The Age Advantage in Youth Football

Stephen Lawrence , Laura Jonker and Jan Verbeek
steve.lawrence.ata@gmail.com laura@xoet.nl jan.verbeek@knvb.nl

Abstract

Average Team Age (ATA) and a Relative Age Index (RAEi) are variables against which performance outcomes in football can be measured and we consistently find that performance advantages are evident when measured against these variables. In 'The Age Advantage in Association Football', Lawrence, S., MSp2015, 6,389 matches played by males at U17 to adult team ages[*] were examined providing evidence of an age advantage. In this new paper analysis of additional match data from U12 to adult, providing a total dataset of 15,088 matches, is presented providing further insight into the development of relative age effects and signifying a causal connection between cut-off date eligibility rules and such effects. We conclude with a proposition that replacement of cut-off date rules with 'average team age' rules will assist with the elimination of such effects.

1 Introduction

For the purposes of this study the Average Team Age (ATA) is calculated as the mean of the chronological ages of the players composing the starting line-up in any given match.[†]

The Relative Age Index (RAEi) is calculated as the number of early-born players divided by the total number of players composing the starting line-up in any given match and is expressed as a decimal proportion between 0 and 1. Early-born players are those players born in the first six months of the competition year defined by the eligibility cut-off date. In our sample of youth players in the Netherlands this is January to June (following FIFA's eligibility rule for youth football).[‡]

We use home team advantage as a comparator variable as a well-known and graphic comparison to age advantage and relative age advantage. An expected ratio in adult professional football of 46% - 24% - 30%, for Home win – Draw – Away win, translates into **1.62** points per game (PPG) for the home team and **1.14** PPG for the away team. We use this PPG parameter as an expression of competitive advantage.

[*] Data courtesy of Gracenote.
[†] 'The Age Advantage in Association Football', Lawrence, S., MSp2015
[‡] 'The Age Advantage in Association Football', Lawrence, S., MSp2015

2 Methodology

The accumulated data including the anonymised team-sheet, match date, date of birth for each player and match result, was compiled into sets for each competition. The percentages of wins, draws and losses accruing to each team according to the three variables (ATA, RAEi & Home/Away) were calculated and then converted into PPG for comparison. For each competition an average PPG score for home teams, away teams, older teams, younger teams, more biased teams and less biased teams was thus established along with the average age of the teams. The competitions were then aggregated into 10 age groupings for clarity. At youth age levels the aggregated group data was drawn from the same competition across 2 seasons in 2013/14 and 2014/15. Figure 1 shows how the data has been aggregated into 10 groupings.

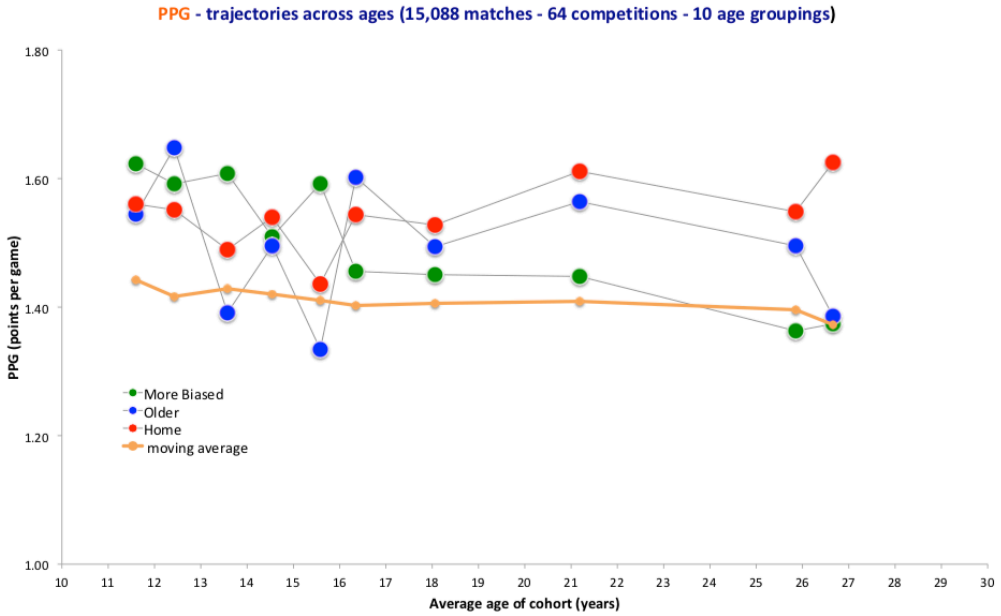

Figure 1. Graph of PPG accrued by the **Home** Team (red), **Older** Team (blue) and **More-biased** Team (green) with the moving average (calculated including away, younger and less-biased teams) shown (orange).

3 Comparing Home Team Advantage with Age Advantage and Relative Age Advantage

Three charts, each including moving average lines and peak height velocity (PHV) curves for reference purposes, show how our three chosen variables relate to the PPG parameter and the average cohort age.

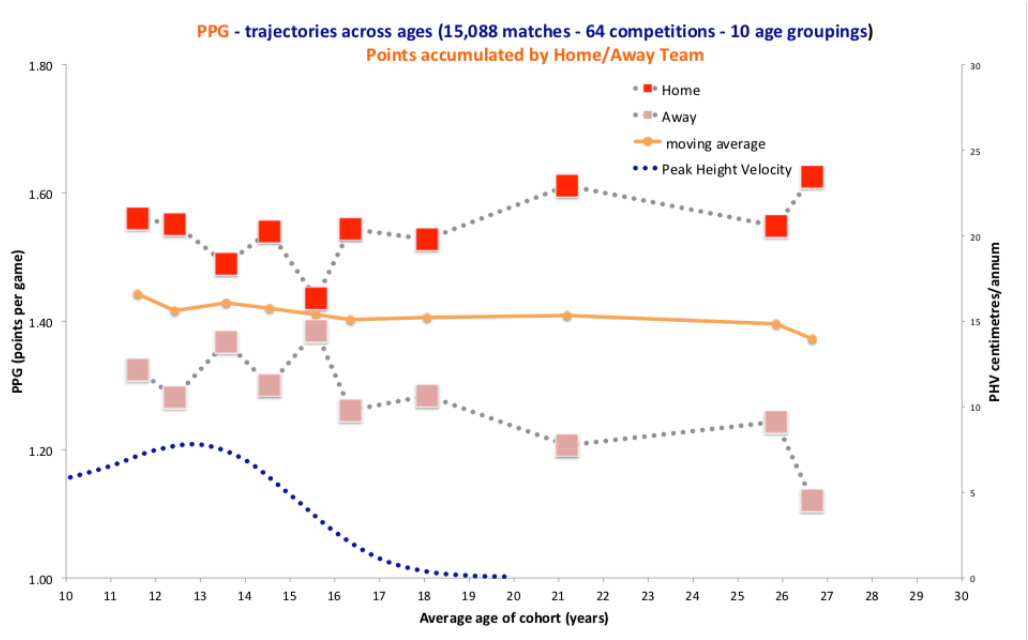

Figure 2. Graph of PPG accruing to the **Home** Team (red) or the **Away** Team (pink).

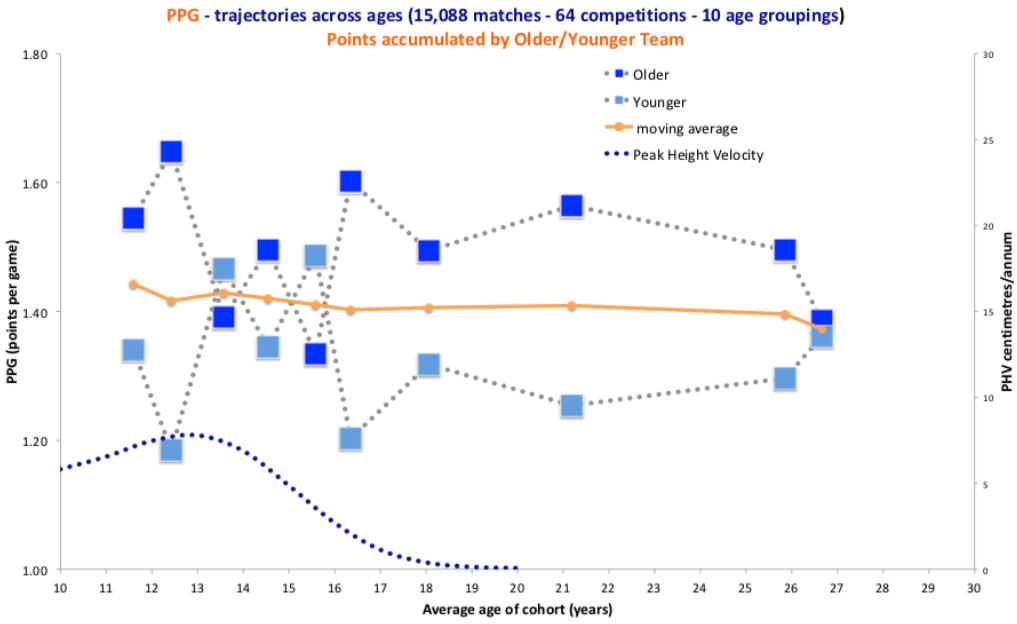

Figure 3. Graph of PPG accruing to the **Older** Team (blue) or the **Younger** Team (light blue).

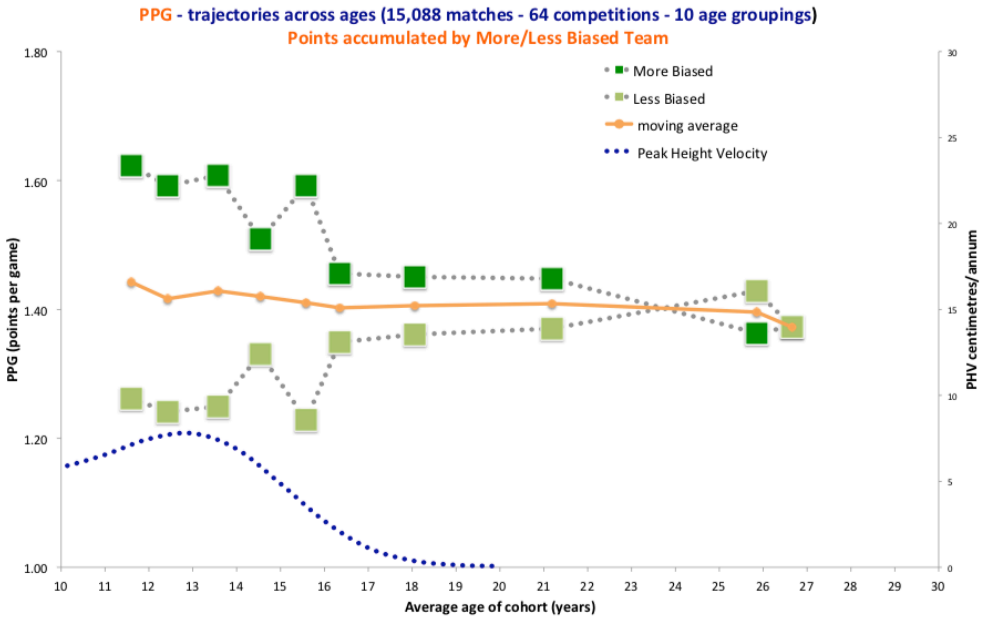

Figure 4. Graph of PPG accruing to the **More-biased** Team (green) or the **Less-biased** Team (light green).

Patterns of diminishing advantage for older or more biased teams, over time, are evident. No such diminution is evident in respect of home team advantage, which in our dataset is at its maximum at the cohort average age of 26.7. This corresponds with the age of minimum age advantage and minimum relative age advantage.

The p-values and chi-square values for the eight youth age groupings plus the semi-pro grouping are shown in Figure 5. The null hypotheses were that home or away teams, older or younger teams and more-biased or less-biased teams all had the same chance of winning matches.

Age Group	ATA	RAEi	n	Home/Away p-value	Home/Away chi-sq	ATA p-value	ATA chi-sq	RAEi p-value	RAEi chi-sq
Under 12	11.598	0.648	191	0.2485632482	1.331	0.3173105079	1.000	0.0646885438	3.413
Under 13	12.427	0.643	1,335	0.0003199830	12.950	0.0000000007	38.162	0.0000016930	22.915
Under 14	13.575	0.659	519	0.3194953451	0.991	0.5377235267	0.380	0.0013462033	10.278
Under 15	14.542	0.691	739	0.0179044401	5.605	0.1376077149	2.205	0.0559740195	3.653
Under 16	15.581	0.683	413	0.7038060054	0.145	0.2540516395	1.301	0.0024318454	9.191
Under 17	16.350	0.646	1,241	0.0002141461	13.703	0.0000001786	27.252	0.1248633764	2.355
Under 19	18.060	0.638	1,243	0.0014746617	10.110	0.0215535829	5.281	0.2003912073	1.640
U23 (2010-2013)	21.187	0.615	890	0.0000086877	19.780	0.0006502320	11.626	0.3424621581	0.901
Netherlands semi-pro	25.853	0.526	3,107	0.0000000002	40.592	0.0000327627	17.250	0.1331310074	2.256
Total			9,678						

Figure 5 Table of p-values and chi-square values for 8 youth age groupings and the semi-pro age grouping. All youth data derives from Netherlands[§] competition during seasons 2014/15 & 2015/16 except for U23 data which is from 2010/11, 2011/12 & 2012/13.

[§] Data courtesy of Koninklijke Nederlandse Voetbalbond.

The data also shows a steady decline in evident relative age bias, as measured by RAEi, as competition cohort ages increase, with parity being achieved in some competitions with cohort average ages of around 27 years. This is what we would expect if the importance of an age advantage and thus a relative age advantage diminished with increasing age. See Figure 6.

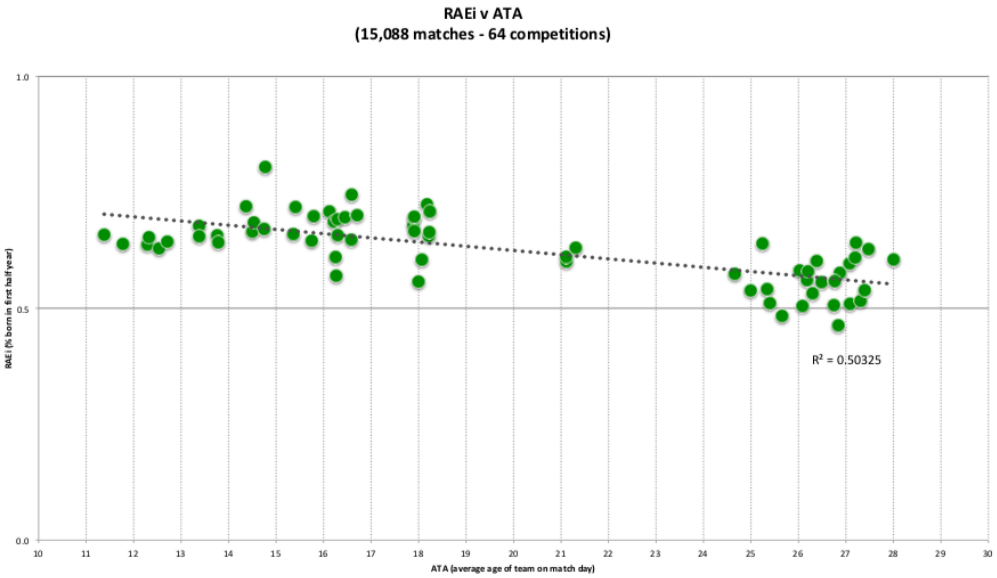

Figure 6. Graph of diminishing RAEi as the ATA of competition cohorts increases.

In respect of the ATA correlation (Figure 3) we observe volatility with higher p-values in the U12, U14 and U16 cohorts (Figure 5). We note that this occurs following a period of maximum peak height velocity and corresponds with a period of minimum evident mean age differences between competing teams (n.b. the mean team age difference is < 0.16 when RAEi is ~ 0.65). See Figure 7.

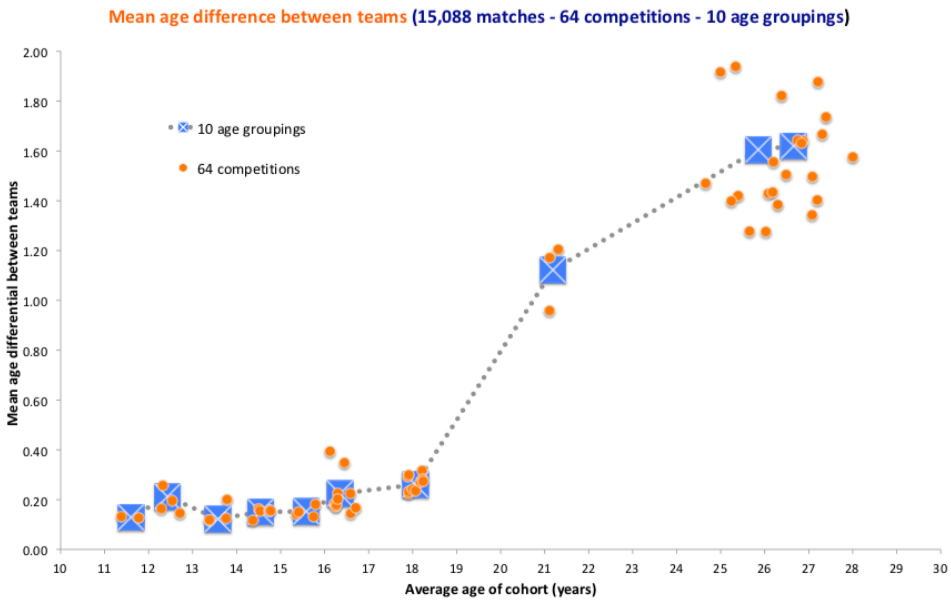

Figure 7. Graph of increasing mean age difference between teams as the average cohort age rises.

4 Conclusion

Our data shows that at youth ages, older football teams experience a competitive advantage when playing against younger teams (Figure 3). The data also shows that the advantage diminishes, as competition cohorts get older. This happens in the context of rigid competition eligibility age rules defined by a cut-off date. In order for any given team to be older than its competitor, it must necessarily consist of more players whose ages are closer to the cut-off date than its competitor. If that is the case the data must show a corresponding competitive advantage accruing to teams exhibiting a higher RAEi (Figure 4). This is precisely what the data shows. The advantage when measured as PPG is similar (as we would expect) for both higher average age and more biased relative age and both are similar to the known home team advantage which allows us to intuit its severity. Furthermore the evident relative age bias in competition steadily decreases as the average age of competition cohorts increases towards parity around the age of 27 when the average age advantage disappears (Figure 6).

The desire for competitive advantage in youth football drives up the average team age, which in turn, within eligibility cut-off date silos, causes relative age bias. The eligibility cut-off date rule is therefore causal in relation to the relative age effect and conversely it follows that removal of the cut-off date rule and replacement with an eligibility rule which disallows the possibility an older team being on the field would be causal in removing that effect.

Such an average team age (ATA) rule can be devised as follows: ‘A competing squad shall consist of no more than ‘X’ players whose average age on the competition start date shall be no more than ‘Y’. The average age of the starting team in any competition match shall be no more than ‘Y’. No player in the squad shall be more than ‘Z’ years older than the youngest player in the squad. The mean and the range of ages are thus defined on a team eligibility basis rather than on an individual eligibility basis allowing any individual player to participate across a spectrum of eligible age groups.