



Munich Personal RePEc Archive

## **The Relative Age Effect Reversal among NHL Elite**

Luca Fumarco and Benjamin Gibbs and Jonathan Jarvis and  
Giambattista Rossi

STATEC, Brigham Young University, Birkbeck, University of  
London

12 December 2016

Online at <https://mpra.ub.uni-muenchen.de/75691/>

MPRA Paper No. 75691, posted 21 December 2016 13:18 UTC

## The Relative Age Effect Reversal among NHL Elite<sup>1</sup>

Luca Fumarco<sup>2</sup>  
Benjamin G. Gibbs<sup>3</sup>  
Jonathan A. Jarvis<sup>3</sup>  
Giambattista Rossi<sup>4</sup>

### ABSTRACT

December 31<sup>st</sup> age cut-offs in junior hockey create an advantage for players born in the first quarter of the year, most likely because they are relatively bigger than their younger counterparts born later in the year. As this Relative Age Effect (RAE) has been well-established in junior hockey and across other professional sports, we argue that the long-term impact of this phenomenon is still poorly understood. Using roster data on North American NHL players from 2008 to 2015, we examine the RAE in terms of birth month distribution and the extent that RAE is associated with points (i.e. goals plus assists) and player salaries. We find evidence of an RAE reversal—that players born in the second half of the year (July-December) score more points per season (29-50% more points) and command higher salaries (30%-50% more salary). Among elite players—the highest scoring and highest paid athletes—the scoring gap ranges between 14% and 26% more points for players born in the second half of the year—whereas the salary gap ranges between 18% and 50% greater salary. We argue that results partly support an “underdog” effect in NHL that is greatest among elite players.

Keywords: relative age effect, hockey, performance outcomes, quintile regression

---

<sup>1</sup> Contact Luca Fumarco at B11.E4, National Institute of Statistics and Economic Studies of the Grand Duchy of Luxembourg (STATEC). E-mail at [luca.fumarco@statec.etat.lu](mailto:luca.fumarco@statec.etat.lu), office +352-247-84271. We are grateful for the assistance of the NHL coding team and Nick Ustach.

<sup>2</sup> National Institute of Statistics and Economic Studies of the Grand Duchy of Luxembourg

<sup>3</sup> Brigham Young University

<sup>4</sup> Brikbeck, University of London

Examining the distribution of players born across the calendar year is a common way to examine the Relative Age Effect (RAE). If more players born in the first months of the year appear on sports rosters and are more represented in professional sports, then this is evidence a relative age effect. While this phenomenon is evident in multiple studies (Deaner et al. 2013; Gibbs et al. 2012; Nolan and Howell 2010; Addona and Yates 2010; Baker and Logan 2007; Barnsley and Thompson 1988) only few examine RAE on players' productivity beyond roster counts. One notable exception is Bryson and colleagues' work (2014). They examine how relative age (and relative birth cohort size) affects salaries among National Hockey League players. Bryson and colleagues find that lower relative age (born towards the end of the calendar year) decreases the chances of drafting into the NHL. But once in, these players score more points and command higher salaries.

In this paper, we build on the previous literature and expand it. Following Ashworth and Heyndels (2007) work, we analyze the link between birth quarter and measures of athletic performance. We add to the literature by examining RAE on points and salary across different quantiles. If the RAE reversal is present as Gibbs et al. (2012) and Bryson et al. (2014) suggest, we argue that the reversal should be greatest among the NHL elite. Because players born at the end of the year were initially at a size disadvantage in the junior leagues, if they successfully make the NHL, we expect these deficits to actually amount to a kind of an "underdog effect," where these temporarily disadvantaged athletes benefited by overcoming size disadvantages that will translate to higher productivity in the NHL than players born at the start of the year.

We also account for another important factor: the NHL drafting rule (only players who turn 18 by September 15<sup>th</sup> but who are not older than 20 before December 31<sup>st</sup> of the draft year are eligible for the draft); this is an unexplored feature of the RAE in NHL.

## LITERATURE REVIEW

There is considerable evidence for RAE—being born closer to age cutoffs is an advantage in a variety of ways. Relative age differences have been found to impact child outcomes such as education (Cobley et al. 2009; Elder and Lubotsky 2009), self-esteem (Thompsons, Barnsley and Battle 2004) or even physical strength (Sandercock et al. 2013). Because of the size and maturity advantage for children born closest to grade and sport cutoffs, they receive more exposure to better competition which contributes to more time for deliberate practice and the development of abilities (Baker and Logan, 2007; Barnsley et al., 1985; Ericsson, 2007; Musch and Gondin, 2001; Fumarca and Rossi, 2015). In this sense, maturity is initially mistaken for talent, and those with greater physical maturity are then provided more opportunity to develop their talent.

The impact of RAE in sports has been examined across a number of different sports and in many different nations. For example, Ashworth and Heyndels's study (2007) of soccer players, found that players born shortly after the age cutoff were more likely to play German professional soccer. Safranyos et al. (2016), found a significant advantage for male and female college volleyball players in Canada born closer to the age cutoff. Helsen et al. (2013) looked at the effect of RAE in European professional soccer players in ten countries over a 10-year period and found a consistent RAE advantage over this time period. More recently,

Garcia et al. (2014) found a significant RAE among male and female international basketball players, but also found that the effects faded as the players got older.

Research on RAE and sports has perhaps been most thoroughly examined among hockey players. Three decades ago, Barnsley et al. (1985) researched players in the NHL and junior hockey (the league that feeds most players to the NHL) and found a strong linear relationship between league participation and birth month. When examining RAE in even younger players (minor hockey), Barnsley and Thompson (1988) found that Canadian children born in the first half of the year were more likely to play minor hockey and more likely to play for top teams. Twenty-five years after Barnsley and colleagues original study, research on RAE continues to suggest a strong relationship between birth month and the proportion of players in junior hockey in Canada (Nolan and Howell 2010). As RAE increased the likelihood of being chosen in the NHL draft (Baker and Logan 2007; Gibbs et al. 2012), paradoxically, RAE appears to have a limited impact on measures of skill and performance in professional play (Wattie et al., 2007) and is not prevalent among Hall of Fame players (Addona and Yates 2010), a surprising conclusion given the considerable research attention on this phenomenon, especially in pre-professional settings.

Furthermore, there is growing evidence the RAE might actually reverse as players advance to professional sports. A reversal has been found in soccer, rugby, handball, cricket, and hockey where relatively younger players appear to suffer disadvantage earlier on, but overcome this disadvantage to earn more money (Ashworth and Heyndels, 2007), have longer careers (Gibbs et al. 2012), and are selected on the most elite squads (McCarthy, Collins and Court 2016; Gibbs et. Al 2012).

There are three potential explanations we will consider to explain any RAE reversal we might find in our analyses. The first is psychological. Smaller players in junior hockey who make it to the NHL demonstrate higher than average resilience due to their ability to overcome their size limitations (McCarthy and Collins 2014; Schorer et al. 2009; Ashworth and Heyndels 2007). To compete against their relatively older and bigger peers, these players born later in the year learn to work harder (Roberts and Stott 2015), that is, they enjoy general positive peer effects. This initial disadvantage then works in their favor when early differences in size reach parity past puberty. The effect, then, is psychological, they are now better equipped to overcome subsequent obstacles and succeed (Schorer et al. 2009). Thus, this early disadvantage (if they can overcome it) becomes a later advantage in professional play—an *underdog effect*.

Another related explanation suggests that the players born later in the year (farther from the cutoff) who are subsequent successful athletes may not only have a degree of resilience, but also superior ability. For these younger, smaller to survive, “a system that discriminates against them,” (372), they must be more talented than their relatively larger counterparts to counteract their size disadvantage (Ashworth and Heyndels, 2007), that is, these players are likely selected from the right tail of the ability distribution. Therefore, while maturity and size (RAE) can delay or postpone the screening of players with inferior ability among experiencing, it in turn, selects the most talented among the smallest and youngest when it comes to professional and elite play in the NHL.

Although these first two explanations (effort plus talent) might provide a persuasive explanation for an RAE reversal among the NHL elite, there might be a less obvious factor influencing the reversal. As the NHL restricts entry into the draft, only players who turn 18 by

September 15<sup>th</sup> but who are not older than 20 before December 31<sup>st</sup> of the draft year are eligible. This means that all 18 year olds who are drafted in the NHL are born in the first three quarters of the year. Thus, the same factors that initially benefitted those born in the beginning of the year may reverse their advantage by making them some of the youngest among the NHL draft picks. If being slightly older at the start of an NHL career is any advantage, than the benefactors of RAE in junior play are now at a disadvantage, simply by another cutoff effect.

As any RAE reversal requires explanation, we will explore these possibilities to the extent that the data allow. Ultimately, our aim here is to first establish the RAE reversal among the highest scoring and highest earning players. Subsequent work, perhaps qualitative, would need to be pursued to further uncover the mechanisms of the RAE reversal.

## DATA AND METHODS

To examine the association between players' quarter of birth and performance outcomes, we compiled data of nearly all NHL players over 8 consecutive NHL seasons (2008 to 2015, N=2,017). This provided a total of 8,760 player-observations. We analyzed a sub-population of these initial data with both descriptive statistics and regression analyses (both OLS and the quantile regression).<sup>5</sup>

---

<sup>5</sup> This approach also allows for potentially co-occurring negative and positive effects on points and salaries to be uncovered across the distribution while mitigating the presence of outliers, which represent an issue particularly when analyzing salary distributions (Adler 1985; Rosen 1981). In addition, the usage of the quantile regression mitigates sample-selection issues, which occurs with the arbitrary segmentation of the sample into sub-samples based on players' salaries (see Lê Cook and Manning 2013).

### *Outcome Measure*

The outcome measures are points (i.e. goals plus assists) and salaries. In accordance with existing literature, we transform salaries data into their natural logarithm; moreover, salaries are deflated at 2015 CPI-U. For interpretational purposes and consistence in the methodology we apply the logarithmic transformation to points as well.

### *Independent Measures*

The key measure of interest is birth quarter. We created dummy variables for quarters of birth, with the aim of capturing possible nonlinear effects (Peña, 2015).<sup>6</sup> The reference is the first quarter (January-March). When we analyze the quarter of birth distribution, we need only this information, whereas in the econometric regressions we use additional information, namely player's age when the draft occurred, age<sup>7</sup> and standardized squared age,<sup>8</sup> body mass index,<sup>9</sup> player position, season, team, country, and draft year.

We do not use all the 8,760 player-observations, but focus on drafted American and Canadian non-goaltenders that currently play in NHL, which totals 4,447 player-observations. While non-drafted players could be included in the investigation with the basic model without control variables, they would be automatically dropped when we control for draft year and draft age, which would hinder results comparability across models; this is the reason

---

<sup>6</sup> Differently from some previous literature (in NFL, Böheim and Lackner 2011; in school, Bedard and Duhey 2006 as well as Ponzio and Scoppa 2014), we do not use quarter of birth as an instrument for age at draft (or, equivalently, as an instrument for age at school entry), because based on the RAE literature RAE can directly affect performance through other mechanisms. Therefore, the so called "exclusion restriction" would be violated—an instrument is valid if we can rule out direct effects of the instruments on the outcome variable.

<sup>7</sup> The minimum value of age (i.e. 18) is subtracted; the same transformation is applied to age at draft.

<sup>8</sup> Standardization of squared age is implemented to break the correlation between age and its square.

<sup>9</sup>  $bmi = weight/(height^2)$ , this is the standard formula, as suggested in from Currie et al. (2008) report.



for their exclusion. Goalies are excluded from the analyses to preserve comparability of the results for different outcomes; in fact, although there are figures on goalies' salaries and draft round there is no statistics on goalies' points. For a similar reason, observations on players who belong to an NHL team but that in a given season play abroad or in minor leagues are excluded: we have information on their salaries and draft round, but not on their points. Finally, the analyses are restricted to Northern American players because we do not have information on when non-Northern-American players started to play professionally, so we cannot construct any measure of experience for these players.

## **Analysis**

### **1.1. Birth dates distribution**

In this section we analyze the birth date distribution. We first present descriptive statistics on the frequency of quarters of birth for the entire population we are considering. Since we have complete information on the population and there is virtually no measurement error on players' birthday, inferential statistics is not necessary (Gibbs et al. 2015). Afterwards, we contribute to the literature by conducting this analysis on different sub-populations based on information from the NHL drafting system. Rules establish that only players within a certain age-range in the draft year are eligible for the NHL draft: those who turn 18 by September 15<sup>th</sup> up until those who turn 20 by December 31<sup>st</sup> in the draft year. To the best of our knowledge, this is the first study to conduct analyses on the RAE on the birth date distribution and accounting for this rule, according to which players born in the last quarter can be drafted only when they are 19 or 20.<sup>10</sup>

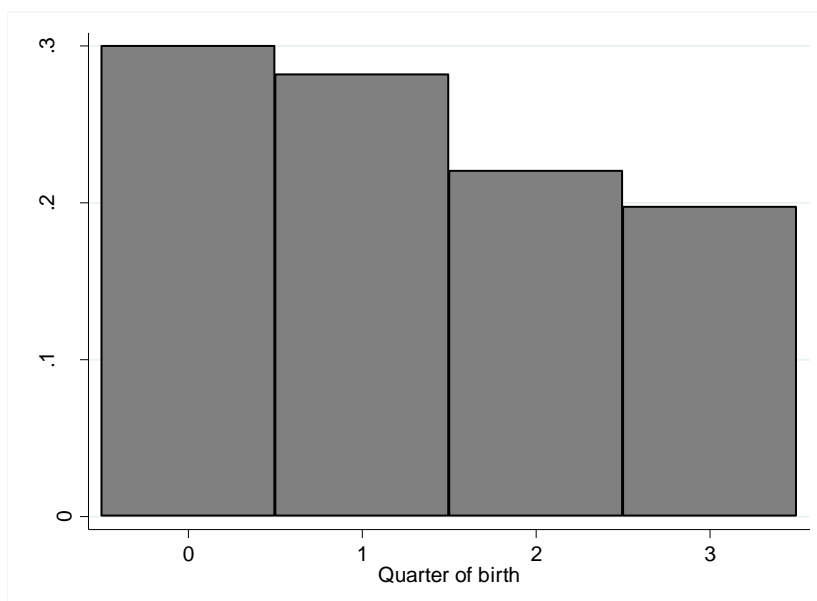
---

<sup>10</sup> This is true also for players born toward the end of the third quarter (September 16<sup>th</sup> to September 30<sup>th</sup>)

This drafting rule has two direct consequences on players born in the fourth quarter. On one hand, it limits their eligibility period compared to relatively older peers (2 years of eligibility against 3 years for players born in other quarters). On the other hand, the best players born in the first three quarters are already drafted when fourth quarter players are non-eligible (i.e. they are 18); so we might expect either a reduction of RAE in terms of representativeness or its reversal among 19 and 20 years old draftees, when also fourth-quarter players are eligible.

Figure 1 reports the frequencies of quarter of births for drafted North-American NHL non-goals (N=4,447); January-March is quarter 0, while October-December is quarter 3.

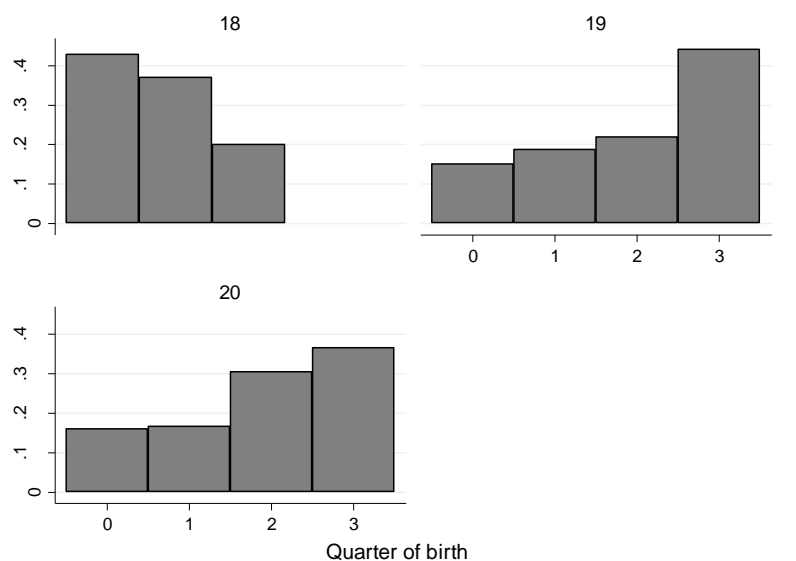
**Figure 1. Quarter of birth rate distribution.**



This graph confirms what observed in previous studies: players born in the third and fourth quarter are under-represented, since they represent approximately 20% of the distribution each.

What happens when we investigate the frequencies of quarter of births based on age at draft, that is, 18 (N=2,363), 19 (N=1,538) and 20 (N=546)? See Figure 2.

Figure 2. Quarter of birth distributions, based on age at draft.



As by the rule, this graph shows that no player born in the fourth quarter is drafted at 18; differently, it shows that at 19 and 20 years of age about 45% and 38% respectively of the players are born in the fourth quarter.

Although we exclude from the focus of this article players who entered NHL as free-agents and currently play in NHL (N=935) and observations on players who currently do not play in NHL (N=819)—note that this sub-population includes players who were drafted as well as players who entered as free-agents, we looked also at these two quarter of birth distributions. These two figures are reported in the Appendix: Figure A.1 shows an approximately uniform distribution of quarters of birth for free-agents, while in Figure A.2 we observe an over-representation of players born in the first two quarters among non-NHL players. These results offer additional pieces of information for our final discussion of the findings.

## 1.2. Points

In this section we investigate RAE on points. We fit an OLS with three model specifications:<sup>11</sup> first, the analysis is conducted only with dummy variables for quarter of birth; second, all the control variables are introduced, except for current age, its standardized square, current team as well as current season;<sup>12</sup> third, we repeat the analysis only for Canadian players, and, only for this sub-population, add a dummy variable for players having played in the Canadian Junior Hockey League. The regression on Canadians works as a robustness check: Canadian and American players trained under the same cut-off (December 31<sup>st</sup>) except for players who grew up in Minnesota (August 31<sup>st</sup>): the restriction to Canadian players insures the same cut-off date applies to everybody. Afterwards, we re-investigate the data with a quantile regression. Since we use repeated observations on individual players through time—scored points potentially change from season to season, these observations are not likely to be independent, thus we cluster standard errors on players. The results are displayed in Table 2.

**Table 2.** RAE by quarter, on natural logarithm of points; OLS.

Variables	North.-Am. Ln_Points (1)	North.-Am. Ln_Points (2)	Canadians Ln_Points (3)
April-June	0.081 (0.088)	0.064 (0.084)	0.154 (0.099)

<sup>11</sup> As this is not a sample, inferential statistics would not be necessary (Gibbs et al. 2015). However, skeptic readers may still claim the possible presence of measurement errors introduce uncertainty in our analyses; therefore we report p-values. Moreover, in doing so we are following other published papers that study RAE on salaries (Ashworth and Heyndels, 2007; Böheim and Lackner, 2012).

<sup>12</sup> We believe that age at draft (similarly to age at school entry) is affected by RAE, as suggested in Bedard and Duhey (2006), Ponzio and Scoppa (2014), and Böheim and Lackner (2011). However, its inclusion as a control variable changes our interpretation of the RAE estimates. A similar approach, where a potentially dependent variable is introduced in the model as an independent variable to refine the interpretation of the initial results, is adopted in studies on salary discrimination based on ethnicity, where performance measures are inserted to the right hand side (for a literature review see Kahn, 2000).

July-September	0.207** (0.094)	0.258*** (0.090)	0.320*** (0.107)
October-December	0.239** (0.094)	0.413*** (0.102)	0.386*** (0.124)
Control variables	N	Y	Y
Can. Jr. Hockey	N	N	Y
Observations	3,927	3,927	2,746
R-squared	0.007	0.111	0.139

Note: Only Northern-American players are investigated; column (3) focuses on the sub-population of Canadian players. In this investigation, repeated observations per player are used. All the control variables are included. This analysis excludes 520 observations on players who scored 0 points. Standard errors in parenthesis are clustered on players. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

This table shows that relatively young players score more points than relatively older counterparts, on average. Players born in the fourth quarter score on average twice the points of players born in the first quarter,<sup>13</sup> and players born in the third quarter score 29% additional points compared to peers born in the January-March quarter. Although the estimates are slightly different, the analyses on the sub-population of Canadians confirm these results (column (3)); the same is true for the regression with no control variables (column (1)).<sup>14</sup>

To investigate whether the RAE on points varies by quantile of hockey players' points distribution, we use the quantile regression technique; to the best of our knowledge, this is

<sup>13</sup> Since the outcome variable is transformed with a natural logarithm, the salary gap in percentage terms is computed as  $[\exp(0.207)-1]*100= 22\%$ , for example.

<sup>14</sup> We conduct three further robustness checks. First, we repeat the analyses using experience and its standardized square in place of age, as well as season dummies. Second, we conduct the analysis using a negative binomial regression, where therefore observations on athletes who did not score are used. Third, we repeat these analyses only on 19- and 20-year old players. These additional analyses provide results that point to the same direction and can be provided upon request.

the first study to do that. We claim this is the ideal econometric technique to study the RAE on performance and its proxies (e.g. salary, as in the next section) for three reasons. First, this is the best suited methodology to investigate the evolution of RAEs along the points distribution. This method drops the assumption that the RAE is the same at the lower/upper tail of the distribution as at the mean while in fact the quantile regression describes the relationship between regressors and conditional quantiles of the outcome variable: it weights different quantiles of the sample to gain the estimated coefficients (Lê Cook and Manning 2013). Therefore, we can obtain a more complete picture of the RAE on points in NHL.

Second, some studies (Ashworth and Heyndels 2007; Gibbs et al. 2012; Bryson et al. 2014; Fumarco and Rossi 2015) suggest the possibility that even if relatively young athletes have been disadvantaged at the moment of the selection into the professional career (i.e. they are underrepresented), they are better players because they enjoyed positive peer effects, and underwent positive selection. In general, if it was true, we could observe three alternative phenomena: i) in case the OLS detected an average positive RAE in terms of points gap, this gap could increase approaching the top quantiles of the points distribution; ii) in case the OLS detected an average negative RAE in terms of points gap, this gap could tend to decrease approaching the top quantiles of the points distribution—and possibly reverse; iii) in case the OLS does not detect statistical significant evidence of the existence of a points gap, that might still exist in lower/upper tails of the points distribution. In this light, the usage of the OLS could lead, not only to an incomplete picture on RAE, but also to misleading results.

Despite this advantage, some researchers might still prefer to use the OLS. In this case, we recommend the quantile regression to be used at least as a robustness check to compare the estimates obtained at the conditional median of the outcome variable, that is, the 50<sup>th</sup>

percentile of the distribution, to the estimates obtained at the mean of the outcome variable with the OLS. Third, for at least two reasons the quantile regression should be preferred over the utilization of the OLS for analyzing different sub-populations/samples based on salary level. First, the investigation of sub-samples reduces efficiency, since the estimates are obtained from smaller samples. Second, the investigation of sub-samples causes sample selection bias, which occurs with the arbitrary segmentation of the sample into sub-samples (Lê Cook and Manning 2013).

We implement the analyses at the 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles of the points distribution of Northern-American players. For sake of brevity, the results in Table 3 are obtained from the specification that includes all the variables, and, also in this case, because repeated observations on individual players are not likely to be independent, standard errors are clustered on players; we use the method suggested by Parente and Silva (2013).<sup>15</sup>

**Table 3.** RAE by quarter, on natural logarithm of points; quantile regression.

Variables	North.-Am. Ln_Points 25%	North.-Am. Ln_Points 50%	North.-Am. Ln_Points 75%	North.-Am. Ln_Points 90%
April-June	0.071 (0.133)	0.079 (0.116)	0.130 (0.099)	0.035 (0.064)
July-September	0.358** (0.159)	0.353*** (0.116)	0.263*** (0.086)	0.131* (0.071)
October-December	0.491*** (0.187)	0.471*** (0.136)	0.380*** (0.108)	0.234*** (0.072)
Control variables	Y	Y	Y	Y
Can. Jr. Hockey	N	N	N	N
Observations	3,927	3,927	3,927	3,927
Pseudo R-squared	0.088	0.097	0.085	0.062

Note: Only Northern-American players are investigated. In this investigation, repeated observations per player are used. All the control variables are

<sup>15</sup> We use the Stata command qreg2 created by Parente and Silva (2013).

---

included; no analysis on the sub-sample of only Canadians is carried out. This analysis excludes 520 observations on players who scored 0 points.. Standard errors in parenthesis are clustered on players. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

The estimates obtained from the quantile regression confirm the initial results and offer additional insights. In no percentile, do the players born later in the year underperform compared to older peers. Moreover, we observe that the reverse RAE decreases in the quantile of the points distribution: the positive points gap in favor of athletes born later in the year is large, being 63% additional points; then it decreases slightly at the median points; and, finally, at the 75% and 90% of the points distribution, relatively young players score 46% and 26% more points than relatively older counterparts. Also athletes born in the third quarter enjoy a positive points gap, which ranges between about 43% and 14% additional points. Finally, if we compare the estimates of the OLS (that are obtained at the conditional mean salary) with the estimates from the quantile regression at the conditional median scored points (i.e. %50 column), we do find that the initial estimates obtained with the OLS are an underestimate of the RAE on points: the quantile regression at the median tells us that players born in the fourth quarter score 63% more points than players born in the first quarter, compared to the additional 51% points suggested by the OLS.<sup>16</sup>

Apparently, these results go against our argument that the RAE reversal is greatest among the NHL elite. However, this result depends on the logarithmic transformation we applied to points, which leads to the interpretation of the estimated coefficients in

---

<sup>16</sup> The analyses in this section are conducted also on free-agents alone. The analysis with the OLS provides weaker evidence of RAE on points, while the quantile regression provides confirm the results. In interpreting these additional analyses, it is important to consider the much smaller sample. These results can be provided upon request.



percentage terms. If we look at points in absolute terms, we find the result we expected: the point gap increases in the quantile of the points distribution. See the results in Table A.1, in the Appendix.

### 1.3. Salaries

In this section we investigate RAE on salaries. The analysis proceeds in a similar manner as for that on points: initially, we fit an OLS with three model specifications; afterwards, we use a quantile regression. Also in this investigation, standard errors are clustered on players because we use repeated observations on individual players through time—salaries change through seasons. Estimates obtained with the OLS are in Table 4.

**Table 4.** RAE by quarter, on natural logarithm of salaries; OLS.

Variables	North.-Am. Ln_Salary (1)	North.-Am. Ln_Salary (2)	Canadians Ln_Salary (3)
April-June	0.126 (0.087)	0.087 (0.076)	0.134 (0.088)
July-September	0.199** (0.094)	0.260*** (0.087)	0.326*** (0.102)
October-December	0.239** (0.093)	0.412*** (0.099)	0.390*** (0.122)
Control variables	N	Y	Y
Can. Jr. Hockey	N	N	Y
Observations	4,447	4,447	3,103
R-squared	0.006	0.200	0.225

*Note:* Only Northern-American players are investigated; column (3) focuses on the sub-population of Canadian players. In this investigation, repeated observations per player are used. Standard errors in parenthesis are clustered on players. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

This analysis confirms the presence of strong reverse RAE. Column (2) shows that North-American players born in the fourth quarter earn on average 50% higher salary than peers born in the first quarter,<sup>17</sup> while players born in the third quarter earn about 29.7% higher salaries. The results obtained investigating the sub-populations of Canadian players alone confirm these findings (columns (3)), while the results in column (1) obtained without control variables suggest a lower reverse RAE.<sup>18</sup>

To investigate whether the RAE on salaries varies by quantile of hockey players' salaries distribution, we use the quantile regression technique; to the best of our knowledge, this is the first study to do so. Besides the advantages discussed in the previous subsection, the usage of the quantile regression is particularly relevant when researchers deal with outcome variables characterized by extremely positively skewed distributions, as it is with the distribution of athletes' salaries because of the presence of superstars (Adler 1985; Rosen 1981). To overcome problems related to the skewedness of the wage distribution, usually scholars implement a logarithmic transformation of salaries, which we implement as well. However, although the resulting distribution is less skewed, it might still not be sufficiently normal, and, as a consequence, OLS estimates might still be biased. Although this is the ideal econometric technique to investigate RAE on salaries, only one study has used it before within a similar context (on Italian soccer players' salaries, Fumarco and Rossi 2015).

---

<sup>17</sup> Since the outcome variable is transformed with a natural logarithm, the salary gap in percentage terms is computed as  $[\exp(0.199)-1]*100= 22\%$ , for example.

<sup>18</sup> We conduct two further robustness checks. First, we repeat the analyses using experience and its standardized square in place of age, as well as season dummies. Second, we repeat these analyses only on 19- and 20-year old players. These additional analyses provide results that point to the same direction and can be provided upon request.

We implement the analyses at the 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles of the salary distribution of Northern-American players. For sake of brevity, the results in Table 5 are obtained from specifications that include already all the variables, and, also in this case, because repeated observations on individual players are not likely to be independent, standard errors are clustered on players; we use the method suggested by Parente and Silva (2013).<sup>19</sup>

**Table 5.** RAE by quarter, on natural logarithm of salaries; quantile regression.

Variables	North.-Am.	North.-Am.	North.-Am.	North.-Am.
	Ln_Salary 25%	Ln_Salary 50%	Ln_Salary 75%	Ln_Salary 90%
April-June	-0.019 (0.043)	0.042 (0.060)	0.041 (0.091)	0.064 (0.083)
July-September	0.064 (0.050)	0.194** (0.077)	0.217** (0.093)	0.166* (0.086)
October-December	0.149** (0.066)	0.289*** (0.083)	0.392*** (0.112)	0.414*** (0.110)
Control variables	Y	Y	Y	Y
Can. Jr. Hockey	N	N	N	N
Observations	4,447	4,447	4,447	4,447
Pseudo R-squared	0.125	0.172	0.158	0.128

Note: Only Northern-American players are investigated. In this investigation, repeated observations per player are used. All the control variables are included; no analysis on the sub-sample of only Canadians is carried out. Standard errors in parenthesis are clustered on players. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

As expected, the estimates obtained from the quantile regression offer deeper insights on the characteristics of the RAE on NHL player salaries. In none of the percentiles that we are

<sup>19</sup> We use the Stata command qreg2 created by Parente and Silva (2013).

considering, do the players born later in the year earn less than their older peers. Moreover, we observe that the reverse RAE is driven by salary disparities in top quantiles: the positive salary gap in favor of athletes born later in the year is small, being 16%; then it increases at the median salary, where players born in the fourth quarter earn 33.5% more than players born in the first quarter; and, finally, at the 75% and 90% of the salary distribution, relatively young players earn about 48-51.3% more than relatively older counterparts. Also athletes born in the third quarter enjoy a positive salary gap, but it is smaller, it does not seem to increase by quantile in a consistent way, and appears only from the median of the salary distribution. Finally, if we compare the estimates of the OLS with the estimates from the quantile regression at the conditional median salary, we observe that the estimates obtained with the OLS are an overestimate of the RAE on salaries: the quantile regression at the median tells us that players born in the fourth quarter earn 33.5% higher salaries than players born in the first quarter, compared to the 50% suggested by the OLS.<sup>20, 21</sup>

## CONCLUSION

RAE in NHL, in terms of skewed quarter of birth rates distribution is a well-known and thoroughly studied phenomenon in the literature. However, the analysis of RAE only in terms of representativeness (i.e. quarters of birth distribution) offers a limited view of the phenomenon. Other important dimensions, that is, RAE in terms of performance and salaries,

---

<sup>20</sup> Both analyses in this section are conducted also on free-agents alone. Neither analysis provides strong evidence of RAE on salaries for free-agents. These results can be provided upon request.

<sup>21</sup> This result is not in contradiction with the RAE on the natural logarithm of points; and we believe it to be related to the superstar effect—as explained in Rosen (1981), that is, relatively few additional points compared to other athletes lead to much greater salaries. In fact, when we analyze RAE on (the absolute quantity of) points, we observe that the point-gap increases (even though in percentage terms decreases) in the quantile of the salary distribution and the OLS provides an underestimation of the RAE—compared to the equivalent analysis with the quantiles regression at the median..

have instead been partially neglected, with the only investigation of the latter two outcomes being Bryson et al. (2014). Our paper contributes to this small strand of literature by exploring the RAE on these two outcomes; this is also the first study to explore how RAE on points and salaries vary along the salary distribution and how the RAE in terms of representativeness varies by age at draft. We restrict our analyses on drafted American and Canadian non-goaltenders that currently play in NHL; we focus our study on this sub-population to provide comparable results across different outcomes, and across model specifications. First, we analyze the RAE on the distribution of quarter of birth. When we investigate the whole sample, we obtain the usual result: relatively young players (that is, players born in the fourth quarter) are under-represented. When we analyze this distribution on different sub-populations based on age at draft, we observe that the drafting rule creates an additional disadvantage to relatively young players in terms of representativeness in NHL: while NHL teams prefer to draft 18-year old players, fourth quarters cannot be drafted at that age. Therefore, although we observe a RAE reversal on the distributions of quarter of birth at 19 and 20, this is not enough to compensate for not being eligible at 18. We also show that the quarter of birth distribution of free-agents is about uniform. Finally, we observe that players born in the first half of the year are overrepresented in the sub-population of non-NHL players.

Second, we find evidence of positive RAE in terms of points. In each season, players born in the fourth quarter score about twice the points of players born at the beginning of the year. These results are confirmed when we investigate only Canadians. The usage of the quantile regression suggests that the gap reduces but does not disappear among the best players. Moreover, the quantile regression at the median provides evidence that players born

in the last quarter score actually about 60% more points than peers born at the beginning of the year.

Third, we find a strong positive RAE on salaries. Our estimates show that, all else equal, “underdogs” earn on average a stunning 50% more than players born in the first quarter. We re-analyze these data with a quantile regression. We find that in none of the percentiles that we are considering (25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, 90<sup>th</sup>), relatively young players earn less than older peers and, finally, we observe that the standard analysis with the OLS gives upward biased results. In fact, the estimates obtained with the quantile regression at the median provide strong evidence that players born in the last quarter earn “only” about 33.5% more than peers born at the beginning of the year.

How can we interpret these combined results? Because of NHL rules, relatively young players are eligible for drafting only when they are 19 or 20, while NHL teams prefer to draft 18 years old players; we recommend future studies to investigate in detail the reason for this age preference. The drafting rule and teams’ preferences reinforce the RAE in terms of representativeness that we would observe because of maturity gaps alone. When finally (at 19 years of age) players born in the fourth quarter are draft-eligible, the best relatively old players have been already drafted when they were 18, which can explain part of the RAE reversal on birth dates distribution at 19 and 20. Positive selection of and positive peer effects on relatively young athletes might provide an edge in terms of performances (and thus wages) in favor of fourth-quarter players, as discussed by the previous literature. As Bryson et al. (2014) as well as Ashworth and Heyndels (2007) explain, despite the expected lower physical and psychological maturity in youth, some relatively young players do not drop out in youth, they overcome adversities, and start a professional career in sports; these athletes have been likely

selected from the right tail of the ability distribution. Symmetrically, more often relatively old players might come from the left-tail of the ability distribution, which concurs to the reverse RAE on performance; this speculation is supported by the observation that these players are often sent to improve their skills in other leagues. Moreover, relatively young players might have benefitted from training with older peers who were more mature throughout the youth career (McCarthy et al. 2016; McCarthy and Collins 2014; Schorer et al. 2009; Ashworth and Heyndels 2007), Finally, since relatively young players have to start their career in minor leagues, when they are drafted they might have cumulated more playing-time than older peers who have been already drafted at 18 by an NHL team, but might not have played an equivalent amount of time. This additional on-ice time may provide an additional edge in terms of performance (and thus wages) to fourth-quarter players during their professional career. This interpretation on the effect of a delayed draft on performances is compatible with evidence from NFL in Böheim and Lackner (2011). However, we also observe that beyond age-draft, additional time in minor leagues does not help any further fourth-quarter players to reduce their under-representation.

The additional on-ice time after high-school and before being drafted, together with positive peer effects and selection, concurs to shape relatively young players that throughout their NHL career perform better, are overrepresented in elite rosters (Gibbs et al., 2012), and earn more than relatively older players, despite being under-represented in NHL.

## REFERENCES

- Addona, V., Yates, P. A. (2010). A Closer Look at the Relative Age Effect in the National Hockey League. *Journal of Quantitative Analysis in Sports* 6(4) Article 9, 1-17.
- Adler, M., (1985). Stardom and talent. *The American Economic Review*, 75(1), 208-212.
- 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.
- Baker, J., Logan, J. A. (2007). Developmental contexts and sporting success: birth date and birthplace effects in national hockey league draftees 2000–2005, *British Journal of Sports Medicine*, 41(8), 515-517.
- Barnsley, R. H., Thompson, A. H., Barnsley, P. E. (1985). Hockey success and birthdate: The relative age effect. *Canadian Association of Health, Physical Education and Recreation (CAHPER) Journal* 51: 23–28
- 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.
- 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.
- Böheim, R., Lackner, M. (2012). Returns to education in professional football. *Economics Letters*, 114(3), 326-328.



- 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.
- 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.
- Currie, C., Nic Gabhainn, S., Godeau, E., Roberts, C., Smith, R., Currie, D., Pickett, W., Richter, M., Morgan, A., Barnekow, V. (2008). Inequalities in Young People's Health – International Report from the 2005/2006 Survey. WHO Regional Office for Europe.
- Deaner, R. O., Lowen, A., Cobley, S. (2013). Born at the Wrong Time: Selection Bias in the NHL Draft. *PLoS ONE*, 8(2), 1-7.
- Elder, T. E., Lubotsky, D. H. (2009). Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers. *The Journal of Human Resources* 44(3), 641-683.
- Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology* 38: 4–34.
- Fumarco, L., Rossi, G. (2015). Relative age effect on labor market outcomes for high skilled workers—evidence from soccer. *Working Papers in Management, University of London, Birkbeck, BWPMA 1501*.
- Garcia, M. S., Romero J. J. F., Lastra, D. F., Oliveira, G. E. (2014). Relative age effect in lower categories of international basketball. *International Review for the Sociology of Sport* 49(5), 526-535.

- 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.
- Gibbs, B. G., Shafer, K., Dufur, M. J. (2012). Why infer? The use and misuse of population data in sport research. *International Review for the Sociology of Sport*, 50(1), 115-121.
- Helsen, W. F., Baker, J., Schorer, J., Van Winckel, J., Williams, M. A. (2012). The relative age effect in European Professional Soccer: Is there any difference after ten years of research? *Journal of Sports Sciences* 30(15), 1665-1671.
- McCarthy, N., Collins, D., Court, D. (2016). Start hard, finish better: further evidence for the reversal of the RAE advantage. *Journal of Sports Sciences* 34(15), 1461-1465.
- Musch, J., Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review* 21, 147–167.
- Kahn, L. M. (2000). The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives*, 14(3), 75-94.
- Lê Cook, B., Manning, W. G. (2013). Thinking beyond the Mean: A Practical Guide for Using Quantile Regression Methods for Health Services Research. *Shanghai Archives of Psychiatry*, 25(1), 55-59.
- Nolan, J. E., Howell, G. (2010). Hockey success and birth date: the relative age effect revisited. *International Review for the Sociology of Sport*, 45(4), 507–512.
- 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.

- Peña, P. A. (2015). Creating winners and losers: date of birth, relative age in school, and outcomes in childhood and adulthood. Unpublished manuscript.
- 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.
- Roberts, S. J., Stott, T. (2015). A new factor in UK students' university attainment: the relative age effect reversal? *Quality Assurance in Education* 23(3), 295 – 305.
- Rosen, S. (1981). The economics of superstars. *The American Economic Review*. 71(5), 845-858.
- Saferanyos, S., Chittle, L., Horton, S., Dixon, J. (2016). An examination of the relative age effect and academic timing in cis volleyball. Proceedings of the SCAAPS 2016 Annual Conference *Journal of Exercise, Movement, and Sport* 48(1), 131.
- Sandercock, G. R.H., Taylor, M. J., Voss, C., Ogunleye, A. A., Cohen, D. D., Parry, D. A. (2013). Quantification of the Relative Age Effect in Three Indices of Physical Performance. *Journal of Strength & Conditioning Research* 27(12), 3293-3299.
- Schorer, J., Cobley, S., Busch, D., Bräutigam, H., Baker, J. (2009). Influences of competition level, gender, player nationality, career stage and playing position on relative age effects. *Scandinavian Journal of Medicine & Science in Sports* 19, 720-730.
- 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.

APPENDIX

Figure A.1 Quarter of birth distributions of free-agents.

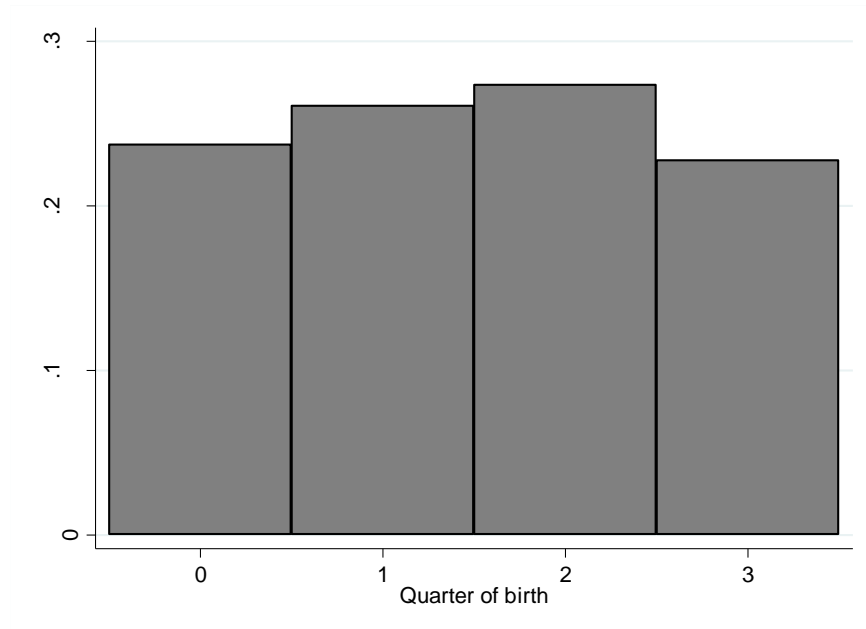
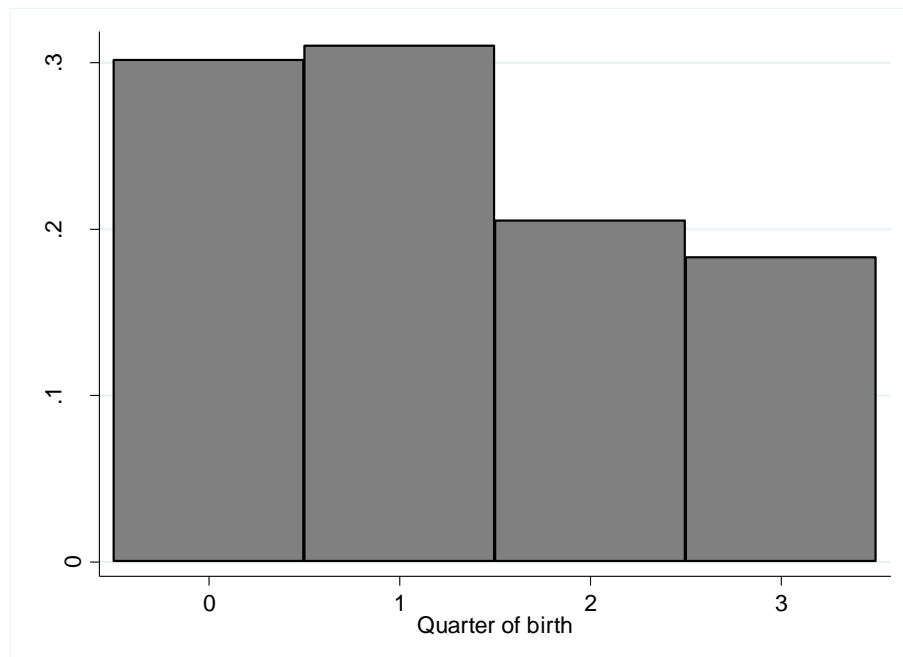


Figure A.2 Quarter of birth distributions of non-NHL players.



**Table A.1** RAE by quarter, on points; quantile regression.

Variables	North.-Am.	North.-Am.	North.-Am.	North.-Am.
	Points 25%	Points 50%	Points 75%	Points 90%
April-June	0.116 (0.540)	0.452 (1.444)	3.135 (2.751)	0.956 (2.737)
July-September	1.122 (0.727)	4.869** (2.036)	7.981*** (2.694)	6.333** (2.786)
October-December	1.819** (0.854)	6.546*** (1.988)	11.46*** (3.068)	9.222*** (2.885)
Control variables	Y	Y	Y	Y
Can. Jr. Hockey	N	N	N	N
Observations	4,447	4,447	4,447	4,447
Pseudo R-squared	0.058	0.120	0.132	0.114

Note: Only Northern-American players are investigated. In this investigation, repeated observations per player are used. All the control variables are included; no analysis on the sub-sample of only Canadians is carried out. Standard errors in parenthesis are clustered on players. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$