Forthcoming *Journal of Sports Economics*

Selection Bias and Peer Effects in Team Sports: the Effect of Age Grouping on Earnings of German Soccer Players

John Ashworth[a]
&
Bruno Heyndels[b,*]

[a] *Department of Economics, University of Durham, 23-6 Old Elvet, Durham, DH1 3HY, England; e-mail: John.Ashworth@durham.ac.uk*

[b] *Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium; e-mail: Bruno.Heyndels@vub.ac.be*

Abstract - This paper analyses how age grouping in youth competitions and soccer education programs affect wage formation at the professional level. A simple theoretical model shows that professional players born late after the cut-off are expected to earn systematically higher wages than their early-born peers. Two discriminating factors are responsible for this: a systematic bias in the talent detection system and peer effects in the production process of human (sports) capital. We demonstrate the existence of such an effect among (native) German professional soccer players. Estimating an earnings function for players in seasons 1997-1998 and 1998-1999 we find clear evidence of a month-of-birth related wage bias. Players born late after the cut-off date earn systematically higher wages, though this effect is not discernible in all positions; it is strongest for goalkeepers and defenders but not evident for forwards.

---

* We wish to thank Erik Lehmann and Jürgen Weigand for providing us the data and information on German soccer. Earlier versions of this paper were presented at the 3rd international conference of the Panhellenic Association of Sports Economists & Managers (PASEM), Athens, and at the 2003 European Sport Management Congress, Stockholm, Sweden. We wish to thank participants and Jochen Musch for helpful suggestions. The authors also thank two anonymous referees for their perceptive comments that have helped the exposition of the paper; the usual caveat applies.

In soccer – as in other sports or at school – children are grouped according to age. Such age grouping brings together children born in different months of the year. This paper analyses how these *relative* age (RA) differences affect human capital returns to soccer education. The basic intuition is that age grouping leads to systematic differences in the treatment of players depending on the month in which they are born. Such differences then translate into productivity and wage differences at adult age.

We identify two discriminating factors that are introduced by age grouping of younger players: *selection bias* in the talent scouting system and *peer effects*. The former implies that early-born children (born soon after cut-off date) are more likely to be identified as being talented at younger ages. In a world where there is a shortage of high quality soccer education programs, the selection bias implies that early-born children are more likely to receive the better education. The latter – peer effects – imply that late born children benefit from being educated together with early born peers who are at a later stage in their individual developmental process.

Both discriminating factors have been discussed in the literature. The problem of selection bias as a function of children's month-of-birth has been analysed extensively in the sports sciences literature on the so-called Relative Age Effects (RAE) (Musch and Grondin, 2001, gives a review). This RAE refers to the overrepresentation in sports teams of players born in certain months of the year. Peer effects in school education have been analysed in the economics of education (see, for example, Hoxby, 2003). Nonetheless, the focus there is not explicitly on the month-of-birth as a determining factor of peer effects. Moreover, to our knowledge peer effects have not yet been investigated in the context of *sports* education. This paper not only integrates two related strands of the literature. More importantly, our paper is the first to explore and test their consequences for wage formation.

The first section presents a simple graphical analysis of the role of age grouping in soccer education. Two predictions follow from our model. We expect (a) a RAE, discriminating *against* late-born players and (b) a wage bias, discriminating *in favour* of a subset of the same players. The model predictions are tested in the second section using data on the birth distribution and wages of (native) German professional soccer players (Bundesliga seasons 1997-1998 and 1998-1999).[1] Confirming earlier studies, we find clear evidence of a RAE. Crucially, estimation of a Mincerian earnings function gives support for the prediction that

1 The choice of these seasons was driven by the fact that recently cut-off dates in German soccer were changed from August 1st to January 1st. By investigating the 1997-1999 period we are certain to have a set of players that have been educated under the same cut-off date.

late-born professional soccer players have systematically higher earnings than their early born peers. A summary and discussion are in the final section.

AGE GROUPING IN SOCCER EDUCATION

From an economic perspective, sports education can be seen as a means to develop one's soccer skills. Education is then a complex production process that transforms players' intrinsic talents into their actual playing qualities through training and competition. More and higher quality training and competitions during childhood increase the player's actual playing quality at adult age. This in turn determines his marginal productivity (MP) in a soccer team's performance and thus his wage.

The skill level (marginal productivity) for a player at any point in time reflects his progress just because he gains height, weight, strength, co-ordination, … but also as a consequence of sport specific training. Age grouping implies that players born in the same year are grouped. Cut-off dates are determined (in soccer the typical cut-off date is January or August 1st) and players born within the next year are considered as one age group. They play in the same teams, competitions, … As the individual player's MP evolves over time, age grouping implies that players who are at different stages of their development are grouped. Consider two equally talented players, E and L. The former (E) is born early after cut-off date – which for expository reasons we will assume to be January 1st ; the latter (L) is born in December of the same year. Their soccer productivity follows an identical pattern with an 11-month lag.

It is clear that – while both players have identical talents and soccer education - at any time t, player E is superior to player L. This has a number of possible consequences for each of them, which have been discussed in detail in the literature on the RAE.[2] An important consequence is that early born players are more likely to be selected for top teams at young age. A clear example of this is reported in Helsen, Starkes and Van Winckel (2000) who analyse the RAE for young players who were transferred to Belgian first division youth teams. Considering the youngest players (age 6-8), they find that of those transferred 45.5 % were born in the first quarter following the cut-off date whereas only 6.8 % were born in the last quarter. Similar results were found for older age groups. Because training facilities, training intensity and the level of competition is generally higher (i.e. better) in top clubs the result is that the early born

[2] Consequences at the individual level can be psychological (E-born players may gain self-confidence, L-born players may be discouraged) or physical (E-born players are stronger and may therefore be

players are more likely to receive a more effective soccer education. This will be reflected in higher soccer productivity.

To see what are the implications of age grouping for wage formation at the professional level, two factors have to be taken into account. First, it should be recognised that the set of professional players is a subset of a much larger set of soccer players in general. Second, the precise nature of soccer education programs should be considered.

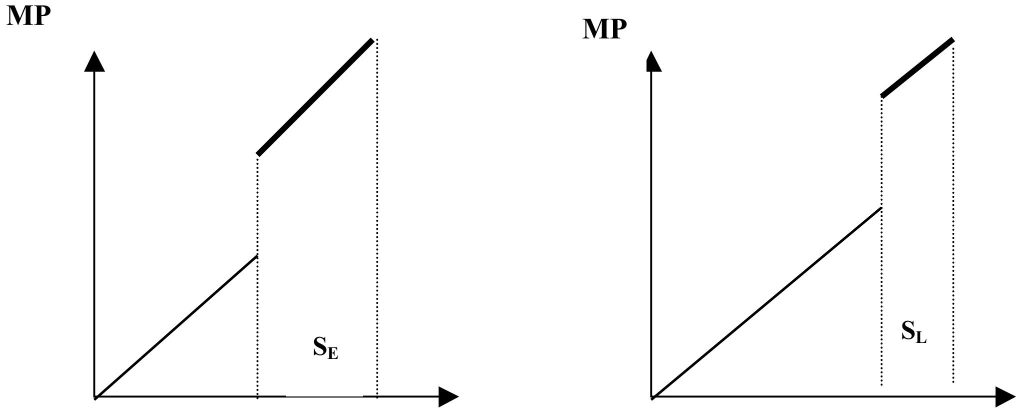
- *Professional players are a subset*: Consider a model in which two types of soccer education are provided: *h* (high quality) and *l* (low quality). This quality is – from the perspective of an individual player – not a choice variable. Some clubs provide *h*, others *l* (for example: first division clubs provide *h*, other teams provide *l*). The clubs autonomously select those players who they think are sufficiently talented to receive *h*-education. If we consider two players of identical quality (E and L in the previous example), the RAE tells us that the E-player will be more likely to receive *h*-education. But it is, of course, possible that both E and L are talented enough to be selected to receive the *h*- education (formally, this means that at the time the selection is made both have attained a soccer productivity level above some selection-threshold).

  To see how a RAE for younger players and other month-of-birth effects translate into professional players' productivity (and thus, their wages) it is important to recognise that the set of professional players is a subset of the set of players having received the "good" *h*-education. The analysis has to shift focus from the individual player to the group of all players.

  We consider the players at adult age (when they have reached their highest MP). Figures 1.a and 1.b give these productivity levels for E-born and L-born players respectively. In the figures players are ranked according to their MP-level reached at adult age. For simplicity it is assumed that the number of E- and L-born players in the population is equal ($T_E = T_L$). We assume that the distribution of players' intrinsic skills is uniform and identical over the months. This means that if all players had received the same level of education, then the skill distribution for each month-of-birth corresponds to a straight line through the origin (see figures 1.a and 1.b). Nonetheless, not all receive the same soccer education. Some are selected at young age and receive *h*-quality education. Players in these subsets (S_E and S_L respectively) reach a higher skill level at adult age than they

---

chosen for specific positions, for example, being a goalkeeper in Ice Hockey as this involves carrying would have had if they had not received *h*-education. The bold lines in figures 1.a and 1.b reflect this. The RAE among young players suggests that the number of players selected in this way is higher for E-born players. In figure 1, this can be seen as $S_E > S_L$.



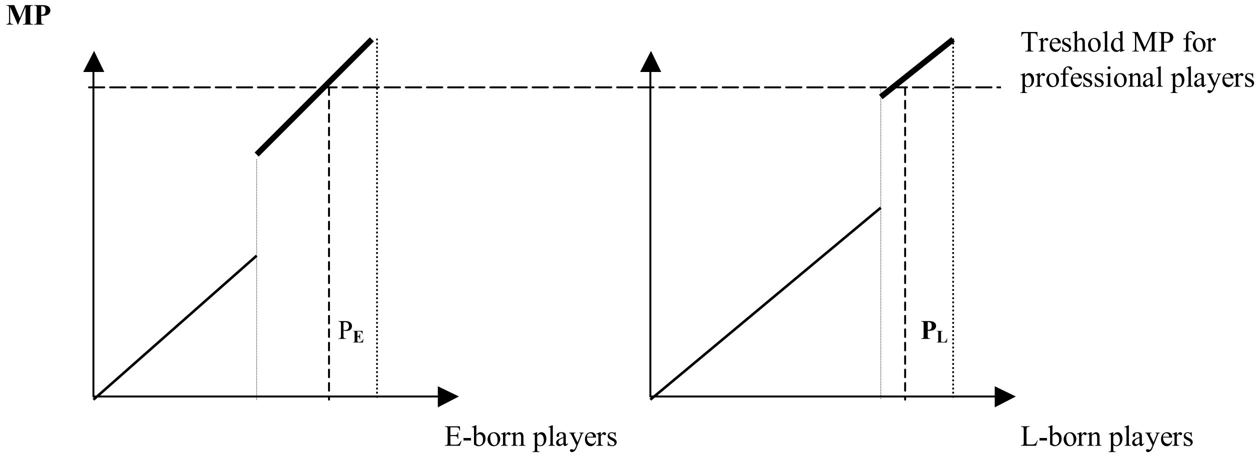
Figures 1. [illegible] ***l L-born playe*** [illegible]

Given the above, what are no [illegible] age formation at the professional level? If there are just two types of education (*h* and *l*) and if *all* players that receive *h*-quality soccer education (S_E+S_L) become professional players, it is clear from figures 1 that the average wage will be higher among L-born players.

However, only a subset of players having received *h*-education will become professional players. The size of this subset – P* - will depend on the need of the professional teams (i.e. the size of the market). Graphically, the selection corresponds with finding a horizontal line through figures 2, so that $P^* = P_E + P_L$, where, P_E and P_L are the subsets of E- and L-players that are selected to become professional soccer player.

In figure 2, it is clear that the average wage of E- and L-players depend on the size of the set of professional players P* relative to the size of the smallest subset of players that received *h*-quality soccer education (in figure 1 this is the subset of L-players).

heavy equipment (see Grondin and Trudeau, 1991).



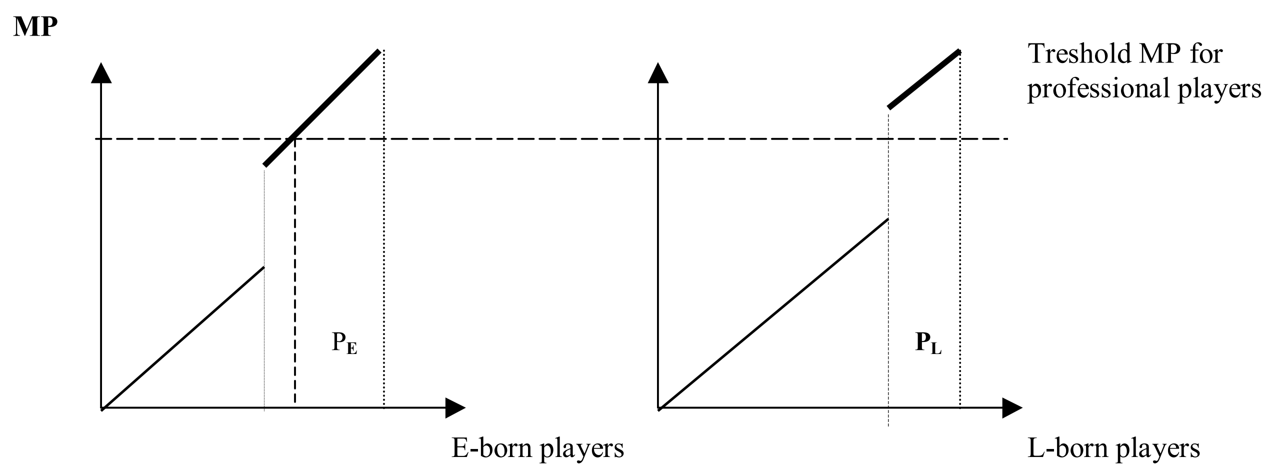
Figures 2.a & 2.b: The subset of professional players $P^* = P_E + P_L$

If the number of players that is selected for a professional career is smaller than twice[3] the size of the smallest subset (in the present case: S_L), then all players that are selected for a professional career have received the same education *h*. This is the situation as depicted in figures 2.a and 2.b. Overall, the expectation must be that average wages of professional E- and L-born players (w_E and w_L respectively) are the same. At the same time, we do not expect a RAE to occur among professional soccer players (despite a RAE among young players): the subsets of E- and L-born professional players are of equal size ($P_E = P_L$). So:

$$P^* < 2.S_L \quad \Rightarrow \quad P_E = P_L \text{ (no RAE)}$$
$$\Rightarrow \quad w_E = w_L$$

If, however, more players are selected for a professional career than this, the threshold line in figures 2 moves downwards. It can easily be seen that if this threshold line moves down sufficiently, the average MP's (and thus wages) for E- and L-born professionals will differ. This is shown in figure 3.

[3] More generally: if we consider n subsets of players, then the condition is that the number or players that is selected is smaller than n times the size of the smallest subset.



Figures 3.a & 3.b: The subset of professional players $P^* = P_E + P_L$ with $P_E > P_L$

In a setting like figure 3 it is clear the marginal player getting a professional contract *after* the $2.S_L$-th is an E-born. As such, a RAE occurs. Moreover, as this E-born player has marginal productivity below that of any of the selected L-born players, the expected average wage for professional E-born players is lower than that for L-born players.

$$P^* > 2.S_L \quad \Rightarrow \quad P_E > P_L \text{ (RAE)}$$
$$\Rightarrow \quad w_E < w_L$$

Thus, depending on the number of players that can enter a professional career, the wage of late-born players is expected to be higher or equal to that of early-born players (note that this argument can be generalised to a situation where we allow for a continuum of education qualities). It should be noted that the occurrence of a wage bias goes hand in hand with the occurrence of a RAE.[4]

- *Peer effects*: In much of the literature on the RAE it is argued that top teams offer higher training effectiveness. They allow a better development of the individual player's skills.

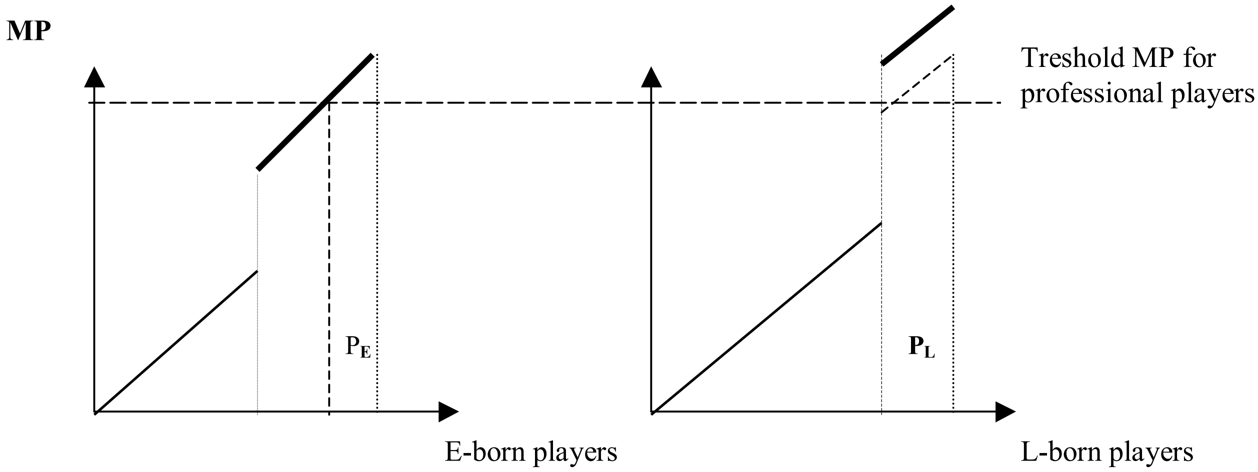
[4] As mentioned, the expected presence of a RAE and of wage differences according to date of birth depend on the 'size of the market' (in terms of the amount of players demanded by professional teams). In practical terms, this means that an institutional change like the Bosman rule - by opening the League to talent from around the world - may have led to a higher threshold for native German players (lower value of P*). The empirical implication to be expected is thus that, ceteris paribus, the Bosman rule will have lowered the probability of an RAE (and of the related wage differentials) occurring. This effect might be (partially) offset by the possible additional outflow of German players. This may actually make the RAE and wage differences more, instead of less, likely. We thank an anonymous referee for pointing this out.

This assumption underlies our analysis in figures 1, 2 and 3 and is not debated in the paper. Nonetheless, it is important to make explicit what is meant by "more effective soccer education". What makes a given education better than another? Soccer education is a human capital production process with many inputs: training accommodation, training frequency, training quality, level of competition, coach quality, and so on. So, ceteris paribus, better accommodation, a higher training frequency, … lead to more effective education.

A crucial characteristic of soccer education programs (including soccer competitions) is that players are inputs. We expect inputs to be complementary: training with better players and/or playing against better opponents is assumed to increase the effectiveness of soccer education. In other words: peer effects are expected to occur. Such effects have been investigated in the economics of education.[5] As Hoxby (2003) in her survey concludes: "As a rule, the evidence suggests that peer effects exist, but their forms (….) are barely understood as yet" (Hoxby, 2003, 4).

For the present analysis, it is crucial to see that these peer effects may lead to an additional month-of-birth effect over and above the RAE. To see this, consider two players of identical quality (E and L from figure 1). Then, if they are educated in the same squad, the late-born player L trains with the early-born E and vice versa. For L this is an advantage, as he gets the opportunity to train with players who are (at a given point in time) better than him. For E, however, this is a disadvantage as he receives his education with players of inferior productivity. The result of this is that L receives more effective soccer education and thus that – despite their identical intrinsic qualities – he reaches a higher productivity at adult age.

[5] For a recent application, see Lefgren (2004). While peer effects as such have – to our knowledge – not been analysed in the context of sports education programs, it should be observed that the implicit complementarity of inputs is formally similar to the complementarity of inputs in the production of sports performances as investigated in Idson and Kahane (2000).



Figures 4.a & 4.b: Peer Effects

An immediate implication of this is that the soccer productivity of L-born players will be higher for *h*-quality soccer education. This is depicted in figure 4.b by the upward shift of the soccer-productivity function for those S_L L-born players who received *h*-quality training.[6] It should be noted that the month-of-birth induced wage bias is now – in contrast to the analyses in figures 1 to 3 – not unambiguously connected to the presence of a RAE among professional players. The wage bias again depends on the size of the set of selected professional players (P*). If the market for professional players is so small that only a very small set of players can enter, then these will predominantly be the late born players (who will be the most productive, and therefore the highest paid).

$$P^* < S_L \quad \Rightarrow \quad P_E < P_L \text{ (“inverse” RAE)}$$
$$\Rightarrow \quad w_E < w_L$$

If more players enter a professional career, then (assuming – as in figure 4 - that the best L-born players with *l*- type education reaches a lower productivity level than any of the E-born players with *h*-type education) the RAE will occur if $P^* > 2.S_L$. The situation is identical to the one in the previous subsection:

$$P^* > 2.S_L \quad \Rightarrow \quad P_E > P_L \text{ (RAE)}$$

[6] For simplicity, we have only drawn the upward shift for those players who received *h*-quality education. It is likely that a similar shift appears for those L-born players receiving *l*-quality education. Still, for this latter shift to be empirically relevant in the context of the present analysis it should be so strong as to ‘wipe out’ the quality difference in soccer education for those L-born players who are

$$\Rightarrow \qquad w_E < w_L$$

In summary, our analysis suggests two testable hypotheses:

H1: Late-born players are underrepresented at the professional level (RAE).

H2: Late-born professional players have higher wages[7].

Discussion

It should be recognised that in reality age grouping is not always applied strictly. In many leagues – like in Germany – young players are selected for squads with older players. For example, talented 10 year olds compete with, say, 11 or 12 year olds. This means that the talented 10 year olds can benefit from an advantage similar to that of the L-born player in our theoretical example. With respect to the empirical work in the next section, it should be noted that our argument does remain valid, however, to the extent that this "crossing" of age categories is not the general rule. Casual observation suggests it is not. Two additional remarks are in place. First, the very fact that it does occur that talented young players are selected for teams of older children can be seen as evidence that the input complementarity discussed in the text is actually recognised to be relevant by practitioners. A second remark is of a theoretical nature. The complementarity issue can explain the occurrence of RAE in a model of rational and fully informed talent scouts. In much of the literature on RAEs, the biases in the talent scouts' selection behaviour is presented as an 'anomaly' in the sense that RAE reflects that talent scouts are incapable to abstract from players' relative age when assessing their intrinsic talents (and future qualities).[8] However, given complementarity of

(marginally) not selected to receive *h*-education. We feel such a strong effect is unrealistic. Note, however, that if it should exist then it would be possible to predict lower wages for L-born players.

[7] Note that the latter assumption can have two sources. First, it is possible that a selection bias at younger ages implies that the subset of L-born players at the professional level is of higher intrinsic quality (on average). Second, peer effects during childhood imply that L-born players receive superior soccer education. Discriminating between both sources could be possible if the wage bias is observed in the absence of a RAE, or under an "inverse" RAE where late born (instead of early born) players are actually over-represented. These situations would suggest the peer effect to be responsible (or at least dominating). Therefore – when analysing the effects on wage formation - it is crucial to first investigate whether a RAE exists or not.

[8] We assume that talent scouts are hired by professional teams and that their purpose is to select young talents to be prepared for a professional career. Of course, talent scouts may have other motives. This may be the case either because their club has other motives (for example, a club may find it important to win youth competitions) or because there is an agency problem. In that case talent scouts' own utility may be increased by short term success of the players that are attracted whereas the team's – the principal's – utility depends on long term success, i.e. on the players' success at professional age. In these cases, scouts pick the players that will win now rather than in the future. In other words: scouts pick E-born players and no inefficiency nor irrationality is involved.

inputs, a RAE may occur under the assumption of rational and fully informed talent scouts. These may be more likely to select E-born players as they recognise the typical characteristic of the soccer-education production function. For example: assume that individual talent is a discrete number between 0 and 10. Assume further that talent scouts can predict with 100 percent accuracy which young player will later become a star player. In other words, assume that talent scouts observe talent levels. Then, scouts will select a player with talent 10 with certainty. Assume, however, that these players are very rare but that there is a larger pool of players with a talent of 9. Now, the talent scout knows that for each year-of-birth he can select one or two players with a talent of 10. Their next question is then: how to "surround" these star players. Given that there is a pool of "9" players, it is a rational choice for the talent scout to select a 9-talent E-born player over a 9-talent L-born player. This is the case, as the perfectly informed talent scout knows that only the 10-talent player will make it to the highest level. Taking into account input complementarities, the scout then maximises the effectiveness of the 10-talent's soccer education by surrounding him with E-born 9-talents whose function it is to be inputs in the "star" players' soccer education.

While we have no knowledge of any existing study concerning the relation between month-of-birth and earnings in sports, such investigations exist in the literature that analyses the effects of schooling. No unambiguous conclusions with respect to RAE are reached, however. Angrist and Krueger (1991) find significant season of birth effects using month-of-birth as an instrument to measure schooling. Birth effects are then attributed to the fact that – due to the institutional organisation of the U.S. schooling system – early born children attend school less long.[9] Angrist and Krueger reject the alternative assumption of RAE type effects on the basis of a number of observations. For example, they find no relationship between earnings and season of birth for college graduates (who are not constrained by compulsory schooling). Bound and Jaeger (1996) and Plug (2001), however, do find a RAE in earnings: being born early after cut-off date leads to systematically higher earnings that can not be attributed to differences in duration of school careers. It should be observed that in these studies early-born people have higher earnings. This contrasts to our own hypothesis. Of course, as discussed in Musch (2002) the context of school education is fundamentally different from the soccer context. A major difference is the fact that unlike sports activities, schooling is compulsory. This makes drop out (as an often mentioned reaction to relative underperformance at young age) less of an option. Also other relevant dimensions that may affect the likelihood of RAE to occur differ systematically. For example, the 'competition for

[9] Children born in a given year start school at the same time, still schooling is compulsory until the day that the child reaches a given age. As a result, early born children are older when they start school but leave school at the same age as late born children (and so have less schooling).

places' may be stronger and the role of children's 'physical development' will be more important in the context of sports. Both factors may contribute to the emergence of RAE (Musch, 2002). Psychological factors (for example children with higher perceived competence showing more intrinsic motivational orientation) will be relevant in both the schooling and the sports context. Russell and Startup (1986) find results on schooling that are in line with those presented here. Investigating month-of-birth effects on academic achievement, they demonstrate a RAE among British students: (proportionately) more early-born students graduate and go to college. But, they also find that the subset of early-born students gets lower grades on average! This observation is, however, explained ex post and rather vaguely: the authors hypothesise that the degree-effect may reflect the fact that "some quality needed for success in examinations increases as a function of age up to around 19 years of age, and then declines" (Russell and Startup, 1986, 845). We think that this explanation is not very helpful in understanding the RAE in the context of college education. Indeed, we (might) suggest an alternative explanation based on differences in the effectiveness of schooling much like the one discussed for soccer players.[10]

One final remark is needed. When we expect that earnings will be biased according to players' ages, this does not mean that we expect wage formation to be inefficient! The earnings bias does possibly reflect inefficiencies in the selection system at younger ages or, possibly, in the composition of squads of players at young age. It's these inefficiencies that we expect to have lasting effects in the sense that their consequences remain in existence until players reach the professional level at adult age.

EMPIRICAL ANALYSIS

A Relative Age Effect among German soccer players

The presence of a RAE has reached the status of a stylised fact in the sports literature. A RAE has been observed among the participants to the 1990 World Cup (Barnsley, Thompson and Legault, 1992). Country specific evidence for an RAE at the professional level is found for the Netherlands and England (Dudink, 1994), Belgium (Verhulst, 1992; Helsen, Starkes and Van Winckel, 1998), the Netherlands and France (Verhulst, 1992), Australia, Brazil,

[10] The Russell and Startup interpretation draws attention to an issue that is of immediate relevance in our own context. Indeed, while being L-born – and thus: younger – may be a "disadvantage" for players who have not reached their maximum productivity level, being young is likely to be an advantage in the later phase of their career. This aspect has to be taken into account in the empirical section (by specifying the players' age in months).

Germany and Japan (Musch and Hay, 1999) and Ireland (O'Loughlin, 2002). In all these countries, players born shortly after the cut-off date (which differs among countries) are strongly over-represented.[11] The presence of Relative Age Effects among adult professional sportsmen is also observed in several other sports (for a review, see Musch, 2002 and O'Loughlin, 2002). While most of the evidence relates to team sports, evidence also exists on RAE in individual sports like tennis and swimming.

Musch and Hay (1999) document the presence of a RAE for German soccer. They analyse birthdays of 355 native German players in the 1995-96 season of the Bundesliga. The cut-off date in Germany is August 1st. We replicate their analysis using data on 285 players in the Bundesliga season 1998-99: the shaded area in figure 5 gives the birth rate (per day) of the general population born in Germany over the period 1970-1980. As can be observed, this birth rate is more or less uniform over the year (with a lower birth rate in the last months of the year).

[11] O'Loughlin (2002) finds a strong RAE for the Irish First Division but not for the Premier Division. Bäumler (1998) shows that there would be an equalisation among German professional player when they grow older. Whereas 68% of the players aged between 18-22 were born in the first half of the year, this percentage falls monotonically and reaches an even proportion of 49 % among the 33-35 years old. This effect does not show up as a significant factor in our empirical analysis. We have looked at this issue from several angles. First, we have split our sample into 2 almost equal sized groups of 'older' and 'younger' players. We found that in each group about 60 % of the players were born 'early', that is between august and January). We have found a RAE to be present in both groups though the Spearman Rank-order correlation is marginally stronger for the younger players (-0,71 compared to –0,57). Finally we re-estimated the wage function (see next section) now including interaction terms of relative age (RA) with the age to test if the RA effect changed with age. These terms are found to be insignificant.

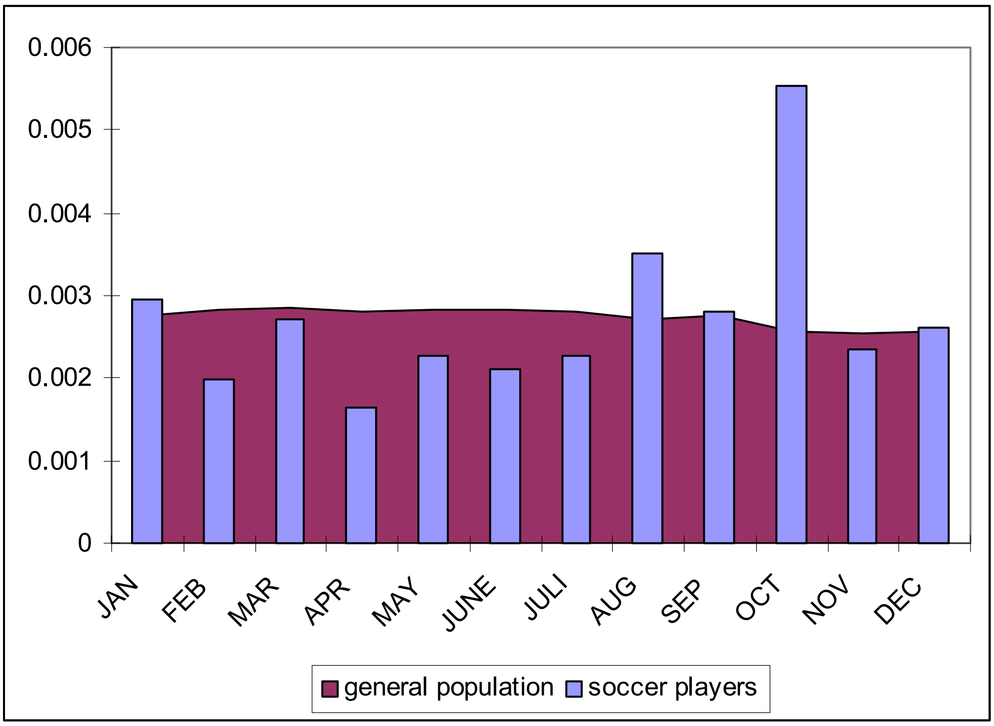


Figure 5: Birth rate by month in general population and among soccer players

If no RAE were present, then we would expect the month-of-birth distribution for Bundesliga players to conform to the shaded area. Actual birth rates for the players are given by the bars in figure 5. It is clear that notable differences occur compared to the birth rates in the general population. More precisely, we observe 5 months in which the soccer players are over-represented. This is most clear for October: while 0.26 % of the German population is born on an average October-day, 0.55 % of Bundesliga players is born on such a day. Put differently: no less than 17.19 % of all players were born in October compared to 7.99 % for the general population. We observe in figure 1 also an over-representation of players born August, September, December and January. Given that the relevant cut-off date for the Bundesliga players in 1998-1999 was August 1st, these results suggest the presence of a RAE.

Table 1 here

The visual inspection is confirmed by formal testing. We calculate a Spearman rank correlation coefficient between rank of the month of the competition year (August=1, September=2,…., July=12) and the rank of the month according to differences between expected and observed number of players. For each month we compute the expected number of players (as expected from the birth distribution of the general population) and compared this to the actual number of players. The results are summarised in table 1. For example, using the numbers already given, in the absence of a RAE we expect 7.99 % of all 285 soccer players to be born in October. This corresponds with “22.8” players. The actual number of players born in October in our sample was no less than 49. The difference between actual and expected number of players for October is thus -26.2, the lowest value over all months. October is thus assigned a rank 12, indicating that players born in this month are most over-represented in the Bundesliga. Similarly, August, January and so on receive ranks 11, 10, … On the other hand, players born in April, February and June are most underrepresented. In more general terms, table 1 shows that 5 of the last 6 months of the year (beginning in August) are among the 6 months with the highest under-representation. This suggests the presence of a RAE. A RAE is identified formally as the Spearman rank correlation takes a value -.74 ($p < .01$). Thus, the evidence on a RAE among German soccer players in Musch and Hay (1999) is confirmed.

A month-of-birth related wage bias among German soccer players

Lehmann and Weigand (1999) were the first to estimate an earnings function for soccer players (more recently, Lucifora and Simmons, 2003, analyse Italian players' earnings). It is their set of data – referring to wages of all professional soccer players in the Bundesliga seasons 1997-1998 and 1998-1999 – which we use in our own investigation. These data are estimated gross wages (including boni, income from commercials etc.) as published in the “Kicker Sportmagazines” and related sources (see Lehmann and Weigand, 1999). For our own analysis we are interested in how wages differ according to month of birth. This information is summarised in figure 6 that gives, for each month of birth, the mean income in our sample[12].

Taking into consideration that the cut-off date in German soccer in the period under consideration, that is, before the season 1998-99 was August 1st, a general pattern emerges from the data (see figure 6). Player incomes are higher for players born in the first half of the year or in December. The observation that the low incomes are earned by August-born players is fully in line with our hypothesis. So are the observed (relatively) low incomes in the successive months (September, October and November). However, although the July-born players are not very high-income earners either, players born between December and June earn the highest incomes. This is, again, generally in line with the model predictions above.

[12] Sample statistics of the data used in the analysis is given in the Appendix, Table A1.

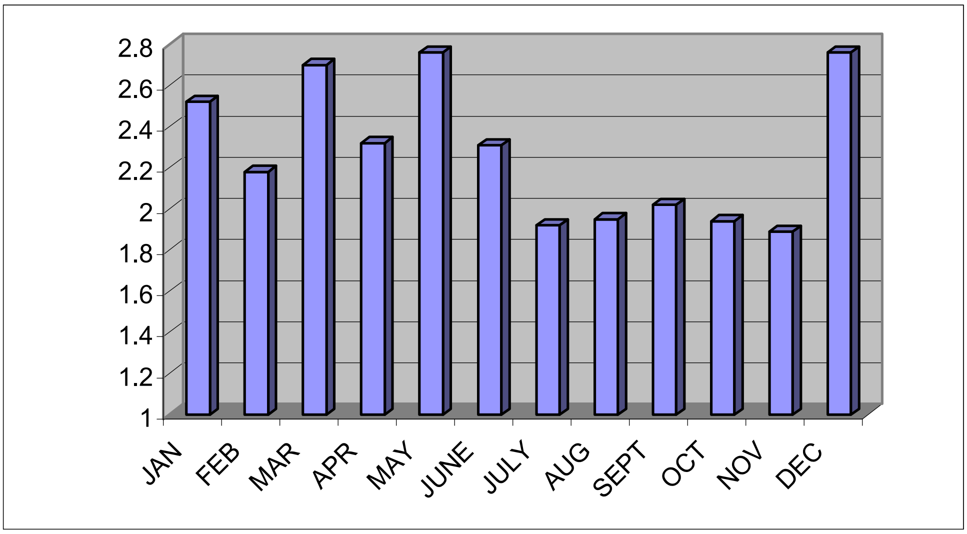


Figure 6: Average Yearly earnings (in millions of DM) of soccer players by month-of-birth

To reiterate, as discussed above, two main sources for month-related wage differences can be identified. First, late-born players may have higher wages as they represent a more selective subset (that 'survived' the scouting system that discriminates against them). Second, higher wages can result because late-born players have benefited from systematically superior soccer education.

Whilst the general pattern in figure 6 is supportive of our wage bias hypothesis, to analyse the RAE on soccer players' incomes more systematically, an earnings function is estimated taking into account other determinants of Bundesliga revenues. Our sample consists of native German players' wages for the seasons 1997-1998 and 1998-1999. We have 481 observations (for 196 players we have information on both seasons, for 89 only for one season due to lack of data – often because those players were not in the Bundesliga in one of the two seasons).

For our empirical analysis, we estimate a Mincerian human capital earnings function (Card, 1999). Typically, such a function relates the (log of) earnings to the individual's level of education and experience. Lacking information on the duration of each player's education we assume for now that this duration is independent of his relative age**.** The estimating equation, which is built up in stages, is

$$\ln(\text{WAGE}) = \alpha^{T}_{0} + \alpha_1\,\text{AGE} + \alpha_2\,\text{AGE}^2 + \alpha_3\,\text{EXPERIENCE} + \alpha_4\,\text{EXPERIENCE}^2 + \alpha_{5j}.\,\text{POSITIONj} + \alpha_6.\text{RA} + \alpha_{7j}.\,\text{POSITIONj*RA} + \varepsilon$$

where WAGE is player *i*'s income in the seasons 1997-1998 and 1998-1999 respectively (player and year subscripts have been left out for convenience), AGE is the player's age expressed in months, EXPERIENCE is the number of games played in the Bundesliga, POSITION is a three-element vector with dummies for the player's position in the field (goalkeeper, defender, midfield), using forwards as a point of reference. Finally RA captures the player's relative age corresponding with the month in which the player is born starting from the cut-off date of August 1st (so RA=1 for August-born players, 2 for September-born players, and so on). The α's are the parameters to be estimated and ε is the random error term. Predicted signs for the coefficients are: α_1 and α_2 positive and negative respectively as we expect revenues to increase at a decreasing rate and we anticipate the same effect to show up in α_3 and α_4.[13] Using forwards as a point of reference we expect the α_{5j}'s to be negative. Our main interest lies in α_6. This is expected to be positive, reflecting the positive wage bias for late-born players. The intercept must be allowed to vary to accommodate the fact that soccer is a team game and so productivity (and thereby wages) are dependent on other players in the team. In addition, the coach and local market conditions are important. The use of a dummy, α^T_0, for the club is a "catch-all" to capture the quality of team mates, the coach and the size of the club with respect to attendance; their position in the league and other factors that allow differential team salaries to be paid. The final term (POSITION*RA) allows for differential month-of-birth effects with respect to positions. Such differential effects may occur for two reasons. First, it may be the case that different positions demand different qualities and that some qualities are more easily detected at young age. For instance, if the talent to score goals ('tor-instinct') is "natural", it may be the case that – even at young age – talent scouts can observe whether a player "has it" or not. If that is the case, then the RAE among forwards will be smaller as will be the wage differentials. A second reason to expect different effects according to position is that the role of complementarities differs. In the extreme case, some qualities are "natural" and cannot be learned. The goal-scoring potential may once again be an example. If scoring cannot be learned, it is self-evident that complementarities will not exist in contrast to what is the case for those qualities that can be learned. In terms of our discussion in section II we expect significant values for the α_{7j}'s for those positions for which the relevant skills are acquired through the soccer education that a player receives.[14]

[13] Given the short duration of a typical soccer player's career, Lucifora and Simmons (2003) expect the negative AGE-effect to eventually dominate reflecting that earnings are expected to fall as players experience declining athleticism at older age. The present estimation supports this with the typical turning point in earning occurring between the age of 27 and 28. With experience, the turning point is around 280 games.

[14] Consideration was also made of whether there was a different age and experience profile for each of the different positions by running regressions that additionally interact the age and experience

The use of age and experience together in the equations follows Lucifora and Simmons (2003) and "represents a departure from the established practice of North American studies which typically include years of experience or number of career appearances but not both, and not jointly with age". The primary reason is that as there is no North American style draft of players in Germany, entering age and experience does not *necessarily* generate multicollinearity in estimated earnings in the way that can be anticipated for North American team sports. Thus, it is an empirical question as to whether there will be separate effects that are discernible. In order to see the effect of this analysis, the estimations are repeated using only age.[15]

It should also be noted, however, that estimating an earnings function following the standard practice in the economics of sports literature would not serve our purposes. This would imply including player's performance indicators, team and market characteristics in a regression that estimates (log of) players' revenues (Kahn, 2000). "Simply" adding a measure for the players' relative ages will inevitably introduce multicollinearity into the model. Indeed, as discussed in section II, the whole idea about how month-of-birth affects the effectiveness of soccer education implies that performance indicators of individual players correlate with their month-of-birth so the estimation here is a reduced form. Indeed, adding a number of the "standard" variables, such as attendance and position of the team, to the equations presented proved to be insignificant, which tends to support the view. Also, given that the equations are well-specified (see table 2 where it can be seem that the test statistics for heteroscedasticity, functional form and normality are all below the critical values), this is taken as evidence that the omission of this type of variable is not unduly biasing the results[16]. Therefore, we proceed as follows: first, we analyse the presence of a month-of-birth related wage effect in a simple specification, just including (exogenous) age (and experience) as explanatory variable(s). Then, we extend the empirical model using results from the existing literature on RAE. More

---

variables with the position dummies. It was found that there was no significant difference between the profiles for the different positions; the closest to significance being keepers; thus ceteris paribus the basic structure of wages with respect to age and experience is the same for all German players.

15 It should be noted that experience relates only to the Bundesliga. Thus the results should be treated with caution, as some players (particularly stars) will have gained some experience in other countries. As can be seen from Table 2, whilst experience is clearly important in addition to age, it is not obligatory to establish the RA effects on wage. In addition, consideration was made of star players (identified as those earning 10m DM or more), see Lucifora and Simmons (2003). No significant effect was found; nor was an effect found if the "stars" were extended to include those with earnings above 8.5m DM.

16 As one final consideration, the major productivity variable of goals was added to the equation. This was found to be overwhelmingly insignificant. Thus the club dummies are picking up the team productivity effects and position and other variables are picking up the goals effect. It is of note that the introduction of goals lowered any forward effect, as would be anticipated, but also over parameterized the equation, as diagnostic test failed with the inclusion of this variable. Also, a test

precisely, we control in the earnings function for team characteristics and the position in the field.

Table 2 here

A summary of the results is in table 2. As can be seen, with one exception, only equations that pass the diagnostic tests are presented in the tables. Thus, for example, equations ignoring position are found to be mis-specified and therefore not reported. This mis-specification implies that in order for a "satisfactory" explanation to be given, position must be taken into account. This is primarily due to the different wage structure of goalkeepers, who are paid less, presumably because their talents can only be used in one place in the team.[17]

In table 2 two sets of results are given. Columns (2), (3) and (4) give results for those specifications that – following Lucifora and Simmons (2003) - pick up age *and* experience as explanatory variables. Columns (5) to (7) replicate these, now excluding the experience variables. It is of note that the experience variable has a significant independent effect in addition to age. Thus the results including experience variables – columns (2), (3) and (4) - are to be preferred to those excluding these variables.[18] The results are built up sequentially across the table. In the second column, the simplest function containing only positional (along with the age and experience) variables is presented. It can be seen that, in this equation, whilst having the correct sign and close to significance at the 10% level, no significant month-of-birth effect emerges. However, it is of note that this equation is mis-specified when age alone is present (column (5)). Column (3) makes an adjustment for each club by the use of dummy variables and it is immediately clear as to the effect. Not only are these club variables significant but their introduction to reflect the "corporate nature" of team sport means that the marginally insignificant coefficient of RA in column (2) now turns to (marginally) significant at a 10 % level. The wage impact of players' RA that emerges is relatively small but

of including goals but excluding RA (via a Davidson and MacKinnon, 1981, J-test) finds in favour of the RA impact and against goals as suitable regressors in the equation.

17 All teams require 2 goalkeepers but rarely use more than one of them whereas outfield squad members are used regularly. The reserve goalkeeper is paid not to play!

18 It should be noted that the addition of the experience variable could be expected to lower the month-of-birth effect as better players are more likely to be selected at a younger age (and so gain more experience). This is akin to the Angrist and Krueger (1991) effect and it would be of interest to investigate the effect of instrumenting the experience variable. Caution would suggest that if the standard results hold, the OLS result will overstate the experience effect. However, it is important to note that even with experience variables in the equation, there is still evidence of a month-of-birth effect.

nonetheless present; players born late after cut-off are ceteris paribus paid more (as they are more productive, i.e. better).

The final column of the set (column (4)) explores the RA wage impact further. It has been observed in the literature that a RAE is more likely to occur in sports with high physical demands (for example, Grondin and Trudeau, 1991). Thus, we allow for possible interactions between the player's RA and his position. It can be seen that the results are stark. There is evidence of a significant RA impact for all except forwards with these most marked for goalkeepers and defenders. In the preferred equations of Table 2 – column (4) - there is no RA influence on wages for forwards yet a significant (at least at 10%) effect for other positions. This is in line with Grondin and Trudeau (1991) who find a RAE for NHL hockey players together with a positional bias. In their analysis, the RAE is strongest among defenders and goalkeepers. The latter result is explained by the fact that ice hockey goalkeepers carry heavy equipment, imposing more physical demands on these players[19]. In our analysis, the month-of-birth effect among soccer players indicates that later (after the cut-off) born defenders and goalkeepers (and to a lesser extent midfielders), in particular emerge as better players than their early-born colleagues. There can be two distinct reasons for this. First, it is possible that the discrimination by talent scouts is more severely biased in the case of goalkeepers and defenders. If later born players in these positions survive the RA discrimination, they will be 'better than average' and thus rewarded for their greater talents. The issue of forwards is different. It suggests that a forward's talent is more easily detectable at young age, meaning that talent scouts will 'find' it irrespective of the month-of-birth. As such there would be no discrimination of forwards in the selection by talent scouts (and we would expect the overrepresentation of early born players to be absent, or less prevalent for forwards). An alternative explanation for the fact that the wage bias is stronger for goalkeepers and defenders would be that forwards are "born" rather than made. Goalkeeping and defending on the other hand would require relatively more skills that can actually be learnt and – most important – learnt through a human capital production process that benefits from

[19] It is possible that position may be subject to an RAE effect and thereby, to an extent, endogenous to the analysis. To investigate this, tests of exogeneity of position were considered. However, the difficulty is to find appropriate instruments, as none readily lend themselves to the purpose. As a tentative test of this issue, instruments, based on the third (and higher) moments, as espoused by Dagenais and Dagenais (1997) were derived. Sargan tests indicated that the instruments were appropriate (i.e. exogenous), the Hausman test indicated exogeneity and the results were insignificantly different from those of the OLS regressions This suggests that this issue is not relevant for German soccer. The instruments used, dictated by the Monte Carlo simulations of Dagenais and Dagenais were a constant, z1 and z4 where (with * designating the Hadamard element by element matrix multiplication operator and the variables in deviation from mean form)

$$z1 = x*x; \qquad z4 = x*x*x - 3x[E(x'x/N)*Ik].$$

with x reflecting the right hand side variables in the equation, see Dagenais and Dagenais (1997, 197-198).

the complementarity of inputs discussed in the previous section. It should be noted that the fact that forwards are not ‘made’ does not imply that their talents can be easily detected at young age; in other words: a RAE can be expected to exist also for forwards. A more detailed analysis of the birth-patterns, however, reveals that the RAE *is* indeed most strongly present among German goalkeepers and midfields. For forwards no RAE is found. These observations lend some support to the previous ‘forwards are more easily detected by talent scouts’ and against the ‘forwards are *not* made’ interpretation of month-of-birth related wage differentials. Still, some caution is warranted as no significant RAE is found among defenders.[20]

CONCLUSION

In youth soccer competitions, players are typically grouped according to their chronological age. We analyse how such age grouping affects wage formation among German professional soccer players. In line with existing empirical work on many team sports, we find the presence of a relative age effect (RAE): players born early after the cut-off date used for age grouping are over-represented even at the professional level.

This over-representation may lead to a month-of-birth related wage bias reflecting that players born late after the cut-off date earn systematically more than those born early after the cut-off date. Such a wage bias can occur for one of two reasons. First, professional players born late after the cut-off date are a more selective subset than professional players born early after the cut-off date. The reason is that the scouting system is biased towards selecting from among the latter group. Thus, players who are born late after the cut-off date and still become professional have actually “survived” a system that discriminates against them. To succeed in such a “survival” is only possible if they are of above average talent. A second reason for a wage premium to occur for late-born players lies in the fact that they benefit from receiving higher quality soccer education. This advantage arises because of the specific nature of the human capital production process that develops their skills. Players at young age benefit from training and competing with better co-players. The age grouping creates a systematic advantage for the late born players who receive their soccer education together with players who – though equally talented - are better than themselves (just because they are a few months older).

[20] Spearman rank correlation coefficients (calculated as in table 1) were –0.69 for goalkeepers, -0,36 for defenders, -0,50 for midfields and –0,36 for forwards. These values reveal a significant RAE for goalkeepers and midfields, not for other players.

Analysing wage data of German Bundesliga soccer players (season 1998-1999), we find clear evidence of a month-of-birth related wage bias. Players born late after the cut-off date (of August, 1st) earn systematically more. Interestingly, the wage impact of the month-of-birth differs with the player's position in the field. The wage premium for being born late after the cut-off date is highest for goalkeepers and defenders. It is absent for forwards. This may reflect that talent scouts looking for young talents are more biased in their evaluation of these defensive players or that the skills required to be a good goalkeeper or defender are more characterised by complementarity of inputs in the human capital production process. In practical terms the latter would mean that it is more important for a goalkeeper or defender to have his soccer education with strong co-players than for a forward.

REFERENCES

Angrist, J.D. and A.B. Krueger, 1991, Does Compulsory School Attendance Affect Schooling and Earnings, *The Quarterly Journal of Economics,* 106 (4), 979-1014.

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

Bäumler, G., 1998, Der Relativalterseffekt bei Fussbalspielern und seine Wechselwirkung mit dem Lebensalter, in G. Bäumler and G. Bauer (eds.) Sportwissenshaft rund um den Fussball, Hamburg, 109-115.

Bound, J. and D.A. Jaeger, 1996, On the validity of season of birth as an instrument in wage equations: a comment on Angrist and Krueger's 'Does compulsory school attendance affect scholling and earnings?', *NBER working paper*, 5835.

Card, D., 1999, The Causal Effects of Education on Earnings in O. Ashenfelter O. and D. Card (eds.), *Handbook of Labor Economics*, V*olume 3*, Elsevier Science B.V., 1801-1863.

Dagenais M.G. and Dagenais, D.L., (1997). Higher Moment Estimators for Linear Regression Modes with Errors in Variables, *Journal of Econometrics*, 76, 193-221.

Davidson, R., and J.G. MacKinnon, J.G., 1981, 'Several Tests for Model Specification in the Presence of Alternative Hypotheses', *Econometrica*, 49, 781-793.

Dudink, A., 1994, Birth date and sporting success, *Nature,* 368, 592.

Grondin, S. and F. Trudeau, 1991. Date de naissance et ligue nationale de hockey: Analyses en fonction de differents parametres.,Revue des Sciences et Techniques des Activités Physiques et Sportives, 26, 37-45.

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

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

Hoxby, C.M., 2003, Economics of Education, *NBER Reporter,* winter 2002/2003, 1-5.

Idson, T.L. and L.H. Kahane, 2000, Team Effects on Compensation: an Application to Salary Determination in the National Hockey League, *Economic Inquiry,* 38 (2), 345-357.

Kahn, L.M., 2000, The Sports Business as a Labor Market Laboratory, *Journal of Economic Perspectives,* 14 (3), 75-94.

Lefgren, L., 2004, Educational peer effects and the Chicago public schools, *Journal of Urban Economics,* 56, 169-191.

Lehmann, E. and J. Weigand, 1999, Determinanten der Entlohnung von Profifussballspielern – Eine empirische Analyse für die deutsche Bundesliga, *Betriebswirtschaftliche Forschung und Praxis*, 59, 2, 124-135.

Lucifora, C. and R. Simmons, 2003, Superstar Effects in Sport: Evidence from Italian Soccer, *Journal of Sports Economics*, 40 (1), 35-55.

Musch, J., 2002, Birthdate and Success in Youth Soccer: Investigating the Development of the Relative Age Effect, *Sportonomics*, 8, 22-28.

Musch, J. and S. Grondin, 2001, Unequal Competition as an Impediment to Personal Developement: A Review of the Relative Age Effect, *Developmental Review*, 21, 147-167.

Musch J. and R. Hay, 1999, The Relative Age Effect in Soccer: Cross-Cultural Evidence for a Systematic Discrimination Against Children Born Late in the Competition Year, *Sociology of Sport Journal*, 16, 54-64.

O'Loughlin, H., 2002, An Investigation Into The Relative Age Effect And Irish Men's Senior Soccer, *mimeo*, Department of Physical Education and Sports Science, University of Limerick.

Plug, E.J.S., 2001, Season of birth, schooling and earnings, *Journal of Economic Psychology,* 22, 641-660.

Russell, R.J.H. en M.J. Startup, 1986, Month of birth and academic achievement, *Personality and Individual Differences*, 7, 839-846.

Verhulst, J., 1992, Seasonal Birth Distribution of West European Soccer Players: A Possible Explanation, *Medical Hypotheses*, 38, 346-348.

Table 1: Divergence between expected and actual number of Bundesliga players				
	Expected number of players	Actual number of Players	Divergence	"actual rank" of month (starting August)
April	24.0	14	+ 10.0	9
February	22.8	16	+ 6.8	7
June	24.1	18	+ 6.1	11
May	25.0	20	+ 5.0	10
July	24.8	20	+ 4.8	12
November	21.8	20	+ 1.8	4
March	25.2	24	+ 1.2	8
December	22.6	23	- 0.4	5
September	23.5	24	- 0.5	2
January	24.4	26	- 1.6	6
August	23.9	31	- 7.1	1
October	22.8	49	- 26.2	3
				$r_s = -.74$ $p < .01$

Table 2: The "Month of Birth Effect" on Wages in the Bundesliga

	(2)	(3)	(4)	(5)	(6)	(7)
Variable	ln Wage	ln Wage	ln Wage	ln Wage	ln Wage	ln Wage
Intercept	**4.431** (1.494)**	**3.970** (1.390)**	**4.859** (1.201)**	**3.614** (1.252)**	**3.435** (1.202)**	**3.576** (1.174)**
Age*	**5.805** (0.858)**	**5.680** (0.771)**	**5.687** (0.765)**	**6.195** (0.792)**	**6.101** (0.758)**	**6.215** (0.745)**
Age^2	**-0.890** (0.136)**	**-0.862** (0.121)**	**-0.858** (0.121)**	**-0.869** (0.124)**	**-0.863** (0.119)**	**-0.878** (0.117)**
Experience*	**0.989** (0.127)**	**0.766** (0.122)**	**0.751** (0.123)**			
$Experience^2$	**-0.186** (0.038)**	**-0.147** (0.036)**	**-0.147** (0.036)**			
Goalkeeper	**-0.579** (0.149)**	**-0.606** (0.134)**	**-1.266** (0.262)**	**-0.736** (0.141)**	**-0.720** (0.121)**	**-1.514** (0.269)**
Defender	-0.173 (0.131)	-0.151 (0.119)	**-0.614** (0.233)**	***-0.225 (0.123)***	-0.183 (0.124)	**-0.600** (0.241)**
Midfield	-0.183 (0.120)	-0.157 (0.109)	**-0.442** (0.206)**	***-0.209 (0.113)***	-0.168 (0.114)	**-0.478** (0.255)**
Relative Age (RA)	0.019 (0.011)	***0.019* (0.010)***	-0.037 (0.025)	0.018 (0.011)	**0.023** (0.011)**	-0.038 (0.026)
Goalkeeper*RA			**0.113** (0.039)**			**0.132** (0.040)**
Defender*RA			**0.078** (0.033)**			**0.069** (0.034)**
Midfield*RA			***0.050* (0.030)***			***0.051* (0.029)***
Club Dummies	No	Yes	Yes	No	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
	Random	Random	Random	Random	Random	Random
Hausman	2.001	1.445	1.338	1.944	1.007	1.117
R^2	0.402	0.538	0.552	0.257	0.480	0.499
F (clubs)		6.421	6.826		9.980	10.869
Heteroscedasticity	0.186	1.344	1.887	0.057	0.668	0.446
Normality	0.389	2.006	1.558	3.667	1.230	1.673
Functional Form	1.650	1.997	1.579	9.227**	0.779	0.114

Notes: * Age in months, divided by 100. Estimated standard errors are in parentheses. The diagnostic tests relating to the OLS regressions are given to indicate the appropriateness of the specification of the initial regressions before applying panel estimation and refer to the fixed effect regressions. Heteroscedasticity is a Breusch-Pagan test; normality is a Jarque-Bera test and Functional Form is a Ramsey RESET test. The inclusion of the year dummy is to accommodate for general inflation of wages; it has a positive sign and a coefficient around 0.225. Results using the group means are in line with those reported here and available on request. The random panel estimation treats each player as different, the F test between this stratification and just pooling is on the borderline of significance with F tests close to unity. ** indicates significant at 5% and * indicates significant at 10%

APPENDIX
TABLE A1
SUMMARY STATISTICS OF VARIABLES

Variable	Minimum	Maximum	Average	Median	Standard Deviation
Wage	100,000	12,000,000	2,169,917	1,500,000	2,123,308
Age in Mnths (in Years)	203 (16.9)	453 (37.7)	321.423 (26.8)	324 (27)	51.210 (4.27)
Experience (Matches)	0	449	95.967	67	100.858
Position					
Keeper	0	1	0.143		
Defender	0	1	0.272		
Midfielder	0	1	0.428		
Forward	0	1	0.158		
Wage by Position					
Keeper	100,000	8,000,000	1,352,173	800,000	1,642,819
Defender	200,000	8,500,000	2,075,190	1,500,000	1,685,691
Midfielder	100,000	11,000,000	2,309,709	1,500,000	2,248,908
Forward	200,000	12,000,000	2,696,711	1,500,000	2,598,453
Age by Position					
Keeper	203 (16.9)	444 (37.0)	321.69 (26.8)	322 (26.8)	58.45 (4.87)
Defender	222 (18.5)	453 (37.7)	328.30 (27.4)	330 (27.5)	46.93 (3.91)
Midfielder	210 (17.5)	432 (36.0)	321.13 (26.8)	326.5 (27.2)	51.72 (4.31)
Forward	219 (18.2)	399 (33.2)	310.12 (25.8)	305.5 (25.5)	48.77 (4.06)
Experience by Position					
Keeper	0	377	77.290	35	100.173
Defender	0	449	107.656	77	108.136
Midfielder	0	381	100.762	66.5	104.769
Forward	0	294	79.776	76	70.581

Notes: The table treats observations of the same player in two years as being different observations and is provided for illustrative purposes.