TI 2025-018/V
Tinbergen Institute Discussion Paper

On Bogey Teams and Circular Triads: Psychological Factors in Team Performance

Jan van Ours[1]

1 Erasmus University Rotterdam, Tinbergen Institute

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

On Bogey Teams and Circular Triads: Psychological Factors in Team Performance

Jan C. van Ours*

March 9, 2025

Abstract

Team performance depends not only on the individual abilities of its members and cooperation between them but also on psychological factors such as confidence, rivalry, and perceived pressure. In a regular work environment, it is challenging to isolate the contribution of psychological factors. Analyzing sports data can be insightful, as performance metrics are widely available. This paper focuses on two phenomena that highlight the impact of psychological factors: bogey teams and circular triads. A "bogey team" refers to a team that consistently outperforms another team, often defying expectations. Circular triads involve non-transitive match outcomes among trios of teams. Using balanced panel data from the top league of Dutch football, the analysis identifies both bogey team relationships and circular triads in match outcomes. Clearly, psychological factors play a significant role in shaping team performance.

Keywords: Non-transitivity, Circular triads, Bogey Teams, Football

JEL-codes: C25, D01, Z2

Conflict of interest: none

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

1 Introduction

The performance of individuals and teams depends not only on individual abilities but also on cooperation between team members and rivalry with other teams. Psychological factors, such as confidence and perceived pressure from management or competitors, can also have significant effects. In a regular work environment, assessing the importance of psychological factors is challenging because of the lack of informative data. Analyzing sports data offers valuable insights into the potential effects of psychological factors on performance. Sports performance data are widely available, and situations where psychological factors may play a role are relatively easy to identify.

In professional football, match outcomes largely depend on differences in team quality and, often, on home advantage. However, because football is a low-scoring sport (Scarf et al. (2019)), match outcomes are inherently uncertain. While individual matches are difficult to predict, over the long term, results typically align with expectations: high-quality teams are more likely to win against lower-quality teams, and home teams tend to have an advantage. This straightforward reasoning does not always hold. Just as psychological factors contribute to the home advantage, other, less-studied psychological influences may also shape match outcomes. These factors warrant further exploration to better understand their impact.

Home advantage in sports matches has been frequently studied, and various explanations have been proposed for why home teams are more likely than away teams to win a match.[1] These explanations include travel fatigue experienced by the away team, the home team's familiarity with the playing ground, and psychological factors associated with stadium crowd noise. Stadium crowds may influence match outcomes directly by affecting player performance or indirectly by influencing referee decisions. For instance, Nevill et al. (2002) found that video referees exposed to crowd noise in a taped football match awarded fewer fouls against home players than those who watched the same match without hearing the noise produced by the stadium crowd.

The impact of stadium crowds on home advantage has been explored in greater depth during the Covid-19 pandemic when stadiums were largely or fully empty due to public health restrictions. Studies such as Bryson et al. (2021) and Benz and Lopez (2023), conducted in twenty-three and seventeen professional football leagues respectively, found that stadium crowds frequently influenced referee decisions and home bias, though the extent of this effect varied across leagues. For the Dutch professional football league, Van Ours (2024b) showed that playing behind closed doors negatively impacted the performance

[1] See Peeters and van Ours (2021) for an analysis of developments in home advantage in English professional football over a period of more than 40 years.

of home teams. This suggests that home players were directly and adversely affected by the absence of psychological support from the stadium crowd. The psychological effects of stadium crowds extend beyond team performance to influence individual players. For example, Caselli et al. (2023) found that some players who faced racial harassment in regular times performed better during matches played without spectators under Covid-19 restrictions. This indicates that the absence of negative crowd behavior can positively affect individual performance. This type of influence of psychological factors on performance is found in other sports as well. Goller and Späth (2024), for example, found in ski-jumping and diving that positive feedback on past performance had a positive effect on subsequent performance.

Stadium crowds play a significant role in shaping psychological effects on performance in sports. However, the magnitude of these effects can vary across matches and situations. In some cases, unique pairwise relationships emerge between two clubs and their supporters, such as when one team is referred to as a "bogey team" of another. A bogey team is a team that, contrary to expectations, frequently performs well against a particular opponent, often defeating a stronger rival. The concept of bogey teams, also known as "angstgegner", is a topic of frequent debate in the popular press. Media discussions often highlight examples of bogey teams in Dutch professional football, including Utrecht against Ajax, Feyenoord against PSV, PSV against Ajax, and Feyenoord against PSV.[2] The phenomenon of bogey teams has received limited attention in academic literature. Rare exceptions are Hankin and Bunker (2016), who conducted a bogey team analysis of the Australian National Rugby League and Bunker et al. (2024), who showed that in professional tennis bogey relationships between players exist. The origins of the bogey team effect remain unclear, but psychological factors likely play a significant role in how these teams influence their rivals. For instance, players and supporters of a stronger team may feel additional pressure or develop a mental block when facing a known bogey team. Circular triads add another layer of complexity. In these triads, match outcomes form a non-transitive relationship: Team A often defeats Team B, Team B regularly beats Team C, and Team C consistently outperforms Team A. These circular patterns challenge traditional notions of team strength and match outcomes. Both phenomena are interesting and surprising but the overlap between the two is not. Persistent circular triads require bogey team relationships. One can imagine that team A often beats team B and team B frequently wins against C because of differences in strength. However, a circular triad requires teams C beating team A regularly. This will only happen if team C is the bogey team of team A because in terms of strength team C should be weaker than team A.

[2] See for example *Voetbal International* (www.vi.nl), 6 November 2020, 3 March 2024, 4 November 2015, and 24 February 2015, respectively.

Therefore, bogey team relationships and circular triads go hand in hand.

Circular triads are frequently studied in the context of preference orderings based on pairwise comparisons. For instance, individuals may prefer A to B, B to C, but C to A, creating a circular pattern of preferences. Edwards (1974) provided an extensive overview of inconsistencies in comparative judgments and their connection to circular triads. Non-transitivity is not limited to preferences but is also observed in games and competitive systems. LiCalzi (2016) discussed several classic examples of non-transitivity, including the rock-paper-scissors game, the magic square, the Condorcet paradox in voting theory, and Efron's non-transitive dice. Poddiakov (2025) expanded this discussion by providing numerous examples of non-transitivity across mathematics and the natural sciences. Non-transitivity has also been explored in sports contexts. Bozóki et al. (2016) identified triads of non-transitivity among male tennis players, while Temesi et al. (2024) found similar patterns among female tennis players. Both studies addressed the challenge of how to manage non-transitivity when rankings are required, highlighting its implications for competitive evaluations. Additionally, Spearing et al. (2023) developed a model that incorporated non-transitive outcomes in the analysis of baseball match data. In professional football, the analysis of non-transitivity is complicated by the possibility of draws, which can obscure the structure of circular triads. Bruss and Ferguson (2018) argued that a low number of circular triads suggests a lack of parity in skill levels among teams. In their analysis of the Greek football league, they used an alternative measure not reliant on circular triads to demonstrate that the league was unbalanced. In the Dutch professional football league, a long-standing pattern of non-transitive outcomes has been observed among its three major clubs. Van Ours (2024a) documented this phenomenon over several decades: Feyenoord consistently defeated PSV more often than expected, PSV frequently outperformed Ajax, and Ajax regularly prevailed against Feyenoord. These results suggest a circular, non-transitive relationship between the three clubs. Van Ours (2025) analyzed circular triads in the English Premier League focusing on the potential profits that non-transitive betting can provide.

The current paper follows up on Van Ours (2024a) demonstrating that bogey teams and circular triads of match outcomes are common phenomena in Dutch professional football, underscoring the significant role of psychological factors in team performance. The study makes several contributions to the literature on psychological effects in performance. First, it provides robust evidence for the existence of bogey teams: Over a period exceeding two decades, specific teams consistently outperformed particular rivals beyond what would be expected based on their relative strengths. Second, the paper identifies persistent circular triads in match outcomes creating a non-transitive pattern of results. A key finding of the study is the strong relationship between bogey teams and circular triads. Most bogey team dynamics are embedded within one or more circular triads, while the majority of circular triads involve at least one bogey team relationship. This interconnectedness highlights how psychological factors, such as mental blocks or heightened pressure, influence both team rivalries and broader patterns of match outcomes. Together, these findings provide compelling evidence of the psychological underpinnings of team performance, offering new insights into how mental and emotional factors shape competitive dynamics.

The structure of the paper is as follows. Section 2 provides an overview of the top tier of professional football in the Netherlands since the 2000/01 season, with a focus on a balanced panel of seven clubs. Section 3 examines the existence of bogey team dynamics, identifying several pairs of clubs where one team consistently outperforms the other beyond expectations. Section 4 analyzes non-transitivity in match outcomes, focusing on circular triads formed among the seven clubs. The analysis highlights that many of these circular triads are at least partially driven by bogey team relationships. Finally, Section 5 summarizes the key findings and concludes the paper.

2 The Dutch Eredivisie 2000/01-2023/24

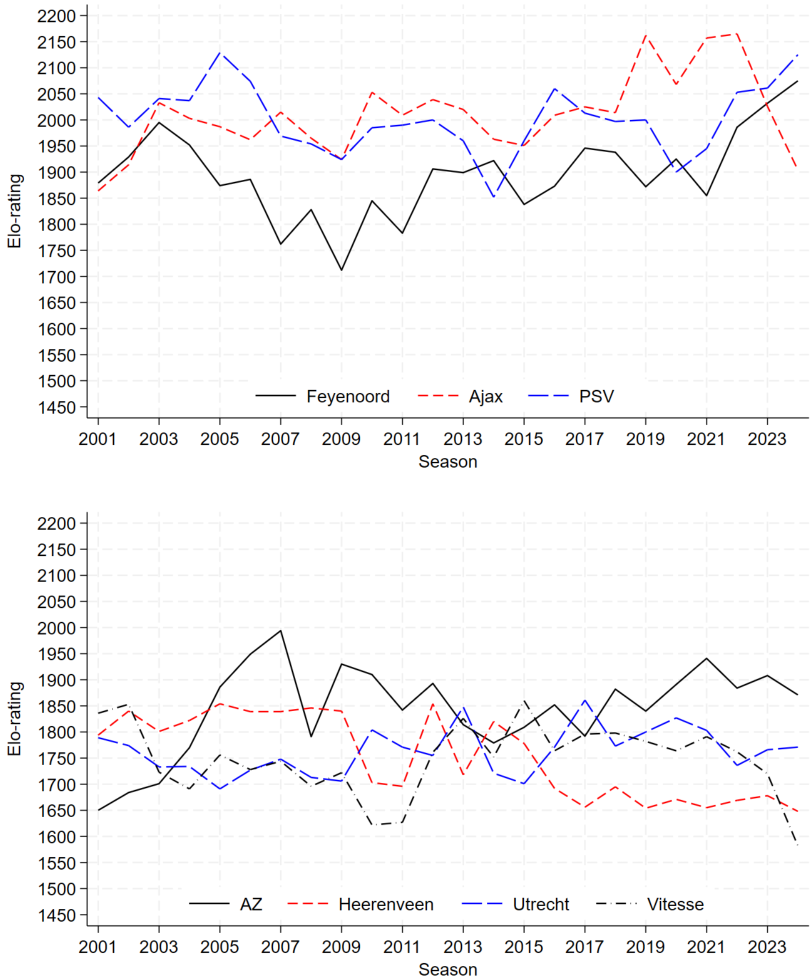
The top league of Dutch professional football is called "Eredivisie". The focus of the analysis is on the period 2000/01 to 2023/24. Season 2019/20 is excluded from the analysis since due to Covid-19 restrictions this season was stopped before all matches were played. The final ranking was based on the matches played at the time the competition was discontinued but no club was declared league champion. The analysis is based on a balanced panel of seven clubs that were present every season in the sample period.

The outcome of a match will depend to a large extent on the difference in strength between the two teams playing against each other. To provide an impression of differences in strength between the seven clubs Elo-ratings are used.[3] If two teams play against each-other the difference in Elo-ratings between the two teams provides an indication of the relative strength. Elo ratings are updated depending on the outcomes of matches played. So, over time an increase in Elo-rating indicates that a team got stronger. The absolute Elo-ratings do not have a particular meaning. A higher Elo-rating simply means a stronger team.

Figure 1 shows the evolution of the Elo-ratings over the sample period. Panel a shows

[3]Originally the Elo-rating system was designed for chess players (Elo (1978)) but nowadays it is used in many sports including professional football. Elo-ratings are available from various sources. The current paper uses data from elofootball.com which was also used by, for example, Gyimesi (2024). Alternative Elo-ratings are available from clubelo.com which were used by, for example, Csató (2024). Both studies provided more details about the set-up of these Elo-rating systems.

Figure 1: **Elo-ratings Seven Clubs; 2000/01 - 2023/24**



the Elo-ratings for the big three of the league. Over most of the sample period PSV and Ajax had the highest Elo-rating but after 2022 the Elo-rating of Ajax went down a lot such that in the last season of the sample Ajax was far below the two other clubs. Panel b of Figure 1 shows the Elo-ratings for the other four clubs. During most of the sample period, AZ had the highest ratings. Heerenveen had lower ratings at the end of the sample period than it had in the beginning. The Elo-ratings of Utrecht fluctuated with no clear trend. The same holds for Vitesse although in the last season there was a strong drop in the ratings and at the end of season 2023/24 Vitesse was relegated. The first column of Table 1 provides information about the average Elo-rating for every club over the period of analysis. On average, PSV and Ajax had the highest Elo-rating followed by Feyenoord. The Elo-ratings of the big three clubs were above average, the rest was below average. AZ had the fourth highest Elo-rating. Vitesse was at the bottom.

Sometimes, a weaker team defeats a much stronger team. That could be a one-off surprising match outcome. From a longer perspective, match outcomes are surprising if actual outcomes differ persistently from expected outcomes. Expectations about match outcomes can be based on bookmaker data. In betting markets participants are in general well-informed, motivated and experienced. Furthermore, news in sports is reported quickly and accurately so it is easy to take this information into account by participants when placing their bets and for bookmakers when setting their odds (Angelini and De Angelis (2019)).

There is randomness in the outcomes of football matches but in an efficient betting market where all relevant information is used expected outcomes based on bookmaker odds can be used to predict match outcomes. Strumbelj and Šikonja (2010), for example, examined six major European football leagues concluding that the effectiveness of bookmaker odds as forecasts increased over time. Nevertheless biases may occur. Probably the most frequently observed bias is the favorite-longshot bias which implies that bettors tend to overvalue underdogs and undervalue favorites. Because of this, bookmakers may offer lower returns on underdogs and higher returns on favorites. However, the effect of these biases on the use of bookmaker odds for forecasting match outcomes may be limited. Winkelmann et al. (2024), for example, provided an overview of 19 empirical studies on betting markets in the top five European football leagues concluding that inefficiencies exist but they do not occur persistently over time or across leagues.

Bookmaker odds can be used to generate expected probabilities of match outcomes using a simple normalization. For example, with decimal odds the expected probability of a win of home team i against away team j (ignoring a subscript for match date) is equal to:

$$p_{ij}^h = \frac{(1/O_{ij}^h)}{(1/O_{ij}^h) + (1/O_{ij}^d) + (1/O_{ij}^a)} = \frac{1}{O_{ij}^h \cdot (1 + B_{ij})} \tag{1}$$

where O_{ij}^h are the odds for a home win of i against j, O_{ij}^d are the odds for a draw and O_{ij}^a are the odds for an away win. B_{ij} is the bookmaker margin, which is equal to the sum of the inverse odds minus one (bookmaker margins cover costs of bookmakers and allow them to make a profit). The probability of a draw is equal to $p_{ij}^d = 1/(O_{ij}^d \cdot (1 + B_{ij}))$ while the probability of an away win is equal to $p_{ij}^a = 1/(O_{ij}^a \cdot (1 + B_{ij}))$. The expected number of points for the home team is equal to $(3.p_{ij}^h + p_{ij}^d)$ and for the away team the expected points are equal to $(3.p_{ij}^a + p_{ij}^d)$. Surprise match outcomes are defined as the differences between the actual and expected match outcomes. A surprise home win, for example, is defined as $S_{ij}^h = W_{ij}^h - p_{ij}^h$, where W_{ij}^h indicates whether or not team i won at home against team j.

In Table 1, the clubs are ordered from high to low by Elo-ratings but the ranking would be very similar if done by actual average points per game obtained by playing against the other six clubs. The second column shows that the range is from PSV that obtained 1.97 points per game to Vitesse that obtained 0.94 points per game. The third column shows the average expected points per game which has about the same ordering. The fourth column shows the number of surprise points per game – the difference between actual and expected points per game. In terms of surprises, PSV performed best with an average of 0.16 surprise points per match. Vitesse disappointed the most, with an average of -0.11 surprise points per match.

Table 1: **Elo-ratings, Points per Game and Wins; seven clubs 2000/01 – 2023/24**

		Points per game			Wins		
	Elo-rating	Actual	Expected	Surprise	Actual	Expected	Surprise
Ajax	2010	1.77	1.82	-0.05	0.52	0.53	-0.01
PSV	2002	1.97	1.81	0.16	0.58	0.52	0.06
Feyenoord	1896	1.60	1.55	0.05	0.45	0.43	0.02
AZ	1844	1.30	1.33	-0.03	0.36	0.36	0.00
Utrecht	1763	1.05	1.05	0.00	0.28	0.27	0.01
Heerenveen	1753	1.03	1.05	-0.02	0.25	0.26	-0.01
Vitesse	1748	0.94	1.05	-0.11	0.23	0.27	-0.04
Average/sum	1859	1.38	1.38	0.00	0.38	0.38	0.00

Note: Match outcomes from the seven teams playing against each other (276 games per team; 966 games in total). Note every match is included twice since in every match two teams play against each other. Covid-19 season 2019/20 excluded.

The last three columns of Table 1 show actual, expected and surprise wins. PSV was expected to win 52% of their matches, won 58% and therefore had an average surprise win of 6% per match. The actual wins are very much in line with the expected wins so the percentages of surprise wins are not very high, ranging from -4% to +6%.

3 Bogey teams

Table 2 shows the cumulative balances of expected wins, actual wins and surprise wins for each pair of clubs. With seven clubs there are 21 pairs of clubs. Since the focus is on the balances, the matrix is symmetric (outcome A-B = − outcome (B-A). Panel a gives the balances of expected wins. Over the period of analysis, Ajax and PSV had the highest cumulative balance of expected wins of about 80 while Feyenoord had a cumulative balance of expected wins of about 32. AZ had a cumulative balance of expected wins of -8.7 while the other clubs had a cumulative balance of expected wins of about -60.

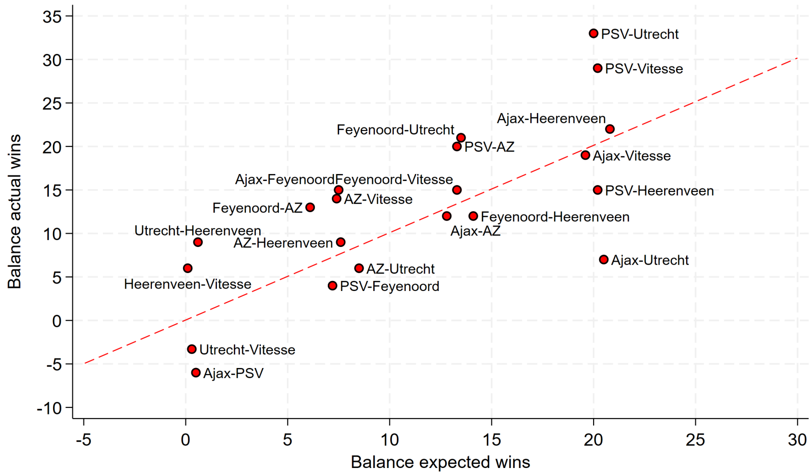
Table 2: **Cumulative Balance of Expected Wins, Actual Wins and Surprise Wins; 2000/01 – 2023/24**

a. Expected wins	AZ	Ajax	Feyenoord	Heerenveen	PSV	Utrecht	Vitesse	Total
AZ		-12.8	-6.1	7.6	-13.3	8.5	7.4	-8.7
Ajax	12.8		7.5	20.8	0.5	20.5	19.6	81.7
Feyenoord	6.1	-7.5		14.1	-7.2	13.5	13.3	32.3
Heerenveen	-7.6	-20.8	-14.1		-20.2	-0.7	0.1	-63.3
PSV	13.3	-0.5	7.2	20.2		20.0	20.2	80.4
Utrecht	-8.5	-20.5	-13.5	0.7	-20.0		0.3	-61.5
Vitesse	-7.4	-19.6	-13.3	-0.1	-20.2	-0.3		-60.9
Total	8.7	-81.7	-32.3	63.3	-80.4	61.5	60.9	0.0
b. Actual wins	**AZ**	**Ajax**	**Feyenoord**	**Heerenveen**	**PSV**	**Utrecht**	**Vitesse**	**Total**
AZ		-12	-13	9	-20	6	14	-16
Ajax	12		15	22	-6	7	19	69
Feyenoord	13	-15		12	-4	21	15	42
Heerenveen	-9	-22	-12		-15	-9	6	-61
PSV	20	6	4	15		33	29	107
Utrecht	-6	-7	-21	9	-33		-3	-61
Vitesse	-14	-19	-15	-6	-29	3		-80
Total	16	-69	-42	61	-107	61	80	0
c. Surprise wins	**AZ**	**Ajax**	**Feyenoord**	**Heerenveen**	**PSV**	**Utrecht**	**Vitesse**	**Total**
AZ		0.8	-6.9	1.4	-6.7	-2.5	6.6	-7.3
Ajax	-0.8		7.5	1.2	-6.5	-13.5	-0.6	-12.7
Feyenoord	6.9	-7.5		-2.1	3.2	7.5	1.7	9.7
Heerenveen	-1.4	-1.2	2.1		5.2	-8.3	5.9	2.3
PSV	6.7	6.5	-3.2	-5.2		13.0	8.8	26.6
Utrecht	2.5	13.5	-7.5	8.3	-13.0		-3.3	0.5
Vitesse	-6.6	0.6	-1.7	-5.9	-8.8	3.3		-19.1
Total	7.3	12.7	-9.7	-2.3	-26.6	-0.5	19.1	0.0

Panel b of Table 2 provides the cumulative balances of actual wins. This was by far the highest for PSV (107), the second highest for Ajax (69), and the third highest for Feyenoord (42). By far the lowest cumulative balance of actual wins was for Vitesse (-80). The most fascinating part of Table 2 is the overview of cumulative surprise wins in panel c. Overall, PSV had the best cumulative balance of surprise wins of about 27. Feyenoord was second best with a cumulative balance of about 10. Vitesse was by far the worst with a cumulative balance of almost -20 surprise wins but also Ajax had a large negative cumulative balance of surprise wins almost 13.

Figure 2 gives a graphical representation of the relationship between the balances of expected wins and the balances of actual wins. Some pairs are close to the diagonal but there are also pairs which have a similar balance of expected wins but a very different balance of actual wins. In terms of the largest number of cumulative surprise wins between two teams, Utrecht-Ajax and PSV-Utrecht stand out. In both pairs, the number of cumulative surprise wins is about 13. There is no clear pattern. Utrecht won much more often from Ajax than expected while PSV won much more often from Utrecht than expected. Feyenoord also won more often than expected from Utrecht while Ajax in its

Figure 2: **Balance of Expected Wins and Actual Wins by Pair of Teams**



Note: the dashed line represents the situations where the balance of expected wins is equal to the balance of actual wins.

turn won more often against Feyenoord than expected.

It should be noted that every separate match outcome is a surprise as the expected outcome based on bookmaker data is always based on a distribution of possible outcomes. Persistent surprise outcomes over a long period of time are needed to establish whether there are bogey teams or circular triads in match outcomes. To establish whether some teams are bogey teams for others the average balance of surprise wins per season between every pair of teams is calculated. If the average seasonal number of surprise wins is significantly different from zero the first team of a pair is considered to be a bogey team for the second team of that pair. Table 3 gives an overview.[4]

Table 3: **Bogey Team Relationships; Balance of Average Seasonal Surprise Wins; 2000/01-2023/24**

Bogey team	Dominated team	Balance of wins	
PSV	Ajax	0.28	(0.21)*
PSV	AZ	0.29	(0.22)*
PSV	Utrecht	0.56	(0.15)***
PSV	Vitesse	0.38	(0.18)**
Feyenoord	Utrecht	0.32	(0.20)**
Utrecht	Ajax	0.59	(0.27)**
Utrecht	Heerenveen	0.36	(0.27)*

Note: In parentheses standard errors; ***(**) indicates significance at the 1%(5%) level based on a one-sided t-test.

[4]Bunker et al. (2024) identified bogey effects in professional tennis applying Fischer's Exact Tests to contingency tables of expected wins and actual wins for matches between players. This approach cannot be used with football data because of the possibility that a match ends in a draw.

As shown there are seven pairs of clubs of which one is the bogey team for the other. PSV is bogey team for Ajax, AZ, Utrecht and Vitesse. Utrecht is bogey team for Ajax and Heerenveen while Feyenoord is bogey team of Utrecht. For three out of the seven pairs the balance of average seasonal surprise wins is somewhat imprecisely estimated (10% level of significance).

A bogey club is a team that wins more often than expected against a particular opponent, without any clear reason for this recurring advantage. It could stem from a long-standing rivalry where, for instance, Team A may consider games against Team B the highlights of their season, while Team B treats them as just another fixture. In the period analyzed, Utrecht's success against Ajax, outperforming expectations, indicates that Utrecht is Ajax's bogey club. However, Feyenoord is a bogey club for Utrecht. It is possible that Utrecht prioritizes winning against Ajax over Feyenoord, but that would not explain why Ajax consistently dominates Feyenoord in expected wins. If anything, the rivalry between Feyenoord and Ajax is likely the longest and most intense in Dutch football history. It is therefore difficult to believe that Ajax's dominance is due to its players valuing the match more than Feyenoord players value the match against Ajax.

In Table 3 there is no obvious circular triad. Feyenoord is the bogey team of Utrecht and Utrecht is the bogey team of Ajax but there is no significant relationship between Ajax and Feyenoord. Ajax has a positive balance of surprise wins against Feyenoord but this is not significantly different from zero. Nevertheless, there is some evidence of a circular triad Feyenoord-Utrecht-Ajax. Something similar holds for PSV-Utrecht and Utrecht-Ajax. But here, Ajax has a negative balance of surprise wins against PSV. So, there is no indication of a circular triad PSV-Utrecht-Ajax.

4 Circular Triads

If in season t team i plays at home against team j the outcome of the match, S_{ijt}, depends on the home advantage of team i, H_{it}, and the difference in qualities Q_{it} and Q_{jt} between the two teams (Clarke and Norman (1995)):

$$S_{ijt} = H_{it} + Q_{it} - Q_{jt}. \tag{2}$$

Similarly, assuming that team qualities and home advantages are constant within a season, if team j plays at home against team i, the match outcome is equal to:

$$S_{jit} = H_{jt} + Q_{jt} - Q_{it}. \tag{3}$$

This implies that from the perspective of team i, the net result R_{ijt} of team i playing against team j in season t is equal to:

$$R_{ijt} = S_{ijt} - S_{jit} = H_{it} - H_{jt} + 2 \cdot (Q_{it} - Q_{jt}). \tag{4}$$

Similarly, for matches between teams j and k and matches between teams k and i, the balances of the match results in season t are equal to:

$$R_{jkt} = S_{jkt} - S_{kjt} = H_{jt} - H_{kt} + 2 \cdot (Q_{jt} - Q_{kt}) \tag{5}$$

$$R_{kit} = S_{kit} - S_{ikt} = H_{kt} - H_{it} + 2 \cdot (Q_{kt} - Q_{it}). \tag{6}$$

Adding up Equations (5) to (7), it follows that the sum of the net balances is equal to 0:

$$R_{ijt} + R_{jkt} + R_{kit} = 0. \tag{7}$$

In other words, there is a transitive relationship in the match outcomes. The balance of the match outcomes between k and i is fully determined by the results between i and j and j and k.

Non-transitivity in match outcomes between the three teams can be introduced by assuming that between one of the pairs, for example k and i, there is an interaction-specific effect, $\theta_{kit} \neq 0$, Then,

$$R_{ijt} + R_{jkt} + R_{kit} = \theta_{kit}. \tag{8}$$

If match outcomes between i, j and k are non-transitive the sum of the balances of outcome is equal to θ_{kit}. This measure of non-transitivity can take any value. If it is not zero, it can be positive as well as negative, depending on the nature of the non-transitive relationship. Note that for every season it holds that $R_{ijt}+R_{jkt}+R_{kit} = -R_{jit}-R_{kjt}-R_{ikt}$. It could be that θ_{kit} is positive one season and negative in the next. The match outcomes of three teams are non-transitive over a period of time if over that period the average measure is significantly larger (or smaller) than zero. Because the measure is a sum of balances of outcomes it is not important whether outcomes are measured in points or wins since draws cancel out and therefore a points measure is simply three times a win measure.

According to Kendall and Smith (1940), with n teams, when n is odd, the number of circular triads ranges from 0 to $\frac{n^3-n}{24}$. Therefore, with seven teams the number of possible triads ranges from 0 to 14. To establish non-transitivity in match outcomes the focus is on surprise wins where two restrictions were used. First, each pairwise balance of surprise wins has to be positive. Second, the average seasonal sum of the balance of surprise wins of the three match outcomes has to be significantly different from zero (at least at a 10%-level). For every season, the average balance of surprise wins was calculated. Then, a one-sided t-test was used to establish whether or not the seasonal averages were positive and significantly different from zero. Imposing these restrictions, there were eight non-transitive triads shown in Table 4.[5]

Table 4: **Circular Triads; Average Seasonal Surprise Wins; 2000/01-2023/24**

			Average seasonal match surprise							
Team 1	Team 2	Team 3	Team 1 - 2		Team 2 - 3		Team 3 - 1		Sum	
Ajax	Feyenoord	Utrecht	0.32	(0.25)	0.32	(0.20)**	0.59	(0.27)**	1.23	(0.44)***
Heerenveen	PSV	Utrecht	0.22	(0.22)	0.56	(0.15)***	0.35	(0.27)*	1.15	(0.41)***
Feyenoord	Utrecht	Heerenveen	0.32	(0.20)**	0.36	(0.27)*	0.09	(0.21)	0.78	(0.42)**
Heerenveen	Vitesse	Utrecht	0.26	(0.22)	0.14	(0.26)	0.36	(0.27)*	0.76	(0.46)*
Ajax	Feyenoord	PSV	0.32	(0.25)	0.14	(0.28)	0.28	(0.21)*	0.74	(0.42)**
AZ	Ajax	Feyenoord	0.04	(0.23)	0.32	(0.25)	0.30	(0.24)	0.66	(0.35)**
AZ	Heerenveen	PSV	0.06	(0.26)	0.23	(0.22)	0.29	(0.22)*	0.58	(0.36)*
Ajax	Heerenveen	PSV	0.05	(0.26)	0.22	(0.21)	0.28	(0.21)*	0.56	(0.39)*

Note: Covid-19 season 2019/20 excluded from the analysis; ***(**,*) indicates significance at the 1%(5%,10%) level based on a one-sided t-test.

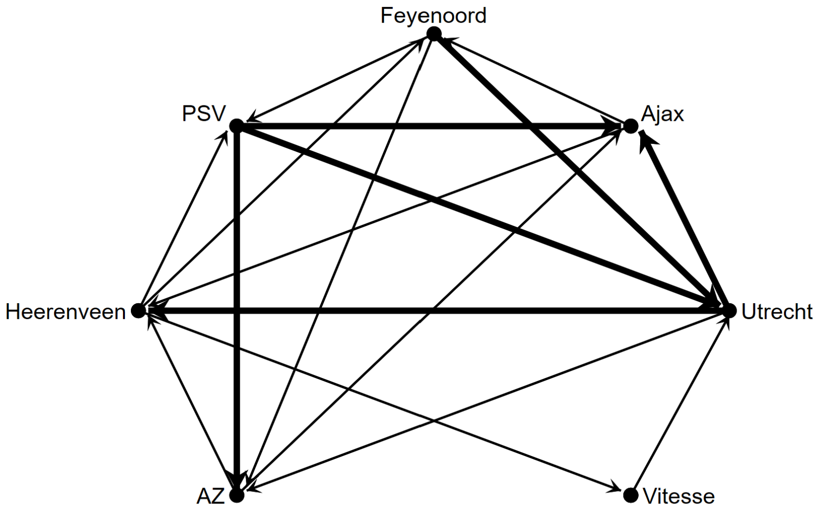
Heerenveen is in five triads, also because the pairs Utrecht-Heerenveen and Heerenveen-PSV are in three triads. Vitesse is present in only one triad. AZ is present in two triads. At least one of the top three of the league, Feyenoord, PSV, Ajax, is present in all but one triad, the triad of Heerenveen-Vitesse-Utrecht. In addition to the eight circular triads there is also a circular tetrad: Feyenoord–PSV–Utrecht–Ajax. In line with Kendall and Smith (1940), this circular tetrad consists of two circular triads, in this case: Feyenoord-PSV-Ajax and Feyenoord-Utrecht-Ajax.

Table 4 presents the average surprise wins over the sample period. Figure 3 gives a graphical representation of the patterns. In terms of significance levels of at least 10%, PSV has three outgoing arrows dominating Ajax, AZ and Utrecht. Utrecht has two incoming and two outgoing arrows: dominating Ajax and Heerenveen, being dominated by Feyenoord and PSV. Ajax has two significant incoming arrows being dominated by Utrecht and PSV. The figure also makes clear that there circular triads of which the sum of the surprise wins is not significantly different from zero and thus not reported in Table 4. For example the circular triads Feyenoord–AZ–Heerenveen. Similarly, there is the significant circular tetrad Feyenoord–PSV–Utrecht–Ajax and the insignificant circular tetrads AZ–Heerenveen–PSV–Utrecht and Heerenveen–Feyenoord–Utrecht–AZ.

Figure 4 shows the evolution over time of the cumulative surprise wins. Clearly,

[5] Appendix C provides information about expected wins, actual wins and surprise wins for each of the triads. Appendix D presents for each of the triads parameter estimates of ordered logit models with actual match outcomes as dependent variables. This appendix also shows estimates of models in which the expected number of points are dependent variables.

Figure 3: **Patterns of relationships between pairs of teams**

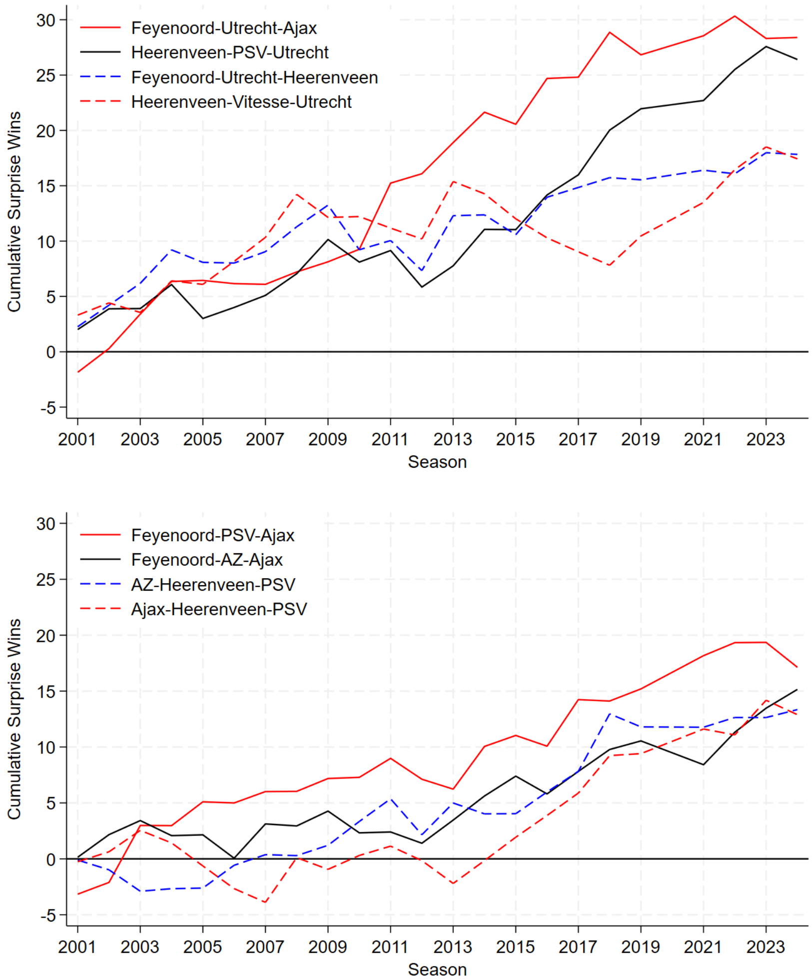


Note: The arrows go from the dominant teams to the dominated teams.
The thick arrows apply to significance levels of at least 10%.

for many triads in the first part of the sample period cumulative surprise wins hardly increased indicating that the average annual surprise win of each triad was close to zero. From 2010 to 2012 onward the cumulative balance of wins started to increase. So, the intransitivity of surprise match outcomes is in particular a phenomenon of the last 10 to 15 seasons. The origin of this phenomenon is not clear. As with all the circular triads and bogey team relationships there is no economical of psychological rational explanation. This also implies that in the near future some of these relationships may fade away.

Fanatic supporters or hooliganism do not seem to be an explanation. In his analysis of hooliganism in the Netherlands, Spaaij (2007) noted that four primary groups — Ajax, Feyenoord, Utrecht, and Den Haag — dominated incidents of football violence in the early stages. By the early 2000s, arrests primarily involved fans from Ajax, Feyenoord, Utrecht, and PSV. This overlap in major hooligan groups suggests intense rivalries between these teams. However, rivalry alone does not fully explain the observed non-transitive patterns in match outcomes. In each circular triad, one team consistently dominates another while being dominated by a third, creating a cyclical dynamic that points to factors beyond simple competitive rivalry.

Figure 4: **Cumulative Balance of Surprise Wins by Triad; 2000/01 - 2023/24**



5 Concluding Remarks

Match outcomes in professional football are subject to uncertainty. While stronger teams are generally expected to win more frequently against weaker teams over time, this is not always the case. This paper explores this discrepancy by analyzing the match outcomes of a balanced panel of seven clubs in the top football league of the Netherlands over a span of 23 seasons. The main conclusion is that surprising match outcomes persist between certain club pairings, and some of these surprises extend beyond individual pairings. The analysis identifies non-transitive circles, including examples involving three teams and one involving four teams. Notably, in almost every circle of non-transitivity, at least one of the "big three" clubs—Feyenoord, Ajax, or PSV—is involved.

While the exact causes of these non-transitive relationships remain unclear, psychological factors are likely significant contributors. One of the unresolved mysteries in sports economics concerns the home advantage in matches. It is widely believed that the presence of home crowds can influence player and team performance, either directly or indirectly, by affecting referee decisions in favor of the home team. Crowd noise may have varying effects on players: a hostile crowd can have a paralyzing effect, while an enthusiastic crowd can energize players to improve their performance. In both cases, the impact is psychological in nature.

The phenomenon of "bogey teams"—teams that consistently underperform against certain opponents—adds to the complexity of these psychological effects. Within the context of circular triads of match outcomes, these bogey team dynamics further contribute to the mystery of psychological influences on performance. Stadium crowds may be more hostile toward certain opponents and/or more enthusiastic in supporting their team against particular rivals. Bogey teams and circular triads are closely linked; almost all bogey teams are part of circular triads, and most circular triads include at least one bogey team relationship. Both phenomena—bogey teams and circular triads—are real and cannot be easily explained. They occur in competitive environments where team strength and quality matter, and psychological factors such as confidence and rivalry influence match outcomes.

With seven teams, the theoretical maximum number of circular triads is 14. Finding eight circular triads is surprising and shows that non-transitivity is not a rare phenomenon. Their persistence over time suggests that psychological factors can have a lasting impact. This influence is shaped by the dynamic interactions between stadium crowds, fans, and teams. Investigating the origin of bogey team pairings and circular triads in relation to the psychological processes is an unexplored yet potentially fascinating area for future research. Understanding these phenomena could provide insights into improving regular production processes in other fields where the performance of workers and teams are similarly affected by psychological factors.

Acknowledgment
The author thanks two reviewers for their helpful comments on a previous version of the paper.

Data availability
The data and codes used in this paper will be made available through a public database.

Declaration of competing interest
The author declares that he has no known competing interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Angelini, G. and L. De Angelis (2019). Efficiency of online football betting markets. *International Journal of Forecasting 35*(2), 712–721.

Benz, L. S. and M. J. Lopez (2023). Estimating the change in soccer's home advantage during the Covid-19 pandemic using bivariate Poisson regression. *AStA Advances in Statistical Analysis 107*(1-2), 205–232.

Bozóki, S., L. Csató, and J. Temesi (2016). An application of incomplete pairwise comparison matrices for ranking top tennis players. *European Journal of Operational Research 248*(1), 211–218.

Bruss, F. and T. Ferguson (2018). Testing equality of players in a round-robin tournament. *Mathematical Scientist 43*, 125–136.

Bryson, A., P. Dolton, J. J. Reade, D. Schreyer, and C. Singleton (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. *Economics Letters 198*, 109664.

Bunker, R. P., C. Yeung, and K. Fujii (2024). An expected wins approach using Fisher's exact test to identify the bogey effect in sports: An application to tennis. *The Journal of Sport and Exercise Science 8*(1), 43–64.

Caselli, M., P. Falco, and G. Mattera (2023). When the stadium goes silent: How crowds affect the performance of discriminated groups. *Journal of Labor Economics 41*(2), 431–451.

Cattelan, M., C. Varin, and D. Firth (2013). Dynamic Bradley–Terry modelling of sports tournaments. *Journal of the Royal Statistical Society, Series C Applied Statistics 62*(1), 135–150.

Clarke, S. R. and J. M. Norman (1995). Home ground advantage of individual clubs in English soccer. *The Statistician 44*(4), 509–521.

Csató, L. (2024). Club coefficients in the UEFA Champions League: Time for shift to an Elo-based formula. *International Journal of Performance Analysis in Sport 24*(2), 110–134.

Edwards, A. L. (1974). Circular triads, the coefficient of consistence, and the coefficient of agreement. In G. Maranell (Ed.), *Scaling: A Sourcebook for Behavioral Scientists (1st ed.)*, Chapter 7, pp. 98–105. Routledge.

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

Goller, D. and M. Späth (2024). 'Good job!' The impact of positive and negative feedback on performance. *Sports Economics Review 8*, 100045.

Gyimesi, A. (2024). Competitive balance in the post-2024 Champions League and the European Super League: A simulation study. *Journal of Sports Economics 25*(6), 707–734.

Hankin, R. and R. Bunker (2016). Bogey teams in sport. Technical report, Auckland University of Technology.

Hvattum, L. and H. Arntzen (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting 26*(3), 460–470.

Kendall, M. G. and B. B. Smith (1940). On the method of paired comparisons. *Biometrika 31*(3/4), 324–345.

LiCalzi, M. (2016). Sic transit, et non. In M. Emmer, M. Abate, M. Falcone, and M. Villareal (Eds.), *Imagine Math 5: Between Culture and Mathematics*, pp. 183–195. Springer.

Nevill, A. M., N. J. Balmer, and A. M. Williams (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise 3*(4), 261–272.

Peeters, T. and J. C. van Ours (2021). Seasonal home advantage in English professional football; 1974-2018. *De Economist 169*, 107–126.

Poddiakov, A. (2025). Are mathematicians, physicists and biologists irrational? Intransitivity studies vs. the transitivity axiom. *Human Arenas*, forthcoming.

Sauer, R. D. (1998). The economics of wagering markets. *Journal of Economic Literature 36*(4), 2021–2064.

Scarf, P., R. Parma, and I. McHale (2019). On outcome uncertainty and scoring rates in sport: The case of international rugby union. *European Journal of Operational Research 273*(2), 721–730.

Spaaij, R. (2007). Football hooliganism in the Netherlands: Patterns of continuity and change. *Soccer & Society 8*(2-3), 316–334.

Spearing, H., J. Tawn, D. Irons, and T. Paulden (2023). Modeling intransitivity in pairwise comparisons with application to baseball data. *Journal of Computational and Graphical Statistics 32*(4), 1383–1392.

Strumbelj, E. and M. R. Šikonja (2010). Online bookmakers' odds as forecasts: The case of European soccer leagues. *International Journal of Forecasting 26*(3), 482–488. Sports Forecasting.

Temesi, J., Zs. Szádoczki, and S. Bozóki (2024). Incomplete pairwise comparison matrices: Ranking top women tennis players. *Journal of the Operational Research Society 75*(1), 145–157.

Van Ours, J. C. (2024a). Non-transitive patterns in long-term football rivalries. *Journal of Sports Economics 25*(7), 802–826.

Van Ours, J. C. (2024b). They didn't know what they got till the crowd was gone. *Economics Letters 236*, 11615.

Winkelmann, D., M. Ötting, C. Deutcher, and T. Makarewicz (2024). Are betting markets inefficient? Evidence from simulations and real data. *Journal of Sports Economics 25*(1), 54–97.

Appendix A: Details about the data

Table A.1 gives an overview of the number of seasons each of the clubs was present in the top league in the period of analysis. As shown, seven clubs were present all the time, two clubs missed out for one season, two clubs for four seasons and so on.

Table A.1: **Number of Seasons Eredivisie by club; 2000/01 – 2023/24**

N	Club
23	AZ, Ajax, Feyenoord, Heerenveen, PSV, Utrecht, Vitesse
22	Groningen, Twente
19	Willem II, NEC
17	Heracles, NAC Breda, Roda JC
16	Den Haag
15	RKC Waalwijk
13	Sparta Rotterdam
12	Excelsior, PEC Zwolle
10	Graafschap
8	VVV Venlo
7	Fortuna Sittard
6	Go Ahead Eagles
5	Cambuur, RBC
4	Volendam
3	Emmen
2	Den Bosch
1	Almere City

Note: N = number of seasons; season 2019/20 not included.

The data are from the following sources:

1. Elo-ratings: elofootball.com
2. Match outcome data: www.football-data.co.uk
3. Bookmaker odds: www.football-data.co.uk. Odds were primarily drawn from William Hill. For 17 of the 966 matches William Hill odds are missing and supplemented by odds from Interwetten.

Appendix B: Bogey team relationships

To formalize the relationship between pairs of teams, ordered logit models are estimated in which the match results in terms of number of points obtained by the home team (0, 1, 3) is the dependent variable and the expected number of points by the home team based on bookmaker odds is among of the right-hand side variables:

$$\Pr(Y_{ijt} \leq z) = \frac{1}{1 + \exp(-(\delta_z + \beta E_{ijt} + \gamma_j \sum_{j \neq i} d_j))}, \qquad z \in \{0, 1, 3\} \tag{9}$$

where the δ_z represent threshold parameters such that $-\infty < \delta_0 < \delta_1 < \delta_3 = \infty$. Furthermore, E_{ijt} represents the expected number of points in the match between i and j in season t and the d's are dummy variables representing the opposing teams.

Table B.1: **Bogey Team Relationships; Parameter Estimates Home Points (Ordered Logit Models); 2000/01-2023/24**

	AZ		Ajax		Feyenoord		Heerenveen		PSV		Utrecht		Vitesse	
a. Expected points	1.68	(0.31)***	1.40	(0.23)***	1.97	(0.30)***	1.50	(0.24)***	1.91	(0.22)***	1.47	(0.26)***	1.66	(0.24)***
b. Expected points	1.36	(0.42)***	1.81	(0.53)***	1.94	(0.48)***	1.79	(0.49)***	0.96	(0.59)	0.51	(0.51)	0.81	(0.48)*
AZ		–	-0.16	(0.32)	0.35	(0.32)	-0.01	(0.31)	0.76	(0.44)*	-0.19	(0.33)	-0.50	(0.33)
Ajax	-0.01	(0.30)		–	-0.32	(0.29)	-0.02	(0.50)	0.30	(0.29)	-0.01	(0.46)	-0.57	(0.50)
Feyenoord	-0.40	(0.29)	0.31	(0.28)		–	0.22	(0.34)	-0.02	(0.34)	-0.89	(0.36)**	-0.43	(0.34)
Heerenveen	0.07	(0.29)	-0.05	(0.50)	-0.30	(0.34)		–	0.07	(0.49)	0.46	(0.31)	-0.31	(0.25)
PSV	-0.50	(0.37)	-0.30	(0.28)	0.25	(0.34)	0.49	(0.42)		–	-1.64	(0.44)***	-1.10	(0.48)**
Utrecht	-0.08	(0.31)	-0.89	(0.46)*	0.24	(0.35)	-0.40	(0.30)	1.30	(0.47)***		–	0.16	(0.28)
Vitesse	0.32	(0.30)	-0.14	(0.50)	-0.02	(0.33)	0.33	(0.26)	1.05	(0.55)*	-0.14	(0.28)		–

Note: Threshold parameters not reported. Expected points based on bookmaker data; 46 matches (two matches per season). Covid-19 season 2019/20 excluded from the analysis; ***(**,*) indicates significance at the 1%(5%,10%) level based on robust standard errors.

Bookmakers are expected to set their odds such that all information relevant to match outcomes is included (Sauer (1998)). Indeed, panel a of Table B.1 shows that the parameter estimates with only the expected number of points are all highly significant. Introducing dummy variables for opponent teams in panel b sometimes reduces this significance.

In panel b of Table B.1 five of the seven parameter estimates of expected points are significantly related to actual points. If the expected number of points based on bookmaker odds is included then match-pair dummy variables should not have an effect. However, the parameter estimates of the effects of the opponent dummy variables are sometimes significant. For AZ, Feyenoord and Heerenveen none of the opponent variables have a significant effect. This indicates that the outcomes of their matches against the other teams in the sample are not significantly affected by their opponents to the extent that these effects are not already absorbed by the expected number of goals.

The other parameter estimates suggest that bogey team relationships are present. It appears that Utrecht is a bogey team for Ajax although this only appears in the estimates for Ajax. The effects of the opponent dummies are conditional on the expected points which are assumed to have the same effect for all opponent teams. Since this is likely not the case the results do not need to be symmetrical.

PSV is a bogey team for AZ, Utrecht and Vitesse. Finally, Feyenoord is a bogey team for Utrecht. The main differences between these estimates and those presented in Table 3 in the main text is that both PSV being a bogey team for Ajax and Utrecht being a bogey team for Heerenveen are not present in panel b while they are in Table 3 although significant only at a 10%-level.

Appendix C: Circular triads: Expected wins, actual wins and surprise wins

Table C.1: **Cumulative Balance of Expected Wins, Actual Wins and Surprise Wins by Triad; 2000/01 – 2023/24**

a. Expected wins						Total
Feyenoord-Utrecht	13.5	Utrecht-Ajax	-20.5	Ajax-Feyenoord	7.5	0.5
Heerenveen-PSV	-20.2	PSV-Utrecht	20.0	Utrecht-Heerenveen	0.7	0.5
Feyenoord-Utrecht	13.5	Utrecht-Heerenveen	0.7	Heerenveen-Feyenoord	-14.1	0.1
Heerenveen-Vitesse	0.1	Vitesse-Utrecht	-0.3	Utrecht-Heerenveen	0.7	0.5
Feyenoord-PSV	-7.2	PSV-Ajax	-0.5	Ajax-Feyenoord	7.5	-0.2
AZ-Ajax	-12.8	Ajax-Feyenoord	7.5	Feyenoord-AZ	6.1	0.8
AZ-Heerenveen	7.6	Heerenveen-PSV	-20.2	PSV-AZ	13.3	0.7
Ajax-Heerenveen	20.8	Heerenveen-PSV	-20.2	PSV-Ajax	-0.5	0.1
b. Actual wins						
Feyenoord-Utrecht	21	Utrecht-Ajax	-7	Ajax-Feyenoord	15	29
Heerenveen-PSV	-15	PSV-Utrecht	33	Utrecht-Heerenveen	9	27
Feyenoord-Utrecht	21	Utrecht-Heerenveen	9	Heerenveen-Feyenoord	-12	18
Heerenveen-Vitesse	6	Vitesse-Utrecht	3	Utrecht-Heerenveen	9	18
Feyenoord-PSV	-4	PSV-Ajax	6	Ajax-Feyenoord	15	17
AZ-Ajax	-12	Ajax-Feyenoord	15	Feyenoord-AZ	13	16
AZ-Heerenveen	9	Heerenveen-PSV	-15	PSV-AZ	20	14
Ajax-Heerenveen	22	Heerenveen-PSV	-15	PSV-Ajax	6	13
c. Surprise wins						
Feyenoord-Utrecht	7.5	Utrecht-Ajax	13.5	Ajax-Feyenoord	7.5	28.5
Heerenveen-PSV	5.2	PSV-Utrecht	13.0	Utrecht-Heerenveen	8.3	26.5
Feyenoord-Utrecht	7.5	Utrecht-Heerenveen	8.3	Heerenveen-Feyenoord	2.1	17.9
Heerenveen-Vitesse	5.9	Vitesse-Utrecht	3.3	Utrecht-Heerenveen	8.3	17.5
Feyenoord-PSV	3.2	PSV-Ajax	6.5	Ajax-Feyenoord	7.5	17.2
AZ-Ajax	0.8	Ajax-Feyenoord	7.5	Feyenoord-AZ	6.9	15.2
AZ-Heerenveen	1.4	Heerenveen-PSV	5.2	PSV-AZ	6.7	13.3
Ajax-Heerenveen	1.2	Heerenveen-PSV	5.2	PSV-Ajax	6.5	12.9

Note: Covid-19 season 2019/20 is excluded from the analysis.

The pairwise information about balances of expected wins, actual wins and surprise wins is presented in Table 2. Table C.1 provides an overview of the same information grouped by circular triad and ordered by the magnitude of the balance of surprise wins. Panel a of Table C.1 shows the balance of expected wins for each pair of teams and for the triad. Whereas for some of the expected win balances can be very positive (PSV-Utrecht +20) or very negative (Utrecht-Ajax –20) for a triad the total balance of expected match wins is close to zero. In other words, according to the bookmaker based expected wins all triads have transitive match outcomes.

Panel b shows the balance of actual wins. Also here, there as substantial balances of wins that are very positive (PSV-Utrecht +33) or very negative (Heerenveen-PSV – 15). Overall, for each triad the balance of actual wins is positive ranging from +13 for Heerenveen-PSV-Ajax to +29 for Feyenoord-Utrecht-Ajax.

Panel c shows the balance of surprise wins. As shown over the 23 seasons in the sample, Ajax beats Utrecht 7 times more than the other way around. However, the expectation was that Ajax would have beaten Utrecht about 20 times more than the other way around. Therefore, Ajax has about 13 less surprise wins against Utrecht than the other way around. Similarly, PSV was expected to win 20 times against Utrecht while in fact PSV won 33 times against Utrecht out of the 46 matches they played against each other in the sample period. Therefore, PSV had 13 surprise wins over Utrecht. The balance of surprise wins for each of the triads is positive and of about the same magnitude as the actual wins.

Appendix D: Modeling Circular Triads

In the seminal Bradley-Terry model, the differences in skills or quality teams possess determine the probability that one team beats another team (Bradley and Terry (1952)). The model for three teams i, j, and k can be specified as a logit model. The pairwise win probabilities are determined by the difference in quality Q between the two teams and an interaction term θ for one of the pairs which allows for non-transitivity of match outcomes (Spearing et al. (2023)):

$$p_{ij} = \frac{e^{Q_i - Q_j}}{1 + e^{Q_i - Q_j}} \qquad p_{jk} = \frac{e^{Q_j - Q_k}}{1 + e^{Q_j - Q_k}} \qquad p_{ik} = \frac{e^{Q_i - Q_k + \theta_{ik}}}{1 + e^{Q_i - Q_k + \theta_{ik}}} \tag{D.1}$$

The third probability can be expressed as a function of the first two probabilities:

$$p_{ik} = \frac{p_{ij} \cdot p_{jk}}{(1 - p_{ij}) \cdot (1 - p_{jk}) \cdot e^{-\theta ik} + p_{ij} \cdot p_{jk}} \tag{D.2}$$

Without the interaction term θ_{ik}, the Bradley-Terry model has cardinal transitivity since in that case p_{ik} is fully determined by p_{ij} and p_{jk}. If $\theta_{ik} \neq 0$, p_{ik} is not fully determined by the other two probabilities and strong transitivity is rejected. If θ_{ik} is negative (positive) this increases (decreases) the probability that team i beats team k relative to the effect of the difference in quality.

In football a match ends in a home win, an away win or a draw. With three possible match outcomes transitivity is not so clear and a binomial logit model does not cover possible outcomes. Therefore, an ordered logit model is used. Cattelan et al. (2013), for example, applied an ordered logit model to the 2008–2009 Italian Serie A football.

For a triad of teams i, j and k the result of a match between home team a and away team b in season t, Y_{abt}, is either 3 points for a home win, 1 point in case of a draw and 0 points in case of a victory of the away team:

$$\Pr(Y_{abt} \leq z) = \frac{1}{1 + \exp(-(\delta_z + \gamma \Delta E_{abt} + Q_a - Q_b + \theta_{ik}))}, \quad z \in \{0, 1, 3\} \tag{D.3}$$

where $a, b = i, j, k$ and $a \neq b$. Furthermore, the δ_z represent threshold parameters such that $-\infty < \delta_0 < \delta_1 < \delta_3 = \infty$, ΔE_{abt} represents the average difference in Elo-ratings between team a and b in season t and Q_k is normalized to zero. Note that it would be straightforward to introduce home advantage by adding an additional parameter but since in none of the estimates home advantage had a significant effect home advantage is ignored. By adding the seasonal difference in Elo-ratings the model allows for the possibility that the quality of a team varies between the seasons. If over the seasons the relative quality of three teams does not change E_{abt} is perfectly correlated with $Q_a - Q_b$.

Hvattum and Arntzen (2010) concluded that Elo-ratings are a useful measure of the strength of professional football teams. This is confirmed in the first column of Table D.1. In every estimate, the difference in Elo-ratings between two teams has a significant and positive parameter estimate. Whether conditional on the Elo-rating difference the qualities of the teams have significant effects on match outcomes varies a lot. However, it is difficult to interpret these team quality estimates as part of the differences in quality is picked up by the difference in Elo-ratings. The effect of the interaction term between the first and the third team in a circular triad is always negative and significant (at at least the 10%-level). This suggests that transitivity of match outcomes does not hold.

To investigate whether the non-transitivity materializes in expected home wins similar equations are estimated with a transformed dependent variable:

$$\log\left(\frac{p^e_{abt}}{1 - p^e_{abt}}\right) = \eta^e + \gamma^e \Delta E_{abt} + Q^e_a - Q^e_b + \theta^e_{ik} + \varepsilon_{abt} \tag{D.4}$$

where p^e_{ab} represents the expected points of home team a playing against away team b. And, as before, $a, b = i, j, k$ with $a \neq b$ and ΔE_{abt} represents the average difference in

Elo-ratings team a and b in season t and Q_k^e is again normalized to zero. In the second column of Table D.1 none of the interaction terms is significantly different from zero. This indicates that in the odds of the bookmakers cardinal transitivity in match outcomes is assumed.

Table D.1: **Parameter estimates home points (ordered logit) and expected points (linear); circular triads**

	Actual points		Expected points	
a. Feyenoord-Utrecht-Ajax				
Quality Ajax	0.50	(0.60)	0.30	(0.06)***
Quality Feyenoord	0.36	(0.36)	0.23	(0.04)***
ΔElo-rating	0.59	(0.21)***	0.16	(0.02)***
Interaction Ajax-Utrecht	-1.56	(0.51)***	-0.02	(0.04)
b. Heerenveen-PSV-Utrecht				
Quality PSV	1.43	(0.61)**	0.38	(0.05)***
Quality Utrecht	0.53	(0.35)	0.02	(0.03)
ΔElo-rating	0.53	(0.18)***	0.12	(0.02)***
Interaction PSV-Heerenveen	-1.94	(0.56)***	-0.02	(0.04)
c. Feyenoord-Utrecht-Heerenveen				
Quality Feyenoord	0.97	(0.44)**	0.27	(0.04)***
Quality Utrecht	0.46	(0.30)	0.02	(0.03)
ΔElo-rating	0.47	(0.15)***	0.15	(0.01)***
Interaction Feyenoord-Heerenveen	-1.00	(0.51)*	-0.01	(0.04)
d. Heerenveen-Vitesse-Utrecht				
Quality Heerenveen	0.51	(0.39)	0.01	(0.04)
Quality Vitesse	0.22	(0.30)	0.02	(0.03)
ΔElo-rating	0.41	(0.16)***	0.20	(0.02)***
Interaction Heerenveen-Utrecht	-0.97	(0.50)*	-0.02	(0.05)
e. Feyenoord-Ajax-PSV				
Quality PSV	0.25	(0.45)	0.04	(0.04)
Quality Ajax	-0.04	(0.33)	0.06	(0.03)*
ΔElo-rating	0.75	(0.18)***	0.17	(0.02)***
Interaction PSV-Feyenoord	-0.90	(0.51)*	0.01	(0.05)
f. Feyenoord-AZ-Ajax				
Quality Ajax	0.34	(0.53)	0.12	(0.05)
Quality Feyenoord	0.40	(0.36)	0.09	(0.03)***
ΔElo-rating	0.79	(0.18)***	0.20	(0.01)***
Interaction Ajax-AZ	-0.95	(0.52)*	-0.03	(0.05)
g. AZ-Heerenveen-PSV				
Quality PSV	-0.26	(0.55)	0.31	(0.04)***
Quality AZ	-0.17	(0.30)	0.12	(0.02)***
ΔElo-rating	0.78	(0.15)***	0.15	(0.01)***
Interaction PSV-Heerenveen	-0.95	(0.55)*	-0.02	(0.04)
h. PSV-Ajax-Heerenveen				
Quality PSV	0.08	(0.58)	0.30	(0.04)***
Quality Ajax	-0.21	(0.49)	0.31	(0.04)***
ΔElo-rating	0.63	(0.15)***	0.15	(0.01)***
Interaction PSV-Heerenveen	-0.92	(0.54)*	-0.00	(0.04)

Note: Actual points: ordered logit models (equation (D.3)) threshold parameters not reported. Expected points: parameter estimates log-odds regression (equation (D.4)); every panel based on 138 matches (six matches per season); in parentheses robust standard errors; ***(**) indicates significance at the 1%(5%) level.