutes are not free, the possibility of redshirting, that is, enter primary school one year later, might affect the estimates of the RAE also via the socio-economic status. In countries as the US, high socio-economic status parents are more likely to afford one extra year of pre-school. In this case the estimate of the RAE might likely be even more downward biased.

The goal of this study is threefold. First, this paper adds to the existing economic literature by investigating different aspects of the RAE on labor market outcomes, including the long-run RAE. The focus is on a particular group of high skilled workers: professional soccer players from the Italian major league, that is, Serie A. Second, this paper aims at providing a descriptive general framework to the RAE, bringing to the attention of the reader articles from different disciplines. The literature review in this article stresses the importance of different mechanisms and of different evidence on RAE sometimes neglected in economic studies. Third, this article proposes the use of the quantile regression to gain more insights on the long-run economic RAE.

What is the reason for analyzing soccer players? The first reason for studying this particular group of workers is that season-of-birth effects seem to play a minor role in the soccer domain. There is evidence that seasonal effects have only an attenuate—if not null—effect on the mechanisms leading to RAE in professional soccer. Munch and Hay (1999) explain that, at the end of the 80s, in the major soccer leagues of Germany, Brazil, Australia and Japan, soccer players born early in the admission year were consistently over-represented. This result is consistent with the RAE: throughout the years of sports activity more early-born

[5] In a similar manner the RAE in sport might be nullified. As documented in Parent-Harvey et al. (2013) and Böheim and Lackner (2012), when the selection of the athletes into professional competitions is based on a draft system, relatively young athletes might delay the entry into professional sport by one year to overcome developmental differences.

soccer players were considered more talented and thus reached the top leagues. This result is obtained despite a number of differences between these four countries: admission dates, reversed seasons, typical climate, biological characteristics, and socio-cultural factors. Also studies on the effect of a shift in the admission date provide results consistent with the RAE, ruling out alternative explanations. Munch and Hay (1999) show that a shift of the Australian admission date by a few months led to a corresponding reduction in the players' birthrate for the early months under the previous admission date. Helsen et al. (2000) study the effect of a similar shift in the admission date occurred in Belgium, and find a corresponding adjustment. Seasonal effects may also hardly offer the explanation to performance gaps between players born in two adjacent months, where one month is before and one after the admission date (e.g., Barnsley & Thompson, 1988; Ponzo & Scoppa, 2014).

A second reason is that the presence of age-groups with heterogeneous ages is limited in soccer. In Italy, which is the context of this analysis, the age-group system for soccer is very strict, so that the bias given by heterogeneous ages within age-groups should be less of an issue.[6] Moreover, related to this point, the effect of the household socio-economic status via redshirting is avoided a priori, since redshirting is not possible; also, there are reasons to believe that it would not matter anyway. For what reason would someone assume that only households with high socio-economic status can afford to have their children starting to play soccer later? In conclusion, no particular identification strategy has to be adopted to address the bias caused by age-groups with heterogeneous ages.[7]

A third reason to study the RAE in the soccer players' labor market is the quality of the available data. As Kahn (2000) writes, within the sport field data are very detailed. For

[6] According to rules set by the Italian Football Federation (FIGC), only in the last juvenile category a team may deploy one overage player in regular matches, and only in this last category as well as in one intermediate category a team may deploy underage players.

[7] For instance, because of the possibility to postpone or anticipate the entry into school, Bedard and Duhey (2006) as well as Ponzo and Scoppa (2014) adopt an instrumental variable estimation strategy, where they instrument the students' actual age with the so called *expected age*, that is, the age children should have at the moment their performance is measured based on both their month of birth and the admission month used in the schooling system.

instance, data on employees' performances and compensations are accessible, data on employees can be easily matched to those of their employers throughout the career and can be more accurate and detailed than usual microdata (Kahn, 2000).

Based on the previous literature, the first hypothesis tested in this paper is the presence of RAE in terms of representativeness. In presence of RAE the observed amount of Italian players born at the beginning of the admission year should be larger than the expected amount, based on the birthrate of the general population; the contrary should be true for those players born at the end of the admission year.[8] The RAE mechanism suggests that relatively old players are often perceived as talented in early ages, they are (more or less formally) streamed (Allen & Barnsley, 1993), and reach Serie A more frequently than relatively young peers.

The results provide evidence for the existence of RAE in terms of representativeness in Serie A among Italian players. Moreover, an additional analysis suggests the presence of a specific trend explainable with the RAE: the over-representation decreases and turns into under-representation as the end of the admission year is approached.

The RAE in terms of wage gaps is also analyzed. The RA framework suggests three different possible results. Traditionally, the RA suggests that on average relatively old players should perform better (Allen & Barnsley, 1993) and thus should receive larger wages, as they have had a relative advantage throughout the pre-labor market period. The opposite result is illustrated in Ashworth and Heyndels (2007), Gibbs et al. (2012), and Bryson et al. (2014). Positive selection and peer effects could positively affect relatively young players' performances and lead to higher wages. The best relatively young children manage to overcome the difficulties and eventually benefit from learning and training with stronger

[8] Moreover, in presence of RAE, players born in January would be over-represented in the sample, players born in March would still be over-represented but to a lower extent, players born in October would be under-represented, players born in December would be the most under-represented. This would be true even when trends in the general population birthrate are accounted for.

peers.[9] Differently, recent studies suggest that the performance gap disappears in the labor market, as the discriminatory streaming criteria which affect per-labor market achievements cease to be relevant (Crawford et al., 2013). To the best of my knowledge only three other studies investigate RAE on wages for high skilled workers: Kniffin and Hanks (2013) for PhD students, Böheim and Lackner (2012) for American football players, along with Ashworth and Heyndels (2007) for German soccer players.

The main set of results provide statistically significant evidence that relatively young players earn lower wages, supporting the theory according to which the RAE negatively affects the performances also in the long-run (Allen & Barnsley, 1993). Additional analyses suggest that this wage gap might be the largest at the entry of the labor market, while in the remainder of the career the wage gap is smaller and tends to increase toward the end of the career. This particular development of the wage gap could be due to players' career choices. As a further contribution to the economic literature, this paper analyzes whether the RAE on wages differs by wage quantile. To the best of my knowledge, none of the existing studies analyzes the wage gap with a quantile regression. This analysis is important when investigating a labor market characterized by a strongly positively skewed wage distribution and when the researchers hypothesize the existence of peer effects or positive selection. The results point to the possibility that the wage gap could increase in the quantile of the wage distribution; in turn, this result could imply the absence of positive peer effects and selection for relatively young players.

The remainder of the paper proceeds as follows. Section II presents a summary of the literature review on RAE in education and sport; Section III discusses the data and presents descriptive statistics; Section IV present the empirical methodology; Section V illustrate the results; Section V concludes.

[9] In alternative, Williams (2010) hypothesizes that in the long-run relatively young players might outperform their relatively older peers, because relatively young athletes have a complete training, while their relatively older peers put less emphasis on skills development, as they are primarily selected upon their physical attributes.

2 The Relative Age Effect: Mechanisms and Evidence

2.1 Mechanisms

RAEs in education and sport contexts appear to be similar in their mechanisms and consequences on people achievements. The similarities between these two contexts are emphasized when a competitive streaming process takes place.

In education, RAE is initially caused by differences in children's cognitive development. These differences trigger misjudgments on pupils' talent and, eventually, more or less flexible streaming (Bedard & Dhuey, 2006). In case of formal streaming, some children are assigned to vocational schools and others to academic schools, or they are divided between ability-based reading groups (Bedard & Dhuey, 2006). When there is no formal streaming, social interactions between children, parents, and educators play a prominent role (Hancock et al., 2013), since stronger students are encouraged to progress while weaker students are allowed to lag behind (Bedard & Dhuey, 2006). An example of social interaction effect is the Pygmalion effect, which predicts that teachers', trainers', and parents' expectations on children's ability trigger self-fulfilling prophecies (Musch & Grondin, 2001; Hancock et al., 2013). Another example is the Galatea effect, which predicts that children's expectations on themselves trigger self-fulfilling prophecies (Hancock et al., 2013).

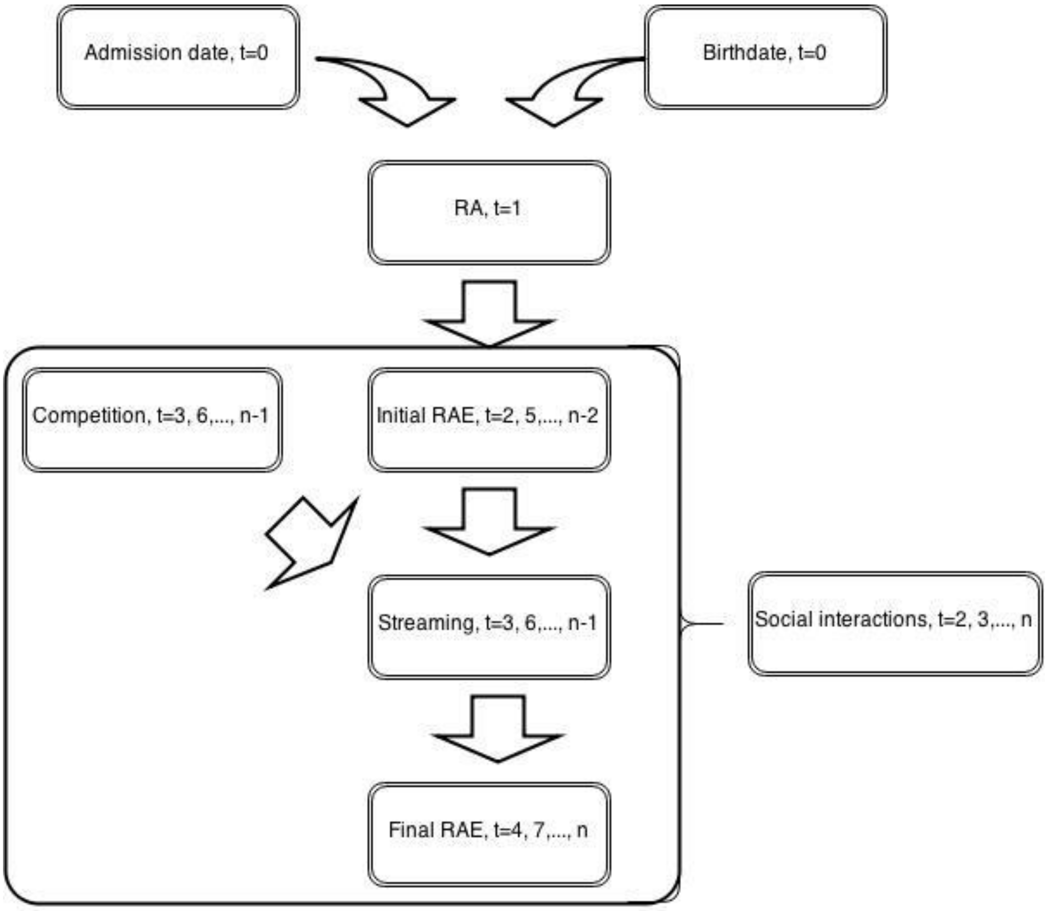
RAE in sport differs from that in education with respect to at least three aspects. First, RAE in sport is caused by initial differences in children's cognitive *and physical* development (Allen & Barnsley, 1993), conveying an additional edge to relatively old children. Second, competition might be tougher from the early stages of youth sports. The competition level is determined by a number of factors, such as the amount of teams within a county, the amount of available spots per team, and the amount of children who eventually can compete regularly (Allen & Barnsley, 1993; Musch & Grondin, 2001). Let's consider the general case in soccer, where the amount of teams per county and the amounts of available spots per team are not binding, and kids enter a team simply by buying a subscription. Only a limited amount of children per team eventually gets to play often in regular matches; the children in the starting team plus the substitutes, who actually enter the pitch, will cumulate experience and skills more rapidly. Since relatively old children are more mature, they perform better, and improve more rapidly. In case the amount of teams and the amount of available spots per team were binding, competition might be fiercer, and teams could select children based on their perceived talent, increasing the effect of competition; for instance, in national youth summer camps (Glamser & Vincent, 2004), or youth national teams (Williams, 2010).[10] Third, in sport children may drop out (Barnsley & Thompson, 1988; Helsen et al., 1998).[11] While school is compulsory in early ages and it is possible to drop out only during late years of high school or during university, sport is based on voluntary participation (Musch & Hay, 1999; Musch & Grondin, 2001).

[10] To the best of my knowledge, only one study investigates the RAE in soccer academies (Carling et. al, 2009). The authors find that the relative age effect might not always determine significant performance gaps. This study analyzes only physical components of young players' performances, however.

[11] It does not mean that children drop sport activities in general; they could simply change sport, opting for one in which the admission date either has lower or no importance (Williams, 2010), or provides them with a positive RAE (Thompson et al., 1999) contributing to the RAE in that sport.

Overall, the RAE mechanism found in education and sport might be summarized by Figure 1.[12]

Figure 1



In t=0 we have a given admission date and a given birthdate, they cause the RA in t=1. The RA in t=1 creates the initial RAE in t=2, then there is a (more or less formal) streaming process which is affected by competition, and generates the final RAE. After the final RAE is created, the cycle begins all over again with a new initial RAE. In all periods from t=2 on, the initial RAE, the streaming process, the final RAE and competition affect, and are affected by, social interactions. Note that initial RAE and final RAE might differ because of the social interactions, even in absence of formal streaming process and competition.

[12] This is an original flowchart which I have produced based on the theories illustrated by articles from different disciplines.

The mechanism which leads professional athletes into the labor market is similar to the mechanism which leads high skilled workers into the labor market. Although initially they differ somewhat, in the last stages they share a number of characteristics: in both education and sports there is more or less formal streaming, the participation to the training / education is based on voluntary participation (e.g., in the last stages of secondary education and in the whole tertiary education) and there is high competition (e.g., in education there is competition for scholarships and for spots in programs with limited amount of seats).

2.2 Evidence from Previous Literature

The short-run evidence on RAE from education and sports reconciles. In education, for example, late-born children are more likely to be retained for an additional year in the same grade or to be assigned to remedial classes (Dixon, Horton, & Weir, 2011); they are more likely to be diagnosed with learning disability (Dhuey & Lipscomb, 2009); they are also more likely to be diagnosed with attention-deficit/hyperactivity disorder and be prescribed ad hoc stimulants (Zoëga et al., 2012); they are characterized by lower performances (Plug, 2001; Bedard & Dhuey, 2006; Ponzo & Scoppa, 2014),[13] and they have a lower school attendance rate (Cobley et al., 2009). The sport context differs in terms of the type of evidence provided for the existence of the RAE. While in education RAE is prevalently measured in terms of actual performances, in sport it is measured in terms of representativeness. In fact, because of the tougher competition and the possibility to drop out, in each age-group early born athletes are over-represented and late born athletes are under-represented with respect to the general population. This result is similar to that from the education context: since the best performers keep on practicing sport (e.g., they do not drop out or are selected into higher tiers) and a

[13] Two articles find opposite results. Fredriksson and Öckert (2005) find that absolute age when starting school, in lieu of relative age, is responsible for different school performances. Cascio and Schanzenbach (2007) find the positive peer effects benefit relatively young pupils.

larger percentage of these performers is born early in the admission year, it follows that relatively old children should on average outperform relatively young children.

Conclusions on the long-run RAE are ambiguous in both education and sports. In university, the RAE might turn in favor of relatively young students in terms of academic performances, although at the cost of lower social skills (Pellizzari & Billari, 2012).[14] Differently, relatively young students seem to earn the Ph.D. at the same age of relatively old students and seem to earn the same salary in postdoc positions (Kniffin & Hanks, 2013). On the general labor market, some other studies provide evidence for null RAE in terms of wages. Perhaps, different performances reflect only chronological age differences (Larsen & Solli, 2012), so that overall there is a null RAE on life earnings. There might even be no wage gap at all, if employers reward employees' productivity irrespectively of their educational achievements, biased in favor of relatively old students (Crawford et al., 2013). Du et al. (2012) find instead a negative RAE, in terms of representativeness in the labor market; they study a sample with the CEOs from the 500 S&P firms, and find that relatively old CEOs are over-represented. Muller-Daumann and Page (2014) find an equivalent result among US congressmen. Finally, Black et al. (2011) and Plug (2001) find a wage gap in favor of relatively old workers. In sport, Ashworth and Heyndels (2007) find reverse RAE in terms of wages, with relatively young athletes receiving higher wages, and RAE in terms of representativeness, with relatively old athletes being over-represented. Also reverse RAE in terms of representativeness among the very best hockey and soccer players has been found,[15] with relatively young players being over-represented (Gibbs et al., 2011; Bryson et al. 2014). Whereas usually, over-representation of relatively old players is found among other professional athletes; for example, in soccer (e.g., Musch & Hay, 1999) and in tennis (Edgar

[14] Examples of lower social skills are leadership skills (Dhuey & Lipscomb, 2008), self-esteem (Thompson et al., 2004), and poorer social lives (Pellizzari & Billari, 2012).
[15] Players selected for the all-stars and for Olympic team rosters in hockey, and teams captains in soccer.

& O'Donoghue, 2005), in both summer and winter Olympic games (Joyner, et al., 2013), and in NFL (Böheim & Lackner, 2012).

Concluding, on one hand, the literature shows that in both contexts the short-run RAEs on children's achievements are qualitatively similar. This comes as no surprise since the RAE is generated through similar mechanisms in sports and education. On the other hand, the evidence on long-run RAE is mixed in both contexts.

3 Institutional Context and Data

The empirical setting of our analysis is the Italian soccer major league, called Serie A. It is currently composed by 20 teams, but these teams do not permanently play in the major league; the Italian soccer has a tiered structure, with promotions and relegations at the end of each season. The last three teams in the ranking are relegated to the second national division, that is, Serie B, which is composed by 22 teams; the top three teams from this second league are promoted to Serie A.

In Italy, the age-group system for soccer is strictly regulated. The 1st of January is the relevant admission date applied to each age-group, although specific age-groups have slightly different rules. There are seven age-groups in youth competitions; some of them are one-year age-groups, while others are two-year age-groups. In the latter case, children with different ages might play in separate games, despite training together, if the rules specify so. In general, children have to train and play in the assigned age-group.[16] The minimum age requirement to

[16] The lowest age category is for children from 5 to 7 years; they are put in the same age-group, called "Piccoli Amici" (i.e., Small Friends), for both training and competing. In the next two categories, children of different ages are still grouped together for training, but they are divided based on year of birth for competing. These categories are "Pulcini" (i.e., Chicks), for children under 11 years of age, and "Esordienti" (i.e., Newcomers), for children under 13 years of age. Up to three underage players may play in "Esordienti" matches. In the next categories, teenagers with different ages are put together for both training and competing. These categories are "Giovanissimi" (i.e., Very Young), for players under 15 years of age; "Allievi" (i.e., Cadets), for players under 17 years of age; and finally "Primavera" (i.e., Spring), for players between 15 and 20 years of age. In all these categories excluded the last one, no overage is allowed; in "Primavera" only one overage player per team may participate to the matches. The rules do not seem to set restrictions on whether children are free to train in an age-group different from the one to which they are assigned, and eventually to play official games in their assigned age-group. More information on the Italian players age-grouping system can be found on the official web-site of the Italian Football Game Federation (FIGC).

play for a professional soccer team is 14 (art. 33, Internal Organization Rules FIGC);[17] however, it is possible to sign a contract with a team in a professional league only from 16 years of age (art. 33, Internal Organization Rules FIGC).

The dataset contains information on players from seven Serie A seasons, 2007-08 to 2013-14. There are observations on 508 Italian soccer players, who played for at least one Serie A team over the seven seasons in analysis.[18] In total, the unbalanced panel data contains 1,704 Italian soccer-season observations. Most soccer players appear in our dataset for one or two seasons, 139 and 100 players respectively; 56 and 48 players are present for 3 and 4 seasons respectively; 53 and 45 players are present for 5 and 6 seasons respectively; 45 players are present in all the 7 seasons. Players may leave the dataset either permanently or temporarily: some players play in teams which are eventually relegated and may or may not be re-promoted to Serie A or sold / lent to a Serie A team; some players may be sold / lent to foreign teams, or to teams in lower leagues, and may or may not be transferred back to Serie A teams; some players may retire.[19]

The empirical analyses use information on players' wage, age, quarter and month of birth, current team, soccer season and role on the pitch. Variables description and descriptive statistics are presented in the Appendix A, respectively in Table A.1 and Table A.2.

Figure 2 illustrates the histogram for Italian players' birthrate per quarter.[20] The division of the admission year into quartiles is a convention adopted within the relative age research (Wattie at al., 2015). The black rhombuses represent the average birthrate per quarter

[17] Norme Organizzative Interne Della FIGC.
[18] The focus is on Italian players so to analyze a set of players who have trained under the same admission date. Moreover, the research of admission dates from other countries is a complex task: admission dates might differ between countries and within countries in different youth categories and different regions or states.
[19] The impact of all these players' movements on the estimates of RAE is not clear. Good Serie A players might move abroad attracted by better contractual conditions and / or greater visibility. However, this is true also for worse performers who may just want to play more often. This point is noted also by Ashworth and Heyndels (2007) who suggest that the amount of foreign players in the domestic league might affect the estimate of RAE, through increased competition.
[20] Where 1 is the quarter January-March, 2 is the quarter April-June, 3 is the quarter July-September, and 4 is the quarter October-December.

in the Italian population, between 1993 and 1998; Italians' birthrates for previous years are unavailable. Appendix B reports the amount of births per month, per year, in the general population. This figure suggests the presence of RAE in terms of representativeness, that is, the relatively young players born toward the end of the admission year are under-represented, while relatively old players are over-represented. Moreover, there seems to be a specific trend: Serie A players' birthrate decreases with the distance from the admission date.

Figure 2

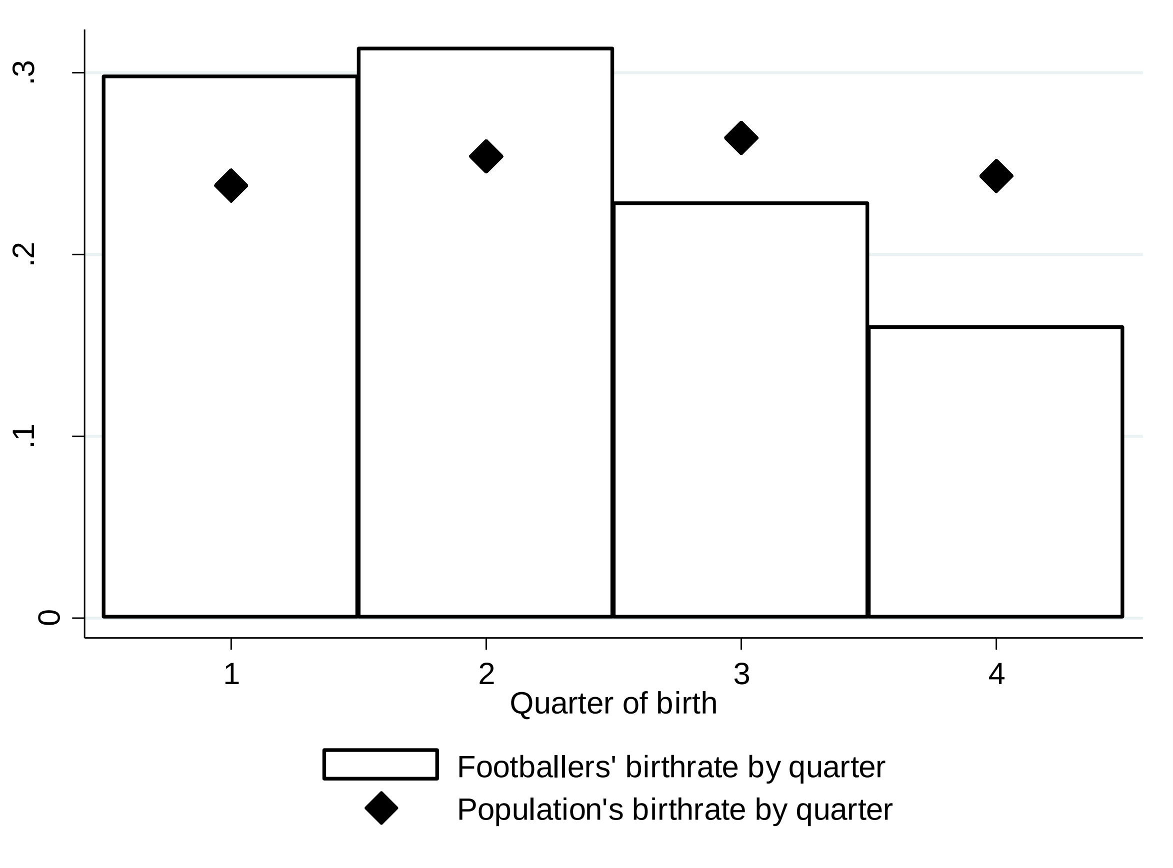
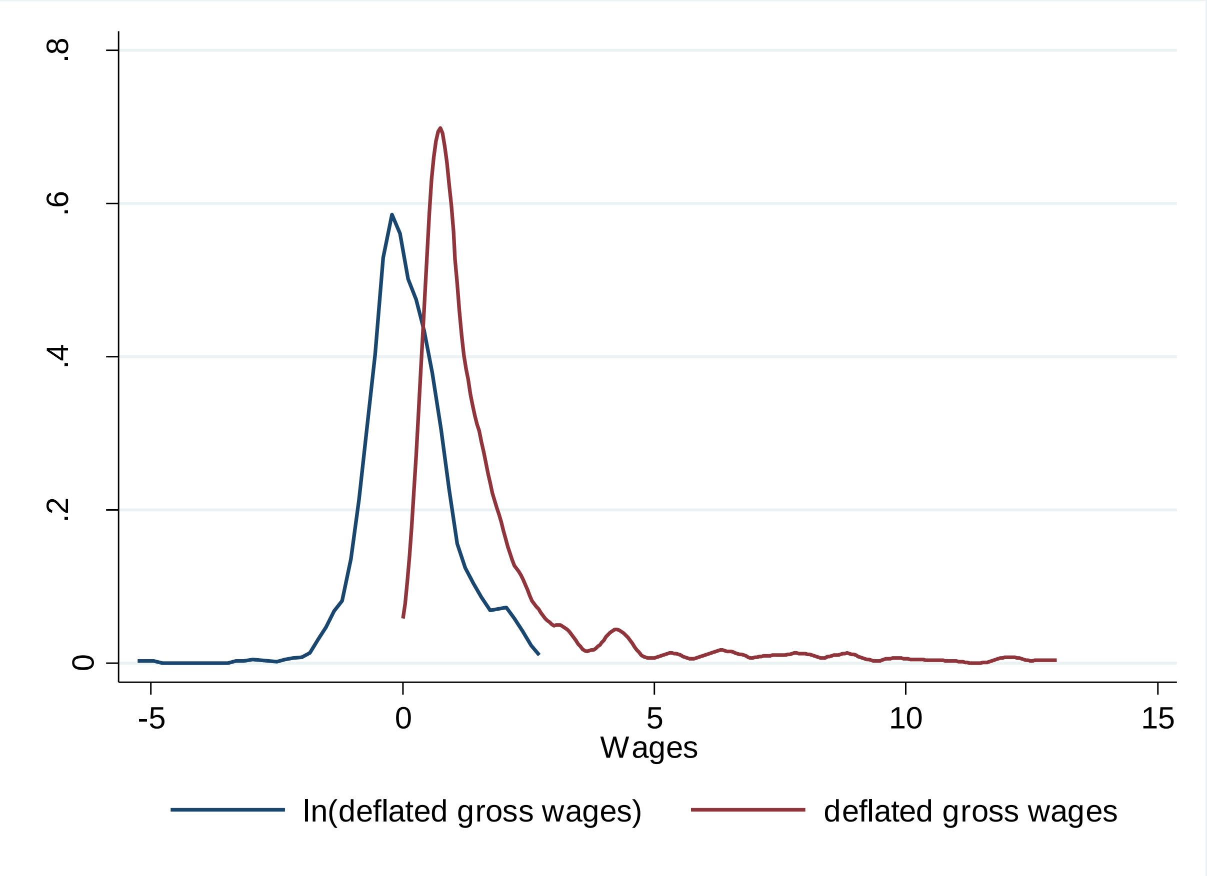


Figure 3 illustrates the players' wage distribution.[21] They are measured before taxation—without either bonuses, image rights or other deals—and they are deflated at the 2013 price level, the annual coefficients are provided by Italian National Institute for Statistics (ISTAT).

[21] The information on wages is obtained from annual reports completed by the Italian sport dedicated newspapers.

Figure 3



The original distribution of the gross wages is highly unequal, with a substantial positive skewedness, as it is expected in labor markets characterized by the presence of superstars (Lucifora & Simmons, 2003).[22] The transformation of gross wages into natural logarithm returns a somewhat fairly normal distribution.

First insights on possible wage gaps can be obtained comparing the distribution of the gross wages for relatively old and young players. Figure 4 compares the kernel density distributions of the gross wages for Italian players, divided by quarter of birth. Since players

[22] Superstar is the term used to refer to extreme wage outliers (e.g., Bryson, Rossi & Simmons, 2014; Kleven et al., 2013; Lucifora & Simmons, 2003; Adler, 1985; Rosen, 1981). These outliers are such that in a labor market there appears to be a convex relationship between wage and skills (Lucifora & Simmons, 2003). The main competing, yet not mutually exclusive, superstar theories are two: Rosen (1981) suggests that superstars enjoy huge salaries because of scarcity of talent, so that little additional talent translates into large additional earnings, whereas Adler (1985) suggests that huge salaries are caused by positive network externalities, which creates additional popularity, even in absence of talent.

can change team during the season, they are assigned the gross wage they receive from the team with which they start the season.[23]

Figure 4

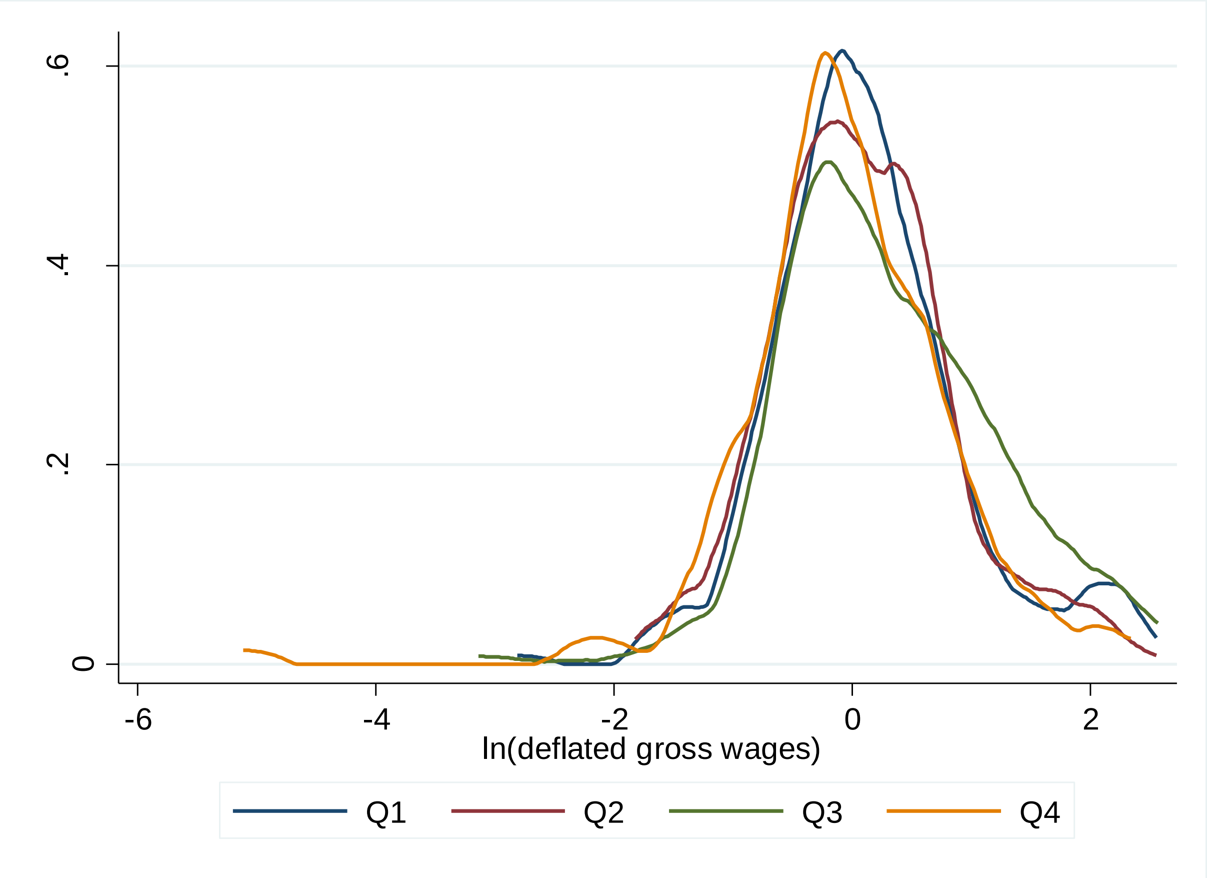
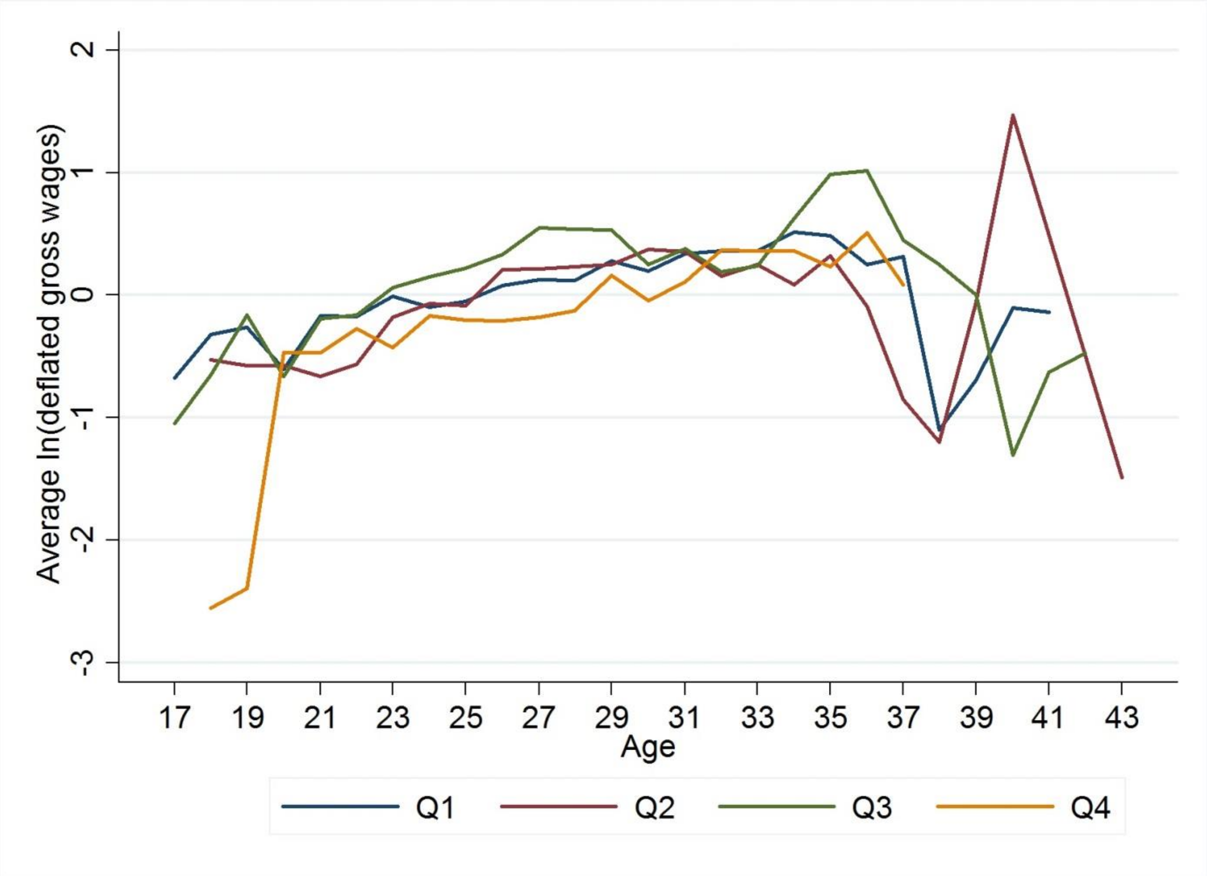


Figure 4 shows that the wage distribution of players born in the last quarter of the year (yellow line, Q4) has a longer and thicker left tail than the distributions for players born in other quarters. This suggests there might be a wage gap penalizing players born in the last quarter. Somehow puzzling, players born in the third quarter (green line, Q3) tend to earn more frequently wages in the top percentiles of the wage distribution.

[23] Teams trade players in two main market sessions: in summer, which separates different soccer seasons, and during the Christmas break and January, which is toward the end of the first half of the season. Players who change team during the latter session and come from another Serie A team are assigned the gross wage they received from the team with which they started the season. Players who join a Serie A team during the Christmas break and come either from abroad or from Serie B—or other lower domestic leagues—are assigned the new wage.

Additional insights on the nature of the wage gap might be gained with the investigation of its dynamics through players' career. With this purpose, Figure 5 plots the average natural logarithm of the gross wages against age for the four groups of players.

Figure 5



For players born in the last quarter of the year (yellow line, Q4), there is an important entry wage gap which disappears in the early twenties; afterwards, a wage gap appears anew and disappears only toward the end of the career. The entry wage gap is smaller for players born in the third quarter (green line Q3);[24] these players also seem to enjoy higher wages in the core of the career and toward its end. Around 40 years of age a gap in favor of players born in the first quarter (red line, Q1) appears. The indications provided by this figure should be

[24] The entry wage gap might be explained with a physical development gap, which disappear at 20 years of age (see the WHO grow charts for children from 5 to 19 years of age, in the WHO web-site, and US grow charts, in Kuczmarski et al, 2002), while the cognitive development gap disappear between 20 and 25 years of age (Salthouse at al., 2004).

considered carefully. This graph illustrates averages and there is a low amount of observations for extreme ages,[25] so outliers may drive the results; this is particularly important because of the presence of superstars.

In conclusion, the visual inspection of the data suggests that players born early in the admission year are over-represented and that the players' birthrate by quarter decreases with the distance from the admission date. The visual inspection also suggests that players born in the last quarter of the admission year receive a lower entry wage than players born in other quarters; they earn lower wages also throughout the career.

4 Methods and Results

4.1 RAE in Terms of Representativeness

In presence of RAE in terms of representativeness, the observed distribution of the quarter of birthrate should differ from the expected distribution. Players born at the beginning of the admission year should be over-represented, while players born at the end of the admission year should be under-represented. Moreover, there should be a specific birthrate trend: the birthrate should decrease with the distance from the admission date.

A chi-square goodness-of-fit test is used to compare the difference between observed and expected amount of players across quarters of birth (e.g., Sims & Addona, 2014; Helsen et al., 2012). The observations from the seven seasons are pooled. The expected amount of players is based on the average quarter of birthrate, in Italy, between 1993 and 1998; data on previous years are unavailable. The uniform birthrate distribution is not assumed, because of the seasonality of birth for the general Italian population (e.g., Rizzi & Dalla Zuanna, 2007; Prioux, 1988). Table 1 shows the results from this analysis.

[25] There are 27 players -season observations at 18 years of age, or less, and 33 players -season observations at 37 years of age, or more.

Table 1. Chi-squared goodness-of-fit by quarter of birth.

Quarter	Observed counts	Expected counts	Difference
Q1 (January-March)	508	406	102
Q2 (April-June)	534	434	100
Q3 (July-September)	389	450	-61
Q4 (October-December)	273	414	-141
$\chi^2(3)$		7.81	
P-value		0.000	

Note: "Expected counts" is the expected amount of players born in each quarter. "Actual counts" is the observed amount of players born per month. "Difference" provides the differences between observed and expected counts, which are used to compute the chi-squared statistics.

The table confirms the insights provided by the descriptive statistics: Serie A is characterized by the presence of RAE in terms of representativeness. The distribution of the observed quarters of birthrate for Serie A Italian players is statistically significantly different from its expected distribution. This result is in line with other studies which analyze the RAE in Serie A (e.g., Salinero, Pérez, Burillo & Lesma, 2013; Helsen et al., 2012).

Furthermore, the column "Difference" suggests the existence of a specific trend in the players' quarters of birthrate. In fact, in presence of RAE players born at the beginning of the admission year are over-represented, the extent of this over-representation reduces with the distance from the admission month, and then turns into under-representation which increases moving toward the end of the admission year. The formal analysis on the existence of this specific birthrate trend is implemented with the Spearman-rank correlation coefficient (for a similar application of this test, see Musch & Hay, 1999, and Ashworth & Heyndels, 2007).

This Spearman-rank correlation coefficient is computed between two measurement variables converted to ranks (McDonald, 2014). One variable is the *months representativeness* in Serie A and it is based on the differences between the expected amount of players, based on the Italian population month of birthrate, and the observed amount of players in each month. When the difference between the expected and the observed amount of observations is negative, the players born in that month are over-represented; vice-versa, a positive difference signals under-representation. The first place in the ranking is assigned to the most under-represented month, while the last place is assigned to the most over-represented month. The measurement variable represents the *admission date distance*, which is based on the distance of the month from the admission date (i.e., January has the first position in the ranking, whereas December has the last position) and measures the RA.

If the Spearman rank-correlation coefficient was computed simply between the ranking based on the observed counts—in lieu of the differences between the expected and observed amount of players—and the admission date distance, the possible presence of trends in the months birthrate for the underlying general population would not be taken into account. If already in the general population the month birthrate was to increase in the distance from the admission date, the results could provide artifactual evidence of RAE.

How to interpret the output of the Spearman-rank correlation coefficient test? On one hand, if the estimate of the correlation was negative and statistically significant, there would be evidence of the specific trend characterizing the RAE, that is, players born in early months are over-represented and, with the increase in the distance of the month of birth from the admission date, the over-representation reduces and eventually players born toward the end of the year would be under-represented. On the other hand, if the correlation was positive and statistically significant, the trend would have opposite direction and there would be evidence of reverse RAE. In this case, we would observe under-representation in early months, the months birthrate would tend to increase with the distance from the admission date, such that players born toward the end of the year would be over-represented. Based on the RAE mechanism previously discussed, we expect to find a negative correlation between months order and admission date distance. The H_0 of the Spearman-rank correlation coefficient test is that the correlation between the two rankings is zero.[26]

Table 2. Correlation between months representativeness and admission date distance.

Month	Months representativeness				Admission date distance
	Expected counts	Actual counts	Difference	Ranking (1)	Ranking (2)
January	138,8	234	-95,17	12	1
February	126,4	120	6,41	7	2
March	140,2	154	-13,81	9	3
April	138,1	139	-0,92	8	4
May	149,2	206	-56,81	11	5
June	145,9	189	-43,07	10	6
July	154,6	138	16,60	5	7
August	146,5	136	10,52	6	8
September	149,0	115	34,01	3	9
October	144,8	111	33,80	4	10
November	132,6	89	43,59	2	11
December	136,9	73	63,87	1	12
Spearman	-0.860				
	(0.000)				

Note: The shaded areas include the figures of interest. "Expected counts" is the expected amount of players born in each month, based on the Italian average monthly birthrate from 1993 to 1998. "Actual counts" is the observed amount of players born per month. "Difference" provides the differences between observed and expected counts. "Ranking (1)" is the month ranking based on the differences between expected and actual counts: the first place is assigned to the most under-represented month, i.e., the month with the largest positive difference; the last place is assigned to the most over-represented month, i.e., the month with the largest negative difference. "Ranking (2)" is the distance of the month of birth from the admission date; this ranking is directly established by the "Admission date distance;" for example, January, February and March receive ranks 1, 2, 3. "Spearman" is the estimate of the Spearman-rank correlation coefficient; the corresponding P-value is in parenthesis.

Table 2 presents a highly statistically significant and negative correlation between months representativeness and admission date distance, providing evidence for a downward trend in the soccer players' birthrate. This result reinforces the evidence of RAE in terms of representativeness.

[26] The usage of the two-tailed test is motivated by the possibility to have a reverse RAE as well.

4.2 RAE in Terms of Wages

The sign of the RAE in terms of wage gaps might differ from the sign of the RAE in terms of representativeness. It is possible to have RAE in terms of representativeness, yet reverse RAE in terms of wages. Relatively young players enjoy higher, equal or lower wages depending on several characteristics of the streaming process. As discussed in the introduction, relatively young players might have been exposed to positive peer effects or might have been positively selected (e.g., Ashworth & Heyndels, 2007; Gibbs et al., 2012; Bryson et al., 2014), which eventually benefits them in terms of wages. Differently, it is possible that the path dependence might have increased original differences in performances (Allen & Barnsley, 1993), which eventually disadvantage them in terms of wages. Finally, discriminatory criteria which cause original differences in achievements might cease to be relevant in the labor market (Crawford et al., 2013), in this case there could be an equalization of the wages. The analysis in this section focuses on the empirical sign of RAE in terms of wages.

The descriptive statistics suggest that players born in the last quarter receive gross wages in the bottom percentiles of the wage distribution more frequently than players born in other quarters. However, other characteristics which determine the wage are not controlled for, so it is not possible to gain clear insights.

The empirical investigation of wage gaps proceeds with a standard methodology used in economics of sports: the pooled OLS regression.[27] The model is the following.

$$Ln(w_i) = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \boldsymbol{\beta RA_i} + \varepsilon_i$$

The natural logarithm of the deflated gross wage for player *i* is the outcome variable. The set of control variables includes age_i and age_i^2, both of them refer to player i's age; age is

[27] See for instance Bryson et al. (2014) as well as Frank and Nüesch (2012).

computed in years, instead of in months as in Ponzo and Scoppa (2014), since there is no information on the date of signing the contract. This variable is rescaled; it ranges from 0 to 26—since in our dataset the minimum registered age for Italian players is 17 whereas the maximum is 43—so that the estimate of the constant is directly interpretable. The variable for squared age captures the decreasing return to age. $\boldsymbol{RA_i}$ is a vector of dummy variables for quarter of birth, where the first quarter of the admission year (i.e., January to March) is the reference quarter. Their estimated coefficients represent the estimates of RAE for different quarters, and reflect differences in both sheer maturity and productivity. The RAE is unbiased if two assumptions hold true: i) date of birth is unrelated to innate ability, this assumption is also called “nonastrology assumption” (Allen & Barnsley, 1993, p. 654); ii) season-of-birth is unrelated to players’ performance, which also implies the absence of a relationship between household socio-economic status and birthdate.[28] Finally, someone may argue that the estimates capture the combined RAE from sports and school, since soccer and school admission years overlap; however, previous literature suggests that education achievements do not affect returns from playing soccer at professional level (Barros, 2001).

As Ashworth and Heyndels (2007) explain, the inclusion of other variables normally used in economics of sports—for instance measures of players’ performance—would cause multicollinearity. For the same reason, controls for players’ experience are not included. There is no problem of collinearity with age,[29] but there might be collinearity with the RA

[28] The existence of the correlation between date of birth and socio-economic status seems to differ by country. For instance, some studies suggest the presence of such a correlation in the US (Bound & Jaeger, 2001; Cascio & Lewis, 2005; Buckels & Hungerman, 2013), in Sweden (Carlsson et al., forthcoming), and in Austria (Doblhammer & Vaupel, 2001). However, a recent study suggests that in Norway there is no such correlation (Black et al., 2011). Moreover, for Italy, this correlation is not systematically analyzed, but a previous study on RAE on school performances suggests that it is non-existent (Ponzo & Scoppa, 2014). Bound and Jaeger (2001) mention the presence of additional season-of-birth effects on health, e.g., mental health, but it can be ignored as it involves only a small number of individuals in the whole population (Plug, 2001).

[29] In European sports athletes may enter professional competitions at different ages. Therefore, the usage of both age and experience does not create multicollinearity. Differently, in studies on US sports, the introduction of both variables would create multicollinearity, as the drafting system is such that athletes enter professional leagues at a somewhat uniform age (Lucifora & Simmons, 2003).

variable: in presence of RAE, relatively old players also benefit from more played time, which increases their experience more rapidly.[30]

As a robustness check, the analyses are re-run adding variables for specific effects on wage heterogeneity. There is one vector for teams and one for season fixed-effects.[31] The favorite model is that including all the fixed-effects.

As additional robustness check, the analyses are re-run on a discontinuity sample (e.g., Ponzo & Scoppa, 2014; Black et al. 2011). This strategy consists in focusing on the narrower sample of footballers born either in January or in December; these are two adjacent months, one is after the admission date (1st of January), whereas the other comes right before it. As suggested in Barnsley and Thompson (1988), in the analysis of a discontinuity sample season-of-birth effects should be eliminated, since the two months are in the same season. All the fixed-effects are included in this analysis.

Since repeated observations on individual players are not likely to be independent, standard errors are clustered on footballers in all the analyses.

Although all the estimates are reported, the focus of the analyses is on the comparison between footballers born in the fourth and first quarters, or in December and January.

The estimates, also for the robustness check, are reported in Table 3.

---

[30] Ashworth and Heyndels (2007) do not find such a correlation, so they introduce also experience in the model.

[31] This paper cannot investigate the effect of the increased competition from foreign players, as the amount of foreign players augments constantly throughout the seven seasons in exam; hence, seasons fixed-effects would be perfectly collinear with this measure of competition.

Table 3. Wage gap between relatively old and young players.

Variables	(1)	(2)	(3)	(4)
Age	0.227***	0.225***	0.224***	0.270**
	(0.039)	(0.039)	(0.028)	(0.105)
Age^2	-0.008***	-0.008***	-0.008***	-0.010**
	(0.002)	(0.002)	(0.001)	(0.004)
Q2 (April-June)	-0.062	-0.061	-0.053	-
	(0.093)	(0.093)	(0.055)	
Q3 (Jule-September)	0.181	0.179	0.045	-
	(0.114)	(0.114)	(0.069)	
Q4 (October-December)	-0.203*	-0.207*	-0.177**	-
	(0.114)	(0.114)	(0.0752)	
December	-	-	-	-0.369*
				(0.186)
Constant	-1.284***	-1.215***	-1.366***	-1.841**
	(0.225)	(0.249)	(0.215)	(0.732)
F-test, quarters of birth (p-value)	0.022	0.022	0.033	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.142	0.151	0.560	0.658
Observations	1,598	1,598	1,598	282

Note: Standard errors clustered on footballers are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects. The sample size for the analyses in column (1)-(2)-(3) is smaller than 1,704 observations because of missing values on wage. Column (4) reports the results obtained with the discontinuity sample, i.e., only players born in January or December; also this analysis is conducted with OLS and includes players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

This table provides evidence for RAE on wage gaps. The results displayed in column (3) provide statistically significant evidence that, ceteris paribus, a player born in the fourth quarter of the admission year receives a wage which is about 19% lower than that earned by footballers born in the first quarter.[32] The p-value from the F-test on the joint significance of the estimates for the quarter of birth coefficients suggests that these three coefficients considered together are statistically significant. The estimates in column (4) are obtained from the analysis of the discontinuity sample; an even larger and statistically significant wage gap is obtained.

[32] The wage gap in percentage terms is computed as [exp(0.177)-1]*100= 19.3%.

These results are robust to two robustness checks. First, the results are confirmed after winsorization at 3% and 97%, see Appendix C, Table C.1. This analysis is conducted to verify whether the evidence of RAE on wages is led by wage outliers. Second, the results are confirmed also when players in their first season in the dataset and players who have changed team are dropped from the dataset. The reason for conducting this robustness check is the following: since the signing date of the contract is not known, it is not possible to differ between players whose first wage refers to the whole season or only to a few months (e.g., a player signs a contract for a new team in January and this first season wage could refer only to the period from January to the end of the season, which is May or June).

5 Heterogeneity Analyses on Wage Gaps

In this section heterogeneity analyses on the RAE in terms of wage gaps are implemented. The wage gap might differ across ages, as the sheer maturity differential decreases[33] and additional selection of the players occurs throughout the career.[34] Additionally, the wage gap might differ across quantiles of the wage distribution; the differences might be particularly important in light of the positive skewedness of the wage distribution. While the main analysis conducted with OLS focuses on the sign of the RAE in terms of wages, these additional analyses might provide additional insights.

Does the wage gap change over the footballers' career? Figure 5 suggests that the answer might be positive. To formally investigate this research question, the analyses are re-run for different categories of age: footballers younger than 21 (players who can still compete in the last youth category, "Primavera," and might present physical development differentials), footballers between 21 and 25 (players for whom complete cognitive maturity is still to be reached), between 26 and 30 (career core), and older than 30 (retirement period).

[33] The introduction mentions that there might still be a maturity differential due to RA up until mid-twenties.
[34] Some players can leave (permanently or temporarily) Serie A during their career for different reasons; other players can start their career in Serie A when they are older. This might change the wage gap at different ages; the direction of this change depends on the performance of the players who leave or join Serie A.

This approach by category of age to study the evolution of RAE is used also in Black et al. (2011).[35] For brevity, only the estimates obtained with season and players' team fixed-effects are reported in Table 4.

Table 4. Wage gap between relatively old and young players, by age group.

	< 21	21-25	26-30	> 30
Variables	(1)	(2)	(3)	(4)
Age	0.825	0.181	0.206	0.330**
	(0.665)	(0.166)	(0.232)	(0.143)
Age^2	-0.049	-0.059	-0.007	-0.011***
	(0.183)	(0.013)	(0.010)	(0.004)
Q2 (April-June)	-0.344	-0.092	-0.038	-0.099
	(0.608)	(0.084)	(0.069)	(0.089)
Q3 (Jule-September)	-0.866*	-0.013	0.076	-0.092
	(0.509)	(0.107)	(0.093)	(0.114)
Q4 (October-December)	-1.129	-0.195*	-0.240***	-0.134
	(0.906)	(0.108)	(0.092)	(0.110)
Constant	-3.517**	-1.011**	-1.357	-1.975
	(1.402)	(0.532)	(1.249)	(1.238)
F-test, quarters of birth (p-value)	0.291	0.223	0.018	0.213
Season F.E.	Y	Y	Y	Y
Team F.E.	Y	Y	Y	Y
R-square	0.841	0.491	0.550	0.635
Observations	51	403	629	515

Note: Standard errors clustered on players are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Column (1) reports the results from the OLS with players' team and season fixed-effects for players younger than 21 years of age; column (2) reports the results from the OLS with players' team and season fixed-effects for players between 21 and 25 years of age; column (3) reports the results from the OLS with players' team and season fixed-effects for players between 26 and 30 years of age. Column (4) reports the results from the OLS with players' team and season fixed-effects for players older than 30 years of age. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

These results confirm the presence of RAE in terms of wage gaps and suggest wage gaps might evolve along players' career. The estimates for players born in the second and third quarter confirm the results from the main analysis. The players born in these two quarters do not suffer from any statistically significant wage gap, at any age-group. The p-value from the F-test on the joint significance of the estimates for the quarter of birth coefficients suggests

[35] The difference with Black et al. (2011) is that they use single ages from 24 to 35, whereas here four age-groups are used; the choice of multiple age categories is due to the sample size.

that these three coefficients considered together are statistically significant only in the age-group between 26 and 30 (career core). These results are similar also after winsorization at 3% and 97%, see Appendix C, Tables C.2.

Although the dataset does not allow for further investigation of the wage gap dynamics through players' career, it is possible to interpret these results based on the existing literature. Players born in the fourth quarter suffer from the largest wage gap under 21 years of age; this estimate is not statistically significant (there is a very limited amount of observations for this age-group). Once in the labor market, the performance gap between these new relatively young players and the older peers might be relevant. Therefore, between 21 and 25 years of age, the worst performers among the relatively young players might decide to leave Serie A to gain experience in lower leagues or abroad, consequently the wage gap decreases sensibly; this gap is not statistically significant though. During this period away from Serie A, the part of the RAE due only to the sheer maturity differential is filled.[36] When, between 26 and 30 years of age, these players re-enter Serie A,[37] they are still worse performers than older peers on average: the sheer maturity gap is now filled, but tangible and intangible skills might still differ because of the RAE. Therefore, the wage gap increases again,[38] even if it is now smaller compared to the entry wage gap they suffered at the beginning of their career. Finally, after 30 years of age, the wage gap might widen even more because the characteristics of the players who decide to retire might differ by quarter of birth. For example, if the best players among those who are born in the first quarter tend to retire late, while the best players from the fourth quarter retire early, the average performance gap diverges, even if on average these two groups of players retire at the same age.

---

[36] Note that the same occurs if the best players among the relatively old players leave Serie A to gain more experience.

[37] Similarly, Parent-Harvey et al. (2013) and Böheim and Lackner (2012) suggest that relatively young athletes might delay the entry into professional sport to wait for the maturity gap to be filled.

[38] The wage gap might increase also because the best relatively old players who left Serie A are now back.

Does the wage gap differ over different quantiles of the wage distribution? This question makes empirical sense: since the distribution of players' wages is positively skewed and characterized by the presence of superstars, the normality assumption does not hold. Moreover, the analysis at the median, that is, the 50th percentile of the wage distribution, is more robust to wage outliers—including superstars—than the OLS. The OLS estimates describe the relationship between the regressors and the conditional mean of the outcome variable, whereas the quantile regression describes the relationship between the regressors and the conditional quantile of the outcome variable; therefore, the quantile regression gives a more comprehensive picture of the relationship. Results from the quantile regression also help to understand whether there are positive peer effects for and selections of relatively young players; this point is illustrated below and is important to understand the RAE mechanism. To the best of my knowledge, this is the first time the quantile regression is used to analyze the RAE on wages. Analyses are implemented at the 25th, 50th, 75th and 90th percentiles of the wage distribution.[39] For brevity, only the estimates obtained with season and players' team fixed-effects are reported in Table 5.

[39] Since repeated observations on individual players are not likely to be independent, standard errors are clustered on players,. The employed method is that suggested in Parente and Silva (2013).

Table 4. Wage gap between relatively old and young players.
Quantile regression at 25%, 50%, 75% and 90% of the wage distribution.

	25%	50%	75%	90%
Variables	(1)	(2)	(3)	(4)
Age	0.258***	0.193***	0.163***	0.132***
	(0.030)	(0.032)	(0.027)	(0.036)
Age^2	-0.009***	-0.007***	-0.006***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Q2 (April-June)	-0.029	-0.030	-0.057	-0.172**
	(0.055)	(0.052)	(0.078)	(0.072)
Q3 (Jule-September)	0.025	0.061	0.096	0.067
	(0.066)	(0.071)	(0.103)	(0.072)
Q4 (October-December)	-0.128	-0.097	-0.113	-0.158
	(0.085)	(0.063)	(0.091)	(0.109)
Constant	-2.055***	-1.222***	-0.355	1.096***
	(0.349)	(0.227)	(0.221)	(0.242)
F-test, quarters of birth (p-value)	0.317	0.172	0.310	0.000
Season F.E.	Y	Y	Y	Y
Team F.E.	Y	Y	Y	Y
Pseudo R-square	0.537	0.544	0.538	0.505
Observations	1,598	1,598	1,598	1,598

Note: Standard errors clustered on players are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The samples for the analyses use less than 1,704 observations because of missing values on wage. Column (1) reports the results from the quantile regression at the 25th percentile of the wage distribution; column (2) reports the results from the quantile regression with players' team and season fixed-effects at the 50th percentile of the wage distribution; column (3) reports the results from the quantile regression with players' team and season fixed-effects at the 75th percentile of the wage distribution; column (4) reports the results from the quantile regression with players' team and season fixed-effects at the 90th percentile of the wage distribution. The sample size for the analyses in column (1)-(2)-(3) is smaller than 1,704 observations because of missing values on wage. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

Two important aspects of Table 5 deserve some attention. On one hand, the analysis at the median, column (3), reports quite a lower wage gap than that obtained with the OLS for players born in the fourth quarter.[40] On the other hand, the wage gap seems to increase in the quantile of the wage distribution. If these results were to be taken seriously, there would be two implications. First, there might not be positive peer effects and selection for relatively young Serie A players; otherwise a reduction of the wage gap, or even its reversal, would have been observed in top quantiles of the wage distribution rather than its increase.[41] Second,

[40] Interpretation of the estimated coefficients is similar to that for OLS estimates (Hao & Naiman, 2007).
[41] Gibbs et al. (2011) show that the RAE in terms of representativeness could reverse among the very best players; therefore, also a reversal of the RAE in terms of wages among the very best players is plausible.

the RAE on wages seems to be driven by the effect on wages in the top quantiles of the distribution; this tendency is not captured by the analysis with winsorized wages (see Table C.1), which is a methodology frequently adopted for robustness checks.

It is worth mentioning that the RAE might differ also based on players' roles. The different effects might depend on at least three factors: i) the innate ability might be more important and compensate to some extent differences in maturity at an early age for some roles, for instance, for forwards and their "tor-instinct" (Ashworth & Heyndels, 2007, p. 368), ii) the identification of role-specific skills might be easier for some roles, for instance, for forwards (Ashworth & Heyndels, 2007), iii) the importance of physical characteristics is greater for goalkeepers and defenders (Salinero et al., 2013). Since the reasons that lead to possible differences in RAE by players' role are peculiar to sports disciplines, the results are reported and commented more extensively in Appendix D. In summary, the RAE seems to be larger for goalkeepers and midfielders, while it seems to be non-existent for forwards.

6 Discussion and Conclusions

Chronological differences between individuals within the same age-group, that is, relative age (RA), determine maturity gaps during childhood, both in school and in sport contexts. These differences translate into a performance gap (e.g., Musch & Hay, 1999; Musch & Grondin, 2001; Bedard & Dhuey, 2006; Dhuey & Lipscomb, 2009), that is, relative age effect (RAE), which should disappear over time. However, because of streaming, competition and social interactions, the performance gap might extend to the long-run (e.g., Allen & Barnsley, 1993; Bedard & Dhuey, 2006), affecting labor market outcomes.

This paper focuses on two aspects of RAE on the labor market for a particular group of high skilled workers: professional soccer players. First, it investigates the existence of RAE in terms of representativeness, that is, whether players born at the beginning of the admission year are over-represented. Second, it investigates the presence of RAE in terms of wage gaps.

Heterogeneity analyses on the evolution of this possible wage gap across ages and quantiles of the wage distribution are also carried out.

This study provides statistical significant evidence for RAE in terms of representativeness, with relatively old players being over-represented in Serie A (e.g., Barnsley & Thompson, 1988; Musch & Hay, 1999; Musch & Grondin, 2001; Böheim & Lackner, 2012). The analyses also suggest the existence of a specific trend in the birthrate distribution, further reinforcing the evidence of RAE: the over-representation decreases and turns into under-representation as the end of the admission year is approached.

This paper also provides statistically significant evidence of RAE on players' wages, with relatively young players, that is, players born late in the admission year, earning on average lower gross wages. The size of the wage gap, caused by the RAE, is economically important; in particular, it is important for players born in December, which is the last month of the admission year. Furthermore, this wage gap appears to increase with age, after an initial reduction. I speculate that this trend could be owed to players' mobility in and out of Serie A, which affects the distribution of the characteristics of the players who do not move. The analyses on different wage quantiles suggest that wage gaps tend to increase in the quantiles of the wage distribution, and thus that the estimates obtained with the OLS are driven by differences in the top quantiles of the wage distribution. These results seem to rule out the presence in Serie A of positive peer effects and selection in favor of relatively young players.

These analyses should be considered carefully. Since the admission date for the players in this sample does not vary it is not possible to definitely rule out any season-of-birth effect on wages.

Also, these results on wage gaps are somewhat different from those in the only other study on RAE in the soccer players' labor market, in Germany (Ashworth & Heyndels, 2007).

However, this might be due to the type of data and the different characteristics of the Italian and German institutional contexts.[42]

Future analyses of RAE on wages should include a quantile regression. This is particularly important when the wage distribution is expected to be positively skewed; consider the case of studies on RAE among CEOs (e.g., Malmendier & Tate, 2009; Frank & Nüesch, 2012). The quantile regression is important also to investigate positive peer effects and selection, which might characterize relatively young workers.

Further work is required to increase the knowledge on short- and long-run RAE. Future studies on wage gaps should exploit variations in the admission date in order to gain clearer evidence of RAE. This variation can be obtained in cross-country analyses, similarly to Bedard and Dhuey (2006) and Munch and Hay (1999). Moreover, different aspects, such as career promotions, retirement and migration decisions, have yet to be analyzed.

Additional evidence in favor of RAE would call for a revision of the age-grouping system. Age-groups could be shortened, for instance, 6 or 9 months (e.g., Pellizzari & Billari, 2012; Barnsley & Thompson, 1988) instead of 12 months, so that within-age-group performance differences would be reduced, potentially decreasing the wage gaps on the labor market. In alternative, the admission date could rotate in different ways, so that the RAE would not consistently provide advantages to people born in a given month (Barnsley & Thompson, 1988; Wattie et. al, 2015).

[42] Italian and German youth categories systems might be differ. Moreover, Ashworth and Heyndels (2007) use data from 1997-98 and 1998-99, immediately after the Bosman ruling, 1995, imposed by the European Court of Justice. This ruling affected the soccer players' labor market, it banned limitations on the amount of players from EU countries and introduced the free-agency. The data used in this paper are almost a decade more recent, so longer time has passed by allowing the Bosman ruling to fully affect the soccer players' labor market. During the seven seasons in analysis the amount of foreign players constantly increased, while until 2008 the amount of foreign players in Italy was among the lowest in EU (Bryson et al., 2014). Since the Bosman ruling affects players' competition, it could also affect the RAE on wages; however, the expected sign of this effect is uncertain: it depends on many factors, such as the top tax rates, which determine the quality of the players who move (Kleven et al., 2013).

Bibliography

Adler, M., (1985). Stardom and talent. *The American Economic Review. , 75*(1), 208-212.

Allen, J., & Barnsley, R. (1993). Streams and Tiers: The Interaction of Ability, Maturity, and Training in Systems with Age-Dependent Recursive Selection. *Journal of Human Resources*, *28*(3), 649-659.

Angrist, J. D., & Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association, 87*(418), 328-336.

Ashworth, J., & Heyndels, B. (2007). Selection Bias and Peer Effects in Team Sports The Effect of Age Grouping on Earnings of German Soccer Players. *Journal of Sports Economics, 8*(4), 355-377.

Barnsley, R. H., & Thompson, A. H. (1988). Birthdate and Success in Minor Hockey: The Key to the NHL. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 20*(2), 167-176.

Barnsley, R. H., Thompson, A., & Legault, P. (1992). Family planning: Football Style. The Relative Age Effect in Football. *International Review for the Sociology of Sport, 27*(1), 77-87.

Barros, C. P. (2001). Economic Return on Schooling for Soccer Players. *Journal of Sports Economics, 2*(4), 369-378.

Bedard, K., & Dhuey, E. (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics, 121*(4), 1437-1472.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *The Review of Economics and Statistics, 93*(2), 455-467.

Böheim, E., & Lackner, M. (2012). Returns to Education in Professional Football. *Economics Letters, 114*(3), 326-328.

Bound, J., & Jaeger, D. A. (2001). Do Compulsory School Attendance Laws Alone Explain the Association Between Quarter of Birth and Earnings? *Research in Labor Economics, 19*, 83-108.

Bryson, A., Rossi, G., & Simmons, R. (2014). The Migrant Wage Premium in Professional Football: A Superstar Effect? *Kyklos, 67*(1), 12-28.

Bryson, A., Gomez, R., & Zhang, T. (2014). All-star or benchwarmer? Relative age, cohort size and career success in the NHL. *IZA Discussion Paper*, 8645.

Buckles, K. S., & Hungerman, D. M. (2013). Season of Birth and Later Outcomes: Old Questions, New Answers. *Review of Economics and Statistics, 95*(3), 711-724.

Carling, C., le Gall, F., Reilly, T., & Williams, A. M. (2009). Do Anthropometric and Fitness Characteristics Vary According to Birth Date Distribution in Elite Youth Academy Soccer Players? *Scandinavian Journal of Medicine & Science in Sports. 19*(1), 3-9.

Carlsson, M., Dahl, G., Öckert, B., & Rooth, D.-O. (forthcoming). The Effect of Schooling on Cognitive Skills. The Review of Economics and Statistics.

Cascio, E. U., & Lewis, E. G. (2006). Schooling and the Armed Forces Qualifying Test: Evidence from School-Entry Laws. *The Journal of Human Resources, 41*(2). 294-318.

Cascio, E. U., & Schanzenbach, D. (2007). First in the Class? Age and the Education Production Function. *Discussion Paper 13663, NBER.*

Cobley, S., McKenna, J., Baker, J., & Wattie, N. (2009) How Pervasive are Relative Age Effects in Secondary School Education? *Journal of Educational Psychology, 101*(2), 520-528.

Crawford, C., Dearden, L., & Greaves, E. (2013). *The Impact of Age Within Academic Year on Adult Outcomes.* Institute for Fiscal Studies, 1-46.

Dixon, J., Horton, S., & Weir, P. (2011). Relative Age Effects: Implications for Leadership Development. *International Journal of Sport & Society, 2*(2), 1-15.

Dhuey, E., & Lipscomb, S. (2009). What makes a leader? Relative Age and High School Leadership. *Economics of Education Review, 27*(2), 173-183.

Doblhammer, G., & Vaupel, J. W. (2001). Lifespan Depends on Month of Birth. *Proceedings of the National Academy of Sciences, 98*(5), 2934-2939.

Du, Q., Gao, H., & Levi, M. D., (2012). The Relative-Age Effect and Career Success: Evidence from Corporate CEO. *Economics Letters, 117*(3), 660-662.

Edgar, S., & O'Donoghue, P. (2005), Season of Birth Distribution of Elite Tennis Players. *Journal of sports sciences, 23*(10), 1013-1020.

Frank, E., & Nüesch, S. (2012). Talent and/or Popularity: What Does it Take to Be a Superstar? *Economic Inquiry, 50*(1), 202-216.

Fredriksson, P., & Öckert, B. (2006). Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance. *Working Paper, IFAU-Institute for Labour Market Policy Evaluation. 12*, 1-52.

Gibbs, B. G., Jarvis, J. A., & Dufur, M. J. (2012). The Rise of the Underdog? The Relative Age Effect Reversal Among Canadian-Born NHL Hockey Players: A Reply to Nolan and Howell. *International Review for the Sociology of Sport, 47*(5), 644-649.

Gladwell, M. (2008). *Outliers: The story of success*. Penguin UK.

Glamser, F. D., & Vincent, J. (2004). The Relative Age Effect Among Elite American Youth Soccer Players. *Journal of sport Behavior, 27*(1), 31-38.

Hancock, D. J., Adler, A. L., & Côté, J. (2013). A Proposed Theoretical Model to Explain Relative Age Effects in Sport. *European journal of sport science, 13*(6), 630-637.

Hao, L., & Naiman, D. Q. (2007). *Quantile regression* (149). Sage.

Helsen, W. F., Starkes, J. L., Van Winckel, J. (1998). The Influence of Relative Age on Success and Dropout in Male Soccer Players. *American Journal of Human Biology, 10*(6), 791-798.

Helsen, W. F., Starkes, J. L., & Van Winckel, J. (2000). Effect of a Change in Selection Year on Success in Male Soccer Players. *American Journal of Human Biology, 12*(6), 729-735.

Helsen, W. F., Baker, J., Michiels, S., Schorer, J., Van Winckel, J., Williams, A. M. (2012). The Relative Age Effect in European Professional Soccer: Did Ten Years of Research Make any Difference? *Journal of Sports Science, 30*(15), 1665-1671.

Joyner. P. W., Mallon, W. J., Kirkendall, D. T., & Garret Jr. W. E. (2013) Relative Age Effect: Beyond the Youth Phenomenon. *The Duke Orthopaedic Journal, 3*(1), 74-79.

Kleven, H. J., Landais, C., & Saez E. (2013). Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review, 103*(5), 1892-1924.

Kniffin, K. M., & Hanks, A. S. (2013). Revisiting Gladwell's Hockey Players: Influence of Relative Age Effects upon Earning the PhD. *CHERI Working Paper*, 157.

Kuczmarski, R. J., Ogden, C. L., Guo, S. S., Grummer-Strawn, L. M., Flegal, K. M., Mei, Z., Wei, R., Curtin, L. R., Roche, A. F., & Johnson, C. L. (2002). 2000 CDC Growth Charts for the United States: Methods and Development. Vital and health statistics. Series 11, Data From the National Health Survey, 246, 1-190.

Larsen, E. R., & Solli, I. F. (2012) *Born to run Behind? Persisting Relative Age Effects on Earnings.* University of Stravenger, Faculty of Social Sciences.

Levitt, S. D., & Dubner, S. J. (2010). Superfreakonomics: Global cooling, patriotic prostitutes and why suicide bombers should buy life insurance. Penguin UK.

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

Malmendier, U., & Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics, 124*(4), 1593-1638.

McDonald, J.H. 2014. Handbook of Biological Statistics. Sparky House Publishing, Baltimore, Maryland.

Muller-Daumann, D., & Page, D. (2013). Political Selection and the Relative Age Effect. *QuBE Working Papers 009, QUT Business School.*

Musch, J., & Grondin, S. (2001). Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport. *Development Review, 21*(2), 147-167.

Musch, J., & Hay, R. (1999). The Relative Age Effect in Soccer: Cross-Cultural Evidence for a Systematic Distribution Against Children Born Late in the Competition Year. *Sociology of Sport Journal, 16*(1), 54-6.

Parent-Harvey C. I., Desjardins C. & Harvey, E. J. (2014). Factors Affecting the Relative Age Effect in NHL Athletes. *Canadian Journal of Surgery, 57*(3), 157-161.

Parente, P. M. D. C., & Santos Silva, J. M. C. (2013). Quantile Regression with Clustered Data. *University of Exeter, Economics Department Discussion Papers Series, 13*(5), 1-24.

Pellizzari, M., & Billari, F. C. (2012). The Younger, the Better? Age-Related Differences in Academic Performance at University. *Journal of Population Economics, 25*(2), 697-739.

Plug, E. J. S. (2001). Season of Birth, Schooling and Earnings. *Journal of Economic Psychology, 22*(5), 641-660.

Ponzo, M., & Scoppa, V. (2014). The Long-Lasting Effects of School Entry Age: Evidence from Italian Students. *Journal of Policy Modeling, 36*(3), 578-599.

Prioux, F. (1988). Mouvement Saisonnier des Naissances: Influence du Rang et de la Légitimité dans Quelques Pays d'Europe Occidentale. *Population,3*, 587-609.

Rizzi, E. L., & Dalla-Zuanna (2007). The Seasonality of Conception. *Demography, 44*(4), 705-728.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*. *71*(5), 845-858.

Salinero, J. J., Pérez, B., Burillo, P., & Lesma, M. L. (2013). Relative Age Effect in European Professional Football. Analysis by position. *Journal of Human Sport & Exercise, 8*(4), 966-973.

Salthouse, T. A., Schroeder D. H., & Ferrer, E. (2004). Estimating Retest Effects in Longitudinal Assessments of Cognitive Functioning in Adult between 18 and 60 Year of Age *Developmental Psychology, 40*(5), 813-822.

Sims, J., & Addona, V. (2014). Hurdle Models and Age Effects in the Major League Baseball Draft. *Journal of Sports Economics,* 1-16.

Thompson, A. H., Barnsley, R. H., & Battle, J. (2004). The Relative Age Effect and the Development of Self-Esteem. *Educational Research, 46*(3), 313-320.

Thompson, A. H., Barnsley, R. H., & Dyck, R. J. (1999). A New Factor in Youth Suicide: The Relative Age Effect. *Canadian journal of psychiatry, 44*(1), 82-85.

Wattie, N., Schorer, J., & Baker, J. (2015). The Relative Age Effect in Sport: A Developmental Systems Model. *Sports Medicine, 45*(1), 83-94.

Williams J., H. (2010). Relative Age Effect in Youth Soccer: Analysis of the FIFA U17 World Cup Competition. *Scandinavian Journal of Medicine and Science in Sports, 20*(3), 502-508.

Zoëga, H., Valdimarsdóttir, U. A., & Hernández-Diaz, S. (2012). Age, Academic Performance, and Stimulant Prescribing for ADHD: A Nationwide Cohort Study. *Pediatrics, 130*(6), 1012-1018.

Appendix A

Table A.1 Variables description.

Variable	Description
Ln(w)	Natural logarithm of the gross wage, deflated at 2013, for player i.
Age	Age of player i. This variable is rescaled; range from 0 to 26 (0 correspond to 17, while 26 corresponds to 43).
Age^2	Squared age of player i.
Quarter dummies	Relative age, in terms of quarter of birth, for player i. The reference quarter is January-March.
December dummy	Relative age for players i born in December. The reference month is January.
Team dummies	Dummies for team; Udinese is the reference team.
Season dummies	Dummies for season; 2013-14 is the reference season.

Table A.2 Descriptive statistics for wage and age, by quarter of birth.

Variable	Whole sample	Q1	Q2	Q3	Q4
Ln(w)	0.126	0.128	0.078	0.310	-0.048
	(0.835)	(0.813)	(0.753)	(0.894)	(0.894)
Age	28.046	27.777	28.095	28.082	28.399
	(4.469)	(4.483)	(4.392)	(4.596)	(4.406)

Note: Mean Ln(w) and age, standard deviation in parenthesis.

Appendix B

Table B.1 reports the amount of births in the Italian general population, from 1993 to 1998. It contains the statistics used to compute the Italian quarterly and monthly birthrates

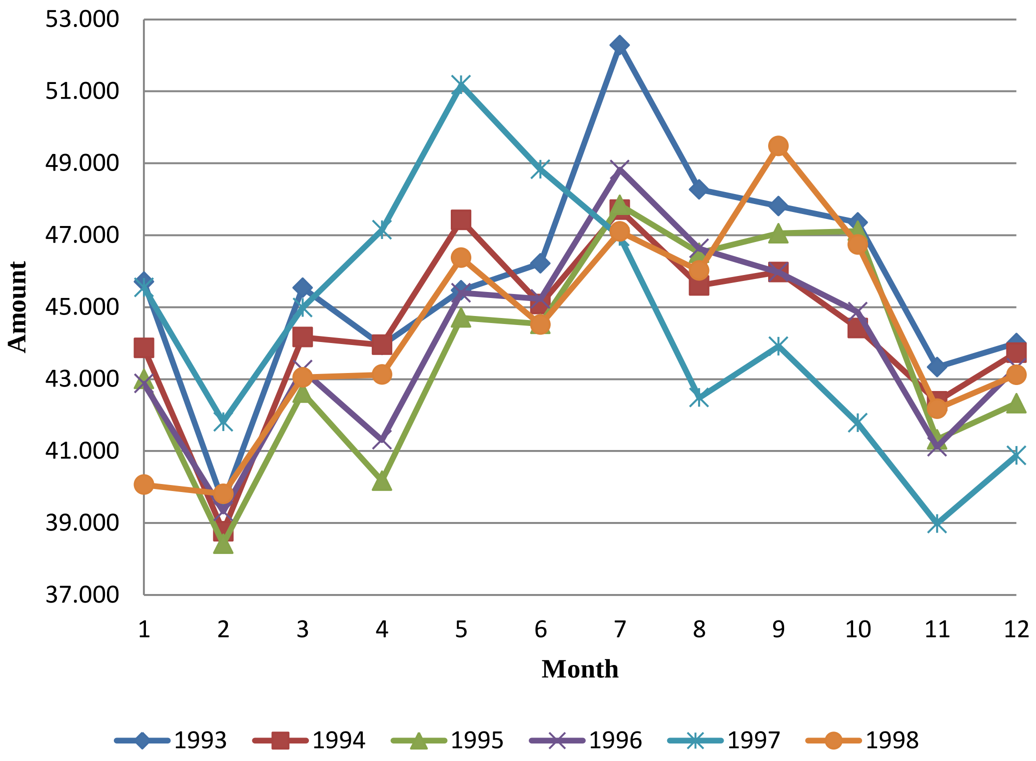
Table B.1 Amount of births in the Italian general population from 1993 to 1998.

	1993	1994	1995	1996	1997	1998	Average counts	Average rate
Period	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
January	45,704	43,859	42,996	42,872	45,549	40,063	43,507.17	0.082
February	39,578	38,767	38,420	39,318	41,809	39,807	39,616.50	0.074
March	45,536	44,163	42,618	43,258	44,982	43,047	43,934.00	0.082
April	43,942	43,948	40,172	41,306	47,143	43,124	43,272.50	0.081
May	45,465	47,419	44,708	45,398	51,174	46,371	46,755.83	0.088
June	46,215	45,082	44,533	45,228	48,824	44,510	45,732.00	0.086
July	52,284	47,703	47,839	48,822	46,955	47,099	48,450.33	0.091
August	48,270	45,599	46,507	46,630	42,482	46,014	45,917.00	0.086
September	47,807	45,976	47,050	45,971	43,910	49,480	46,699.00	0.087
October	47,353	44,419	47,118	44,866	41,778	46,737	45,378.50	0.085
November	43,330	42,382	41,321	41,118	38,981	42,176	41,551.33	0.078
December	44,000	43,733	42,327	43,316	40,874	43,120	42,895.00	0.080
Year	549,484	533,050	525,609	528,103	534,461	531,548	533,709.17	100

Note: Column (1)-(6) report the amount of births in the Italian general population per month and per year, from 1993 to 1998. Column (7) reports the average amount of births per month and per year, from 1993 to 1998. Column (8) reports the average monthly birth rate (month average / year average).

Figure B.1 plots the amount of births per month, by year, for the general Italian population.

Figure B.1



The bulk of the births is concentrated in summer and beginning of fall, while the lowest levels of births are reached in winter.

Appendix C

Main analysis - winsorization at 3% and 97%.

Table C.1 Wage gap between relatively old and young players. Winsorization at 3% and 97%.

Variables	(1)	(2)	(3)	(4)
Age	0.170***	0.168***	0.168***	0.149***
	(0.029)	(0.029)	(0.019)	(0.045)
Age^2	-0.006***	-0.005***	-0.006***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.002)
Q2 (April-June)	-0.061	-0.060	-0.056	-
	(0.090)	(0.090)	(0.051)	
Q3 (Jule-September)	0.179*	0.178*	0.052	-
	(0.108)	(0.108)	(0.065)	
Q4 (October-December)	-0.162	-0.165	-0.131**	-
	(0.103)	(0.103)	(0.063)	
December	-	-	-	-0.176**
				(0.078)
Constant	-0.953***	-0.855***	-1.028***	-1.025***
	(0.163)	(0.174)	(0.166)	(0.356)
F-test, quarters of birth (p-value)	0.026	0.026	0.043	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.112	0.123	0.566	0.700
Observations	1,598	1,598	1,598	282

Note: Standard errors clustered on footballers are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The samples for the analyses use less than 1,704 observations because of missing values on wage. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects; column (4) reports the results obtained with the discontinuity sample strategy, i.e., only players born in January or December, they are obtained with OLS including players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

Column (3), with season and players' team fixed-effects, returns estimates similar to those displayed in Table 3, whereas the estimate for the December coefficient from column (4) is almost halved.

Analysis by age group - winsorization at 3% and 97%.

Table C.2 Wage gap between relatively old and young players, by age group. Winsorization at 3% and 97%.

	< 21	21-25	26-30	> 30
Variables	(1)	(2)	(3)	(4)
Age	0.480*	0.173	0.217	0.280**
	(0.270)	(0.155)	(0.228)	(0.128)
Age^2	-0.070	-0.006	-0.008	-0.010***
	(0.075)	(0.012)	(0.010)	(0.004)
Q2 (April-June)	-0.352	-0.073	-0.036	-0.091
	(0.344)	(0.079)	(0.068)	(0.084)
Q3 (Jule-September)	-0.575*	0.014	0.072	0.092
	(0.338)	(0.103)	(0.090)	(0.106)
Q4 (October-December)	0.116	-0.190*	-0.221**	-0.121
	(0.235)	(0.103)	(0.0880)	(0.101)
Constant	-1.257**	-0.986**	-1.430	-1.593
	(0.559)	(0.496)	(1.230)	(1.128)
F-test, quarters of birth (p-value)	0.263	0.229	0.019	0.197
Season F.E.	Y	Y	Y	Y
Team F.E.	Y	Y	Y	Y
R-square	0.889	0.502	0.556	0.640
Observations	51	403	629	515

Note: Standard errors clustered on players are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The samples for the analyses use less than 1,704 observations because of missing values on wage. Column (1) reports the results from the OLS with players' team and season fixed-effects for players younger than 21 years of age; column (2) reports the results from the OLS with players' team and season fixed-effects for players between 21 and 25 years of age; column (3) reports the results from the OLS with players' team and season fixed-effects for players between 26 and 30 years of age; column (4) reports the results from the OLS with players' team and season fixed-effects for players older than 30 years of age. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

The estimates are similar to those displayed in Table 4.

Appendix D

In this appendix the main analysis is repeated on different sub-samples, based on players' role. Unfortunately, the sample sizes for the analyses by players' role mostly lead to inconclusive results, except for the subsample of midfielders. Therefore, any interpretation of the estimates in Table D.1, D.2, D.3, and D.4 should be considered carefully.

The analysis on the goalkeepers' sub-sample is reported below in Table D.1.

Table D.1 Wage gap between relatively old and young players, only goalkeepers.

Variables	(1)	(2)	(3)	(4)
Age	0.286***	0.286***	0.201***	0.347***
	(0.089)	(0.095)	(0.063)	(0.099)
Age^2	-0.010***	-0.0098***	-0.007***	-0.0125***
	(0.003)	(0.003)	(0.002)	(0.003)
Q2 (April-June)	-0.527	-0.513	-0.295**	-
	(0.387)	(0.419)	(0.122)	
Q3 (Jule-September)	-0.042	-0.038	0.032	-
	(0.345)	(0.351)	(0.147)	
Q4 (October-December)	-0.571	-0.556	-0.389	-
	(0.361)	(0.368)	(0.339)	
December	-	-	-	0.187
				(0.320)
Constant	-1.706***	-1.600**	-1.623**	-2.461***
	(0.483)	(0.601)	(0.616)	(0.176)
F-test, quarters of birth (p-value)	0.181	0.221	0.048	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.208	0.219	0.729	0.864
Observations	163	163	163	33

Note: Standard errors clustered on footballers are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. These analyses are run on the sub-sample of Italian goalkeepers, 175; the missing values on wage are such that the analyzed sub-sample is smaller. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects; column (4) reports the results obtained with the discontinuity sample strategy, i.e., only players born in January or December, they are obtained with OLS including players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

The estimates for quarters of birth are rarely statistically significant. The analysis on the goalkeepers' sub-sample seems to suggest the presence of RAE and, as expected, that the RAE is stronger for goalkeepers; the literature suggests this is due to the importance of physical characteristics in this role (Salinero et al., 2013). However, the large and statistically significant RAE for players born in the second quarter, see column (3), is puzzling. Moreover, column (4) suggests a reverse RAE for goalkeepers born in December.

The analysis on the defenders' sub-sample is reported in below Table D.2.

Table D.2 Wage gap between relatively old and young players, only defenders.

Variables	(1)	(2)	(3)	(4)
Age	0.196**	0.195**	0.227***	0.727***
	(0.081)	(0.079)	(0.059)	(0.240)
Age^2	-0.006*	-0.006*	-0.008***	-0.028***
	(0.003)	(0.003)	(0.002)	(0.010)
Q2 (April-June)	-0.041	-0.039	-0.006	-
	(0.144)	(0.143)	(0.084)	
Q3 (Jule-September)	0.037	0.035	-0.078	-
	(0.197)	(0.197)	(0.104)	
Q4 (October-December)	-0.108	-0.113	-0.135	-
	(0.182)	(0.183)	(0.135)	
December	-	-	-	-0.233
				(0.293)
Constant	-1.266***	-1.338***	-1.632***	-5.557***
	(0.443)	(0.506)	(0.376)	(1.684)
F-test, quarters of birth (p-value)	0.896	0.888	0.675	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.139	0.155	0.586	0.817
Observations	582	582	582	83

Note: Standard errors clustered on footballers are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. These analyses are run on the sub-sample of Italian defenders, 604; the missing values on wage are such that the analyzed sub-sample is smaller. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects; column (4) reports the results obtained with the discontinuity sample strategy, i.e., only players born in January or December, they are obtained with OLS including players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

The estimates for quarters of birth are never statistically significant. The analysis on the defenders' sub-sample seems to suggest the presence of RAE. However, the estimates are smaller than those obtained in the main analysis; according to the literature (e.g., Salinero et al., 2013), they are expected to be larger in presence of RAE.

The analysis on the midfielders' sub-sample is reported below in Table D.3.

Table D.3 Wage gap between relatively old and young players, only midfielders.

Variables	(1)	(2)	(3)	(4)
Age	0.264***	0.262***	0.213***	0.198***
	(0.050)	(0.050)	(0.038)	(0.070)
Age^2	-0.010***	-0.010***	-0.007***	-0.00598*
	(0.002)	(0.002)	(0.002)	(0.003)
Q2 (April-June)	0.090	0.087	0.024	-
	(0.145)	(0.146)	(0.075)	
Q3 (Jule-September)	0.120	0.109	-0.064	-
	(0.165)	(0.165)	(0.083)	
Q4 (October-December)	-0.331**	-0.343**	-0.176**	-
	(0.133)	(0.135)	(0.079)	
December	-	-	-	-0.177**
				(0.085)
Constant	-1.485***	-1.403***	-1.438***	-1.426***
	(0.273)	(0.289)	(0.233)	(0.444)
F-test, quarters of birth (p-value)	0.009	0.008	0.074	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.170	0.183	0.634	0.804
Observations	572	572	572	126

Note: Standard errors clustered on footballers are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. These analyses are run on the sub-sample of Italian midfielders, 612; the missing values on wage are such that the analyzed sub-sample is smaller. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects; column (4) reports the results obtained with the discontinuity sample strategy, i.e., only players born in January or December, they are obtained with OLS including players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

The analysis on the midfielders' sub-sample provides statistically significant evidence of RAE, which is confirmed also by the estimate of the RAE for midfielders born in December, see column (4). The preferred model, which is that that includes seasons and team fixed-effects, provides estimates for RAE that are close to those obtained in the corresponding model in the main analysis.

The analysis on the forwards' sub-sample is reported below in Table D.4.

Table D.4 Wage gap between relatively old and young players, only forwards.

Variables	(1)	(2)	(3)	(4)
Age	0.217**	0.229**	0.234***	0.0324
	(0.0968)	(0.0975)	(0.053)	(0.089)
Age^2	-0.007	-0.007*	-0.008***	-0.003
	(0.004)	(0.004)	(0.002)	(0.004)
Q2 (April-June)	-0.092	-0.096	-0.001	-
	(0.211)	(0.214)	(0.130)	
Q3 (Jule-September)	0.319	0.327	0.212*	-
	(0.209)	(0.210)	(0.126)	
Q4 (October-December)	0.0511	0.0601	0.037	-
	(0.277)	(0.282)	(0.152)	
December	-	-	-	-
Constant	-0.992*	-0.964*	-1.297***	0.781
	(0.514)	(0.527)	(0.371)	(0.650)
F-test, quarters of birth (p-value)	0.234	0.224	0.335	-
Season F.E.	N	Y	Y	Y
Team F.E.	N	N	Y	Y
R-square	0.209	0.224	0.664	0.906
Observations	281	281	281	40

Note: Standard errors clustered on footballers are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. These analyses are run on the sub-sample of Italian forwards, 310; the missing values on wage are such that the analyzed sub-sample is smaller. Column (1) reports the results from the OLS without fixed-effects; column (2) reports the results from the OLS with season fixed-effects; column (3) reports the results from the OLS with players' team and season fixed-effects; column (4) reports the results obtained with the discontinuity sample strategy, i.e., only players born in January or December, they are obtained with OLS including players' team and season fixed-effects. "F-test, quarters of birth" gives the p-value from the F-test on the joint significance of the quarters of birth estimates.

The estimates for quarters of birth are rarely statistically significant. The analysis on the forwards' sub-sample does not seem to suggest the presence of RAE. No estimate for the RAE is displayed in column (4) since there is only 1 forward born in December for the period under consideration.