How outcome uncertainty, loss aversion and team quality affect stadium attendance in Dutch professional football

Lucas M. Besters[1]
Jan (J.C.) van Ours[2]
Martin A. van Tuijl[3]

[1] Rabobank
[2] Erasmus School of Economics
[3] Tilburg University

How outcome uncertainty, loss aversion and team quality affect stadium attendance in Dutch professional football

1 November, 2018

Abstract

We investigate stadium attendance in the highest level of Dutch professional football for the seasons 2000/01 – 2015/16 focusing on outcome uncertainty, loss aversion and team quality. We find that for individual football matches, attendance is related to reference-dependent preferences with loss aversion dominating the preference for uncertain outcomes. Furthermore, team quality is an important determinant of stadium attendance. Towards the end of the season, outcome uncertainty regarding the final ranking becomes important. For this seasonal uncertainty, we find a positive and stable, but rather small impact of the introduction of a unique and large end-of-season play-off scheme for the qualification for European football.

1. Introduction

Attendances at professional (team) sport events are a popular research area. A central topic is the relation between the uncertainty of the outcome of a contest and consumer demand. Rottenberg (1956) and Neale (1964) were the first to formulate the well-known uncertainty of outcome hypothesis (henceforth UOH). They suggest that attending a match is more attractive if the outcome is uncertain. The concept has been introduced for single matches, referring to *match uncertainty*. Two other types of outcome uncertainty are used in sports research as well (e.g., Cairns, Jennett and Sloane, 1986; Borland and Macdonald, 2003). *Seasonal uncertainty* is the uncertainty related to some end-of-season outcome, such as winning a league, promotion or relegation. *Long-run uncertainty* refers to the (lack of) dominance of certain teams during a considerable number of seasons. The UOH is often related to the concept of competitive balance, as proposed by Rottenberg (1956). No universally accepted definition of this concept exists. It generally relates to the degree in which competitors (such as sports teams) are balanced in terms of resources, quality, and talent etc. The *ex-ante* outcome of a match or competition between fairly equal competitors is more uncertain than the outcome of a contest between rather unequal competitors. If consumers, i.e. sports fans, derive utility from outcome uncertainty, a more balanced competition will attract more attendants. Therefore, sports bodies have an incentive to increase competitive balance, as they want to attract attendants in order to serve their members. Rules and regulations, such as salary-caps/wage-bill caps and talent allocation schemes (drafts), which are fairly common for team sports in the US, may be used to achieve this goal. In Europe, sports bodies are more reluctant to apply such restrictive regulations, since they are frequently bound by both domestic and European labor legislation, the latter primarily aimed at the free movement of labor.

Despite a large body of research, it is still not clear whether outcome uncertainty matters for attendance in professional football. Pawlowski (2013) provides an overview of studies concerning football attendance showing mixed evidence with respect to the UOH. His review is dominated by studies that focus on match uncertainty and seasonal uncertainty. Schreyer, Schmidt and Torgler (2016) give an overview of studies on match uncertainty only. Few studies cover long-run uncertainty. Studies generally use match-level data from a limited number of seasons in a single

country, and investigate whether stadium attendance depends on the uncertainty of outcomes.[1] Several indicators of uncertainty have been used and have found to be statistically significant (Borland and Macdonald, 2003; Pawlowski, 2013; Schreyer, Schmidt and Torgler, 2016). Nevertheless, the debate with regard to the UOH is still going on.

Recently, behavioral economic principles and decision making under uncertainty were introduced in empirical studies on attendance (see Budzinski and Pawlowski (2017) for an overview). Coates and Humphreys (2012) discuss the role of loss aversion from prospect theory (Kahnemann and Tversky, 1979) for attendance demand in the National Hockey League (NHL). Furthermore, Coates, Humphreys and Zhou (2014) develop a model of attendance demand that includes loss aversion combined with reference-depended preferences as described by Koszegi and Rabin (2006). They find that attendance is a function of the home win probability and its squared value. In their model, a concave relation between the home win probability and attendance is in line with the classical UOH. It emerges as a special case within the model, where fans prefer tighter matches above certain home wins. For this to happen, the marginal utility of attending an unexpected win has to be at least as large as the marginal utility of an unexpected loss. A convex relation suggests that fans are loss averse. In that case, fans value home wins and the potential to attend an upset, i.e. a home win in case the home team is expected to lose. Fans attend such upsets, if the expected utility of this unlikely event outweighs the utility of attending a home loss in a relatively uncertain match. With controls for several match- and team-characteristics, such as team quality, an empirical test with data from the Major League Baseball (MLB) suggests a convex relation and, thus, the rejection of the UOH (Coates, Humphreys and Zhou, 2014). Humphreys and Zhou (2015) extend this model with a league standing effect, i.e. a type of seasonal uncertainty. Furthermore, they argue that a convex relation between the home win probability and attendance does not rule out the existence of fan preferences for uncertain matches. It only means that loss aversion dominates the preference for uncertainty. Humphreys and Zhou (2015) show that the structural model contains three parameters of interest: home win preference, preference for outcome uncertainty and loss aversion. However, in the reduced form of the model only two parameters can be estimated, i.e. parameters related to the home win probability and the squared value of the home win probability. Home win preference is present if the sum of these parameters is larger than zero.

---

[1] Over time, an increased number of studies have looked at the demand for TV audience (e.g. Forrest, Simmons and Buraimo, 2005; Buraimo and Simmons, 2015; Cox, 2015). In general, such data is not publicly available and, therefore, rather difficult to obtain. In the present study, we focus on stadium attendance.

It is impossible to separately identify preference for outcome uncertainty and loss aversion. One can only observe which of these effects dominates. Cox (2015), for eight seasons of English Premier League football, and Martins and Cró (2016), for four seasons of Portuguese first division football, are examples of studies that empirically test the relationship between stadium attendance and the home win probability and its squared value. Both studies find that loss aversion dominates the preference for uncertain outcomes. Furthermore, the sum of the relevant parameters suggests the presence of a home win preference.

The abovementioned studies use objective measures to investigate fans' preferences regarding the uncertainty of the outcome of a match, such as the home win probability. In a recent study, Pawlowski, Nalbantis and Coates (2018) use survey data from German football fans to investigate how these objective measures of uncertainty relate to the perception of uncertainty by fans. Their results reveal that perceived game uncertainty is comparable to such objective measures of match uncertainty and fans exhibit loss aversion. In a similar way, they conclude that measures of seasonal uncertainty reflect perceived suspense. The concept of suspense itself, and its difference from surprise, has been modeled theoretically by Ely, Frankel and Kamenica (2015). In their model, new information becomes available over time. Suspense is experienced *ex ante* and is high if this new information is about to resolve crucial uncertainty. In contrast, surprise is experienced *ex post* and manifests if new information significantly alters someone's beliefs. Although the paper often refers to occasions in sports, such as minute-by-minute football scores and point-by-point tennis scores, it does not directly relate to attendance demand. However, as pointed out by Pawlowski, Nalbantis and Coates (2018), the notion of suspense could be interesting in that respect. Without discussing this in detail, it seems that suspense as formulated by Ely, Frankel and Kamenica (2015) can be related to seasonal uncertainty, and in particular to the uncertainty with respect to some end-of-season achievement. Their notion of surprise seems less related to the attendance demand studies as discussed before, since these mainly deal with the *ex-ante* decision to attend a match or not.

In our study, we use data from the highest level of Dutch professional football, covering the seasons 2000/01 - 2015/16. Dutch professional football clubs have only rarely been among the leading teams within European club competitions in the past two decades. Therefore, the interest in the Dutch league from abroad might be modest. Still, the league remains very popular within the Netherlands, with high average attendance numbers. Furthermore, over the period of analysis,

average Dutch attendance is substantial compared to other European leagues. Usually, the Dutch league is ranked just after the 'big five'. For example, in the season 2015/16, the highest football league in Germany attracted 43,500 attendants per match, which was about 36,500 in England, 28,500 in Spain, 22,000 in Italy, and 21,000 in France. The highest football league in the Netherlands attracted about 19,500 attendants per match, which is substantially more than in Belgium, the next in line, where on average about 12,000 people visited the football stadium in the season 2015/16 (source: www.european-football-statistics,co,uk). Interestingly, the Dutch league has been the only (big) European league that introduced end-of-season play-offs for the qualification for European football, amongst others to increase interest and attendance. More background information on these play-offs and the Dutch league follows in the next sections.

We follow a similar approach as recent previous studies and investigate the UOH in multiple ways. We use both a measure of match uncertainty and one concerning seasonal uncertainty and empirically test the consumer choice models of Coates, Humphreys and Zhou (2014) and of Humphreys and Zhou (2015). As an alternative for the home win probability, we introduce the expected number of points for the home team. This expected number of points is based on bookmaker odds and takes into account the possibility of a draw, which is the result of approximately one out of four matches in Dutch football. Measuring the expected number of points in this way has been used as an expectations-based reference point in football (Bartling, Brandes and Schunk, 2015). However, to the best of our knowledge, it has never been used before in attendance demand studies. Seasonal uncertainty is measured following Jennett (1984), who uses the significance of a match in relation to some end-of-season outcome. We adapt Jennett's significance measure to use it for winning a league as well as for other end-of-season outcomes, including the qualifications for (end-of-season play-offs for) European football competitions or relegation matches.[2]

Our results provide evidence for the rejection of the classical UOH related to match uncertainty. We find a convex specification for both the home win probability as well as our match-expectation variable. This suggests that fans exhibit reference-dependent preferences and loss aversion. Team

---

[2] Note that Humphreys and Zhou (2015) include a league standing effect as seasonal uncertainty. We deviate from their model, because a league standing effect is calculated per matchday within a league. Therefore, it is not club specific and this complicates the use of threshold amounts of points related to club specific end-of-season achievements. A match significance indicator such as the one by Jennett (1984) deals with these thresholds rather easily.

performance and team quality are important for the determination of stadium attendance. For seasonal uncertainty, many results are in line with the UOH. Of particular interest is the impact on attendance during regular league matches of the introduction of a large play-off scheme in Dutch professional football in the season 2005/06. In line with the UOH, one would expect a positive effect on attendance demand (see Bojke, 2007). We find a significant, though fairly small effect.

Our paper adds to the existing literature in multiple ways. First, we introduce the expected number of points as a new measure for match uncertainty. This measure includes the probability of a draw, which is important in football, and appears to outperform other measures. Furthermore, we use a new way to measure the quality of teams. Bookmaker odds are used to calculate the expected number of points in the previous matches. This measure of quality contains expectations and has its own contribution in addition to pre-match rankings. Finally, with respect to stadium attendance, we evaluate the introduction of the play-off scheme in the season 2005/06 empirically, contributing to the discussion on how to organize competitions in professional football taking into account the notion of suspense.

The structure of our paper is as follows. Section 2 provides an overview of the data we use in our analysis and some background information on Dutch professional football. Subsequently, section 3 discusses the set-up of our analysis. Section 4 presents our parameter estimates while section 5 focuses on the impact of the play-offs on stadium attendance. Section 6 concludes.

2. Our data

2.1 Dutch professional football

There are 18 clubs in the top league in Dutch professional football, the so-called *Eredivisie*. Every club plays all of the 17 adversaries twice, once at home and once away. Thus, each club plays 34 regular league matches during a season. Clubs obtain three points for a victory, a single point for a draw, and nil points for a loss. Direct relegation into the second tier, the so-called *Eerste Divisie*, is the consequence of the bottom position. The clubs ranked 16th and 17th have to play promotion/relegation play-offs with six clubs from this *Eerste Divisie*. The play-off scheme for qualification for European football was introduced in the 2005/06 season. As mentioned in the previous section, amongst others to attract extra attendance. After the introduction, many final rankings made a club qualify for some (end-of-season) competition. The champion always directly qualified for the UEFA Champions League. The effects of finishing at one of the positions between

2nd and 13th changed multiple times. Mainly, because the Intertoto Cup was abandoned by the UEFA (Union of European Football Associations) after the 2006/07 season and the KNVB canceled the play-offs for Champion League football after the 2007/08 season. The design of the play-offs always remained the same. They typically involve four clubs. The 'winner' qualifies, while without play-offs, the best ranked team would have qualified. Thus, especially for the qualification for (the end-of-season play-offs for) UEFA Europa League football not much changed after the play-offs were introduced (see also Appendix A).[3]

The stadiums are generally quite crowded during matches. Occupancy rates are approximately 80-90 percent throughout the sample period. The majority of attendants consist of season ticket holders. Most of these regular attendants are very loyal and, thus, a major part of the fan base and stadium attendants. They generally buy their season ticket before the start of the competition for a pre-announced fee. Single-match tickets are generally sold in the weeks prior to a match. Although the prices for these tickets are not necessarily fixed during the season, they have to be announced some period (i.e. weeks) in advance of the match. Furthermore, it seems reasonable to assume that the public knows what pricing strategies are used by clubs and that these strategies are rather constant throughout a season.[4]

2.2 Stadium attendance
We use match data from the highest tier of Dutch professional football for the seasons 2000/01 – 2015/16. For each match we have information on the attendance, general weather conditions during the match-day, the competing teams, including the location (stadium and stadium capacity, city, province), as well as fixed betting odds (see Appendix A for details). In every season, at least one, but maybe two or even three teams were relegated, depending on the results of the end-of-season promotion/relegation play-offs, while a similar number of teams were promoted. In total, 28 different teams were active in the *Eredivisie* during the sample period.

Figure 1 shows the development of the average attendance rate, average match attendance and average stadium capacity within our sample. Averages are calculated per season for all 18 clubs within the *Eredivisie*. It follows that the attendance rate is quite high in all seasons and moves

---

[3] Additional information about Dutch professional football is provided in Van Ours and Van Tuijl (2016).
[4] There are also business-club members (sponsors) and attendants with free tickets, mostly children from local amateur clubs and primary schools. Our data does not allow distinguishing between the various groups of attendants but, presumably, these groups are quite small.

around 0.88. Average attendance increased from approximately 15,500 in 2000/01 to 19,500 in 2015/16. In the same period, average stadium capacity increased from approximately 19,500 to almost 22,000. In general, these graphs develop in a similar way. Any difference in development is reflected in the attendance rate. Although the changes in this rate seem quite large, one should note that the scale only covers the range from 0.84 to 0.92. Furthermore, one should note that season-by-season changes in stadium capacity might result from actual changes in the capacity, because of renovations/expansions or the move to a new stadium, which generally happens in between seasons. Also, these changes may arise because of changes in the composition of clubs due to promotion and relegation. Certain clubs have bigger (smaller) stadiums than others. Correspondingly, these clubs might also have a bigger (smaller) fan base, which is reflected in the average attendance figures. In general, Figure 1 suggests that people's interest in live stadium attendance increased during the sample period, while the attendance rate remained high and stable.

Figure 1: Average match attendance rate (left hand scale), average match attendance and average stadium capacity (right hand scale); seasons 2000/01 – 2015/16
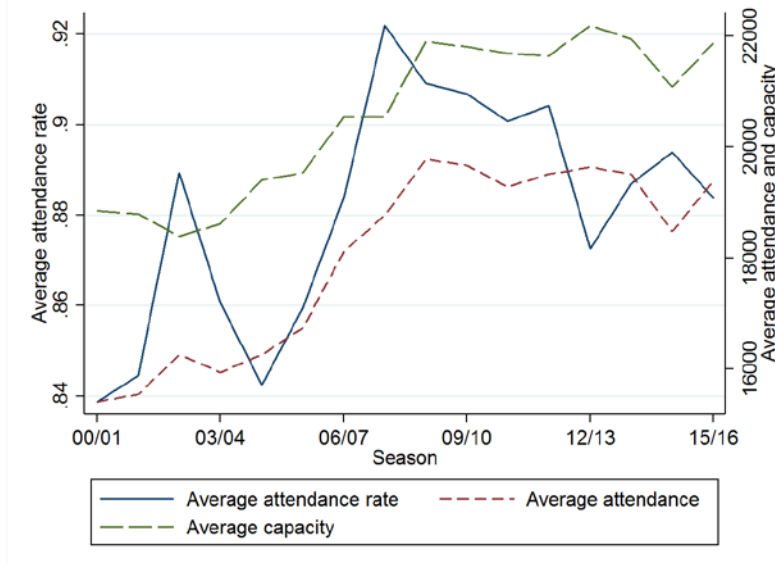
3. Set-up of the analysis

Stadium attendance is our focus variable. In line with many previous studies, we take the natural logarithm of the attendance number as dependent variable. However, in the analysis, we have to consider stadium capacity as well. In quite a few matches, the stadium was a sell-out. Since the formal stadium capacity may slightly differ for security reasons we assume a stadium to be full at

95 percent of the official capacity. This results in about 40 percent of the observations being censored. Obviously, stadium capacity differs between clubs and over time. Our approach should also account for this flexibility. Therefore, we set upper limits at 95 percent of capacity and allow differences per observation (in line with the differences in capacity). Then, we use a Tobit model with flexible upper limits to establish the determinants of the natural logarithm of attendance $\ln A_{ist}$. So:

$$
\ln A_{ist} = \begin{cases} \ln A_{ist}^{*} & \text{if } \ln A_{ist}^{*} < \ln(0.95 \cdot cap_{st}) \\ \ln(0.95 \cdot cap_{st}) & \text{if } \ln A_{ist}^{*} \geq \ln(0.95 \cdot cap_{st}) \end{cases} \tag{1}
$$

with

$$
\ln A_{ist}^{*} = z_{ist}'\beta + \varepsilon_{ist}
$$

where $\ln A_{ist}^{*}$ refers to latent attendance, i denotes the club, j indicates the match, while t refers to the season. Furthermore, z_{ist} represents a vector of club- and match-characteristics, β denotes a vector of parameters, while ε_{ist} denotes the error term. Our vector of club- and match-characteristics contains the following variables:

Natural logarithm of stadium capacity: This variable is usually constant for a club within a season. Only in 5 club-season combinations out of a total of 270 we observe an increase in stadium capacity, while for one club-season there is a decrease in capacity. Between-season changes in capacity are more common. This happens in 34 cases if we only count the changes for clubs that played both the season before and after the change on the highest level. In general, bigger, more popular and richer clubs have bigger stadiums. Thus, this variable (at least partly) reflects certain club and club-season specific aspects.

Average number of attendances in previous season: This variable captures habit persistence of fans of the home team and is constant within a season.

Average number of attendances for the opponent team in previous season. This variable captures some club specific elements of the opponent, such as attractiveness.

Team Quality: For team quality, we introduce a new measure, defined as the sum of the expected number of points in the previous 34 matches. Expectations are based on bookmaker odds. These odds reflect the relative strength or quality of the competing teams. They also include match-specific aspects such as home advantage, current form and player injuries that may influence the

expected match result. Therefore, inference of the quality of a team, based on the expectations of a single match (or a small number of matches) would be wrong. For example, given home advantage, the quality of the home team would be biased upwards, while the quality of the away team would be biased downwards. To correct for these match specific factors that may bias the inferred quality, we consider the previous 34 matches of a team for the construction of the quality indicator. Thus, the measure includes expectations that are based on all sorts of matches, such that match-specific elements cancel out against each other. For example, the number of home matches will be about equal to the number of away matches. Furthermore, the variable contains matches against strong teams as well as weak opponents. Although not all of the impact of the match-specific elements may disappear, the descriptive statistics suggest that the proposed indicator measures quality in line with expectations based on league-tables. For example, the average quality corresponds with an average number of points for an average league rank. Furthermore, the variable contains an arbitrary element, since we use the 34 previous matches without any weight factors, but it seems to work quite well in our attendance demand model. It allows for a continuum of match-by-match differences and an easy comparison between clubs and seasons. Note that, because of the *Team Quality* variable, we lose the 2000/01 season in our analysis, since we lack odds prior to this season.

Cumulative Surprise: The actual performance of a team may differ from the quality of the team as measured through bookmaker odds. To take this into account, we use the cumulative surprise over the previous 34 matches as a measure of actual performance. The variable is equal to the sum of the differences between the actual number of points obtained and the expected number of points, based on bookmaker odds, i.e. our team quality measure.

Home win probability: In line with the consumer choice models proposed by Coates, Humphreys and Zhou (2014) and Humphreys and Zhou (2015), we include the *Home win probability* and its squared value.

Match-Expectation: As an alternative to the home win probability, we introduce a new measure of match uncertainty, i.e. the expected number of points for the home team. We also use the squared value. With this, we attempt to include, in an easy and straightforward way, the probability of a draw. This is important since approximately 25 percent of the matches end in a draw. In comparison to the *Home win probability*, *Match-Expectation* adds the probability of a draw, which is weighted by one-third of the probability of a home win, because teams earn three points for a

victory and one for a draw. Note that the correlation between *Home win probability* and *Match-Expectation* is very high.[5] If the probability of a home win is high, so is match-expectation; if the probability of a home win is low, so is match-expectation. The difference between the two measures is driven by the probability of a draw.

Promoted: Dummy variable indicating whether a team is new in the league because it was promoted from the lower division.

Pre-match difference in rank: Most studies that use a measure of current performance or quality of a team, use the pre-match rank of both the home team and the away team. This is arguably a crude measure since the difference in pre-match rank between the number three and the number six of the table is treated the same as the difference between the numbers twelve and fifteen. Nevertheless, in line with previous studies, we use the difference in rank as an explanatory variable. If the rank in the league table is very different before a match starts, this may cause fewer attendees. We take the absolute difference in rank as explanatory variable.

Seasonal uncertainty: With respect to seasonal uncertainty, we use Jennett's (1984) measure of significance. We test for seasonal significance related to championship victory, qualification for the UEFA Champions League, qualification for the UEFA Europa League and relegation. Seasonal significance if based on the (mathematical) possibility to obtain a certain end-of-season achievement, i.e. to finish at a certain league rank. Whether or not a team is still able to obtain such a rank depends on the potential in-season number of points and the minimum number of points needed to achieve a certain end-of-season target. This minimum number of points is easily determined, since Jennett (1984) assumes that the final table is known in advance. This is a valid assumption as the necessary number of points to obtain a certain target is rather constant across seasons (see also Figure A1 in Appendix A). Note that these threshold numbers already include the (season specific) changes in the structure of the play-offs. Furthermore, it also incorporates the number of Dutch teams that could qualify for European competitions, which slightly decreased over time. Both items mainly impact the necessary number of points for the qualification for (the play-offs for) the UEFA Champions League. In general, the stability of the required number of

---

[5] We also test with two measures of match uncertainty that are frequently used within attendance demand studies, i.e. the *Theil-index*, based on bookmaker odds, and the points-per-game (*PPG*) measure, based on in-season performances, as proposed by Forrest, Simmons and Buraimo (2005). The parameter estimates including these alternative variables are presented in Appendix B, accompanied with a brief discussion on these variables.

points means that clubs and fans might use previous results to form a fairly precise expectation of the numbers of points that a club needs to finish at a certain league rank.

Local variables: dummy for *Derby*, capturing within-province rivalry; dummy for *Weekday* matches, capturing short-term opportunity costs; *Temperature* which partly captures a seasonal pattern; *Precipitation* which mainly captures short-term opportunity costs.

Month dummies: to capture potential seasonal influences on attendance not being seasonal uncertainty and weather conditions.

Some of our variables – stadium attendance of both the home team and away team in the previous season, team quality and cumulative surprise measured over the previous 34 matches – use values from the previous season. Since each season contains at least one, but potentially two or three promoted teams, we must decide how to deal with this. One solution would be to use values from previous season, when the club played on the second level of professional football. However, this approach seems suboptimal for at least the following two reasons. First, related to attendance numbers, we can be sure that these are lower if a club plays on the second level. However, the accuracy of the available data is doubtful, especially for the earlier seasons in our sample. Second, and related to both attendance numbers and (expected) results, the second level is substantially different from the highest level in Dutch football. For example, the majority of matches are played on Friday evenings, with rather limited media coverage. Also, the structure of the division is different and has certain special regulations with regard to the qualification for end-of-season play-off matches for promotion. This may provide clubs with unclear incentives (i.e. if a club already qualified for the play-offs in October, it might not feel the urgency to perform in the rest of the season. Therefore, as an easy and practical solution, we use the values of the relegated teams as if they were related to the promoted teams. The promoted teams are generally quite comparable to the relegated teams, for example in terms of budget, stadium and fan base. Thus, this approach seems applicable. Furthermore, we also include a dummy variable for promoted teams that will account for potential biases of this practical approach.

4. Parameter estimates

Our baseline parameter estimates are shown in Table 1. Column (1) shows the parameter estimates with match uncertainty measured by *Home win probability* while column (2) has match uncertainty

measured by *Match-Expectation*. In general, the parameter estimates of the common variables in the two columns are very much the same.

Table 1: Baseline parameter estimates

	(1)		(2)		(3)	
Stadium capacity (log)	0.636	(0.013) ***	0.636	(0.013) ***	0.635	(0.013) ***
Av. attendance previous season (log)	0.249	(0.014) ***	0.249	(0.014) ***	0.250	(0.014) ***
Opponent av. attend. prev. season (log)	0.063	(0.006) ***	0.062	(0.006) ***	0.062	(0.006) ***
Team quality (previous 34 matches)	0.623	(0.054) ***	0.627	(0.054) ***	0.639	(0.055) ***
Cumulative surprise (prev. 34 matches)	0.364	(0.046) ***	0.362	(0.046) ***	0.358	(0.047) ***
Home Win Probability	-0.661	(0.107) ***			0.694	(1.045)
Home Win Probability-squared	0.662	(0.114) ***			-1.152	(0.714)
Match Expectation			-0.278	(0.043) ***	-0.622	(0.340) *
Match Expectation - squared			0.082	(0.014) ***	0.229	(0.086) ***
Promoted	0.022	(0.010) **	0.022	(0.010) **	0.024	(0.010) **
Pre-match rank difference (absolute)	-0.003	(0.001) ***	-0.004	(0.001) ***	-0.003	(0.001) ***
Championship	0.242	(0.090) ***	0.246	(0.090) ***	0.254	(0.090) ***
UEFA Champions League	0.163	(0.053) ***	0.164	(0.053) ***	0.167	(0.053) ***
UEFA Europa League	0.040	(0.038)	0.040	(0.038)	0.038	(0.038)
Relegation	0.121	(0.039) ***	0.122	(0.039) ***	0.123	(0.039) ***
Derby	0.044	(0.010) ***	0.043	(0.010) ***	0.044	(0.010) ***
Weekday	-0.046	(0.011) ***	-0.046	(0.011) ***	-0.047	(0.011) ***
Temperature	0.014	(0.008) *	0.014	(0.008) *	0.014	(0.008) *
Precipitation	-0.003	(0.007)	-0.003	(0.007)	-0.004	(0.007)
July/August	-0.085	(0.018) ***	-0.085	(0.018) ***	-0.085	(0.018) ***
September	-0.056	(0.017) ***	-0.056	(0.017) ***	-0.056	(0.017) ***
October	-0.024	(0.015)	-0.024	(0.015)	-0.024	(0.015)
November	-0.013	(0.014)	-0.013	(0.014)	-0.013	(0.014)
December	-0.015	(0.014)	-0.015	(0.014)	-0.015	(0.014)
February	-0.006	(0.013)	-0.006	(0.013)	-0.005	(0.013)
March	0.021	(0.014)	0.021	(0.014)	0.020	(0.014)
April	0.046	(0.015) ***	0.046	(0.015) ***	0.046	(0.015) ***
May/June	0.001	(0.018)	0.001	(0.018)	0.001	(0.018)
-Loglikelihood	358.49		356.45		354.94	

Note: Tobit regression with log(attendance) as dependent variable and with the upper limit set at 95% of the stadium capacity. The estimates are based on 4,586 observations (1,896 censored). Standard errors in parentheses. ***$p<0.01$, **$p<0.05$, *$p<0.10$

Stadium capacity has a positive effect on stadium attendance, but there is not a one to one relationship. Ignoring capacity constraints, a stadium that has a 1% higher capacity on average attracts 0.64% more attendants. The number of attendants is persistent in the sense that the average number of attendants in the previous season has a positive effect on attendants in the current

season. Also, the average number of attendants of the opponent in the previous season has a positive effect, although the magnitude is much smaller. A 1% higher home match attendance in the previous season induces a 0.25% higher attendance in the current season. Both team quality and cumulative surprise, measured over the previous 34 matches, have positive effects on attendance.

In line with the recent stream of literature, our main interest is in the results for the variables measuring uncertainty. For match uncertainty, we use *Home win probability* and *Match-Expectation*. In general, the results contradict the classical UOH. For both measures, the relation with attendance is convex, meaning that loss aversion dominates the preference for uncertain outcomes. In column (1), attendance has a minimum value at a home win probability (HWP) of 0.50. In column (2), attendance has a minimum value if the expected number of points is 1.70. If we assume the probability of a draw to be 0.25 (approximately equal to the in-sample mean value), this corresponds to a HWP of 0.48. Since our sample has a mean value of 0.46 – with a minimum of 0.06 and a maximum of 0.88, this implies that for some matches the HWP lies within the downward sloping part of the convex curve, while for other matches in the upward sloping part of the convex curve. If the HWP lies within the downward sloping part, an increase in the home win probability results in a decrease in stadium attendance. This seems to contradict the concept of loss aversion, since that would predict fans to favor higher home win probabilities. In their model, Coates, Humphreys and Zhou (2014) provide an explanation, i.e. the interest of fans in upsets. Fans may want to attend live matches, because they enjoy the possibility of an unexpected win. These cases are characterized by relatively low home win probabilities. Otherwise a home win is not unexpected. The interest in upsets may account for a downward sloping part of the convex curve. Loss aversion indeed accounts for the upward sloping part of the curve, i.e. the part where attendance increases with an increase in the home win probability. Furthermore, Humphreys and Zhou (2015) show that one can identify the presence of a home win preference by the sum of the coefficients for the home win probability and its squared value. If this sum is positive, a home win preference exists. The sum of the coefficients in column (1) is positive but is very small and not significantly different from zero, suggesting that fans do not exhibit home win preferences (the F-statistic on the sum of the parameters being equal to zero equals 0.00). A similar conclusion follows

from the results in column (2) with *Match-Expectation*.[6] Thus, we find a convex relation that fits within an attendance demand model with reference-dependent preferences and loss aversion, but without a clear home win preference.

Note that the results for match uncertainty represent an *average* consumer, i.e. a hypothetical fan that does not prefer a specific football club but, given our model, has a general preference for home teams. Furthermore, this fan is loss averse and likes to attend unlikely home wins (upsets). Although the average fan does not prefer any club, she is faced with heterogeneity between teams, for example regarding home win probabilities. Not surprisingly, better teams have higher win probabilities. Within our sample, PSV has the highest average home win probability of 0.69 (n = 255, minimum = 0.33 and maximum = 0.88). In contrast FC Dordrecht has the lowest average home win probability of 0.26 (n = 17, minimum = 0.09 and maximum = 0.39). Given that the minimum of the convex curve for the average fan is about 0.48, it follows that the reason to attend a match is club-specific. In general, average fans like to attend PSV-matches because they are loss averse and want to experience a home win, while they would like to attend matches of FC Dordrecht to see an upset. Thus, when interpreting the results, one should note that it relates to an average consumer, who is faced with differences between clubs.

The dummy variable for teams that were promoted is positive, which suggests that teams that were promoted have more attendants in the first year after promotion than it has in later years. The difference in pre-match league ranking has a negative effect on stadium attendance.

Several of the seasonal uncertainty variables have a positive effect on stadium attendance. Competing for the championship, competing for a position in (the play-offs for) the UEFA Champion League and competing to avoid relegation have positive effects on stadium attendance, while competition for the play-offs in the UEFA Europa League has no significant effect. We return to the issue of the play-offs in relation to the stadium attendance in the next section.

The positive and significant coefficient for *Derby* suggests that matches against local rivals attract more attendance. Furthermore, people are less likely go to a stadium if the match is played on a *Weekday*. This suggests that opportunity costs play a role in people's decision to attend a match. The result for *Temperature* show that attendance is higher if the temperature is higher, while

---

[6] Assuming that the probability of a draw is 0.25 (approximately the sample average), the match expectation equals 3 times the home win probability + 0.25. Taking this into account, we test whether 3 times the parameter of the match expectation plus 10.5 times the parameter of the match expectation-squared is significantly different from zero. We find that this is not the case (F-statistic =0.33).

precipitation has no effect. The results for the weather variables are conditional on the inclusion of dummies for calendar months. Attendance is low in the early months of the season and high in April when the competition reaches the final stages.

The parameter estimates in columns (1) and (2) show that both HWP and Match Expectation perform well, i.e. both the linear term and the squared term are highly significant. To find out which of the two indicators for match uncertainty performs best, we include them both in one estimate. The results are presented in column (3). The Likelihood-Ratio test statistic comparing the estimates presented in columns (1) and (3) has a value of 7.10, which is significant for 2 degrees of freedom (the critical Chi-square value at a 5% significance level is 5.99). So, the contribution of the two match expectation variables in explaining the attendance is significant. The LR-test statistic comparing the estimates in columns (2) and (3) is 3.02 which is not different from zero at conventional levels of significance, i.e. ignoring the contribution of the two home win probability variables does not affect the overall estimation results significantly. Clearly, *Match-Expectation* outperforms *Home win probability*. From this, we conclude that the specification represented in column (2) is our preferred specification.

In addition to the parameter estimates reported in Table 1, we performed a range of sensitivity analysis. First, we investigated the effects of using attendance rate as an alternative dependent variable, finding that the main results are hardly affected. Second, we ignored censoring at the maximum stadium capacity using a simple linear model with log attendance as dependent variable. Again, our main findings are hardly affected. Third, we investigated the sensitivity of our main findings to the inclusion of fixed effects. Non-linear fixed effects models suffer from the incidental parameters problem such that maximum likelihood estimators are biased and inconsistent when the length of the panel is small. According to Greene (2004a, 2004b) slope estimates in the Tobit model are not much affected by the incidental parameter problem although standard errors are likely to be underestimated. Nevertheless, to explore the sensitivity we included team and calendar year fixed effects and included team-year fixed effects in the Tobit model. In neither case did our conclusions change. Fourth, we investigated whether the effect of the cumulative surprise was stronger for recent matches finding that this is not the case.

5 The impact of play-offs

Thus far, we obtained results for seasonal uncertainty that include the impact of the play-offs. We implicitly assumed that fans experience the qualification for the end-of-season play-offs in a similar way as a certain qualification for a European competition, i.e. qualification that is based on the final league table without end-of-season play-offs. In other words, we assumed that the interest shifts from the qualification for UEFA competitions, to an interest in the qualification for the end-of-season play-offs.[7] This seems reasonable, if clubs and fans experience these two types of qualification as comparable rewards for the performance during the regular league. Then, the play-offs become a sort of separate competition with its own reward, i.e. the qualification for European football. Given that many club officials and coaches nowadays state that their goal for the season is to qualify for the play-offs, the assumption, at least partly, seems to make sense. Seasonal uncertainty is modeled in line with this reasoning.[8]

We evaluate the impact of the play-offs in terms of additional regular league match attendance. Since we compare the period before and after the introduction, and given that club-specific factors might matter, we prefer to work with a balanced panel of nine clubs that were active in the *Eredivisie* throughout the sample period. The clubs in the balanced panel are Ajax, AZ, FC Groningen, FC Twente, FC Utrecht, Feyenoord, PSV, SC Heerenveen and Vitesse. Appendix A provides separate details on the data for this set of clubs.

The impact of the play-offs is empirically tested by the inclusion of an interaction term for the variables *UEFA Champions League* and *UEFA Europa League*. The *UEFA Champions League* variable interacts with a dummy that takes on the value of one for the three seasons 2005/06 – 2007/08. During these seasons, play-offs for an UEFA Champions League ticket were organized.

---

[7] It would be more precise to say, "the qualification for the end-of-season play-offs, with a possibility to qualify for a European competition", since there is always a possibility that a club earns the entry ticket to Europe after the qualification for the play-offs. However, we prefer to discuss it as the qualification for end-of-season play-off. First, because it is shorter and easier to distinguish from the other case, i.e. a direct or certain qualification for Europe. Second, and more important, because we want to stress that it has to do with the interest in the play-offs and not that this provides the possibility to qualify for an UEFA competition.

[8] The other extreme would be that fans only value the qualification for European football, and the end-of-season play-offs are not a separate reward at all. Then, seasonal uncertainty related to the introduction of the play-offs should be modelled in a different way. As discussed by Bojke (2007) and Koning (2007), two opposing effects will emerge by the introduction of end-of-season play-offs. First, the number of significant matches increases, since more final rankings provide the possibility to qualify for European football. Second, the significance per match decreases, since the number of available tickets for UEFA competitions remains unchanged. This extreme situation does not seem to correspond with the current situation in Dutch professional football.

The *UEFA Europa League* variable interacts with a dummy variable that takes on the value of one for all seasons after the introduction of the play-offs in 2005/06. Note that for both periods, the design of the play-offs did not change. Furthermore, for both types of play-offs and for all relevant years, four clubs compete against each other in two rounds in which a team plays the opponent twice, once at home and once away. The winner of round one will continue to round two, where the play-off champion is decided. In general, the introduction of the play-offs changed seasonal uncertainty, for example by increasing the number of matches with a non-zero significance and a redistribution of uncertain matches over a different set of teams. As far as the utility function of attendants with regard to seasonal uncertainty remained the same and given that the introduction of the play-offs was the main change with regard to seasonal uncertainty related to the qualification for European club competitions, we suggest that our approach is able to measure the impact of play-offs on attendance demand.

Table 2 presents the results of this approach. Column (1) reports the results for seasonal uncertainty of our baseline model presented in column (2) of Table 1, with data from the balanced panel. It follows that, for this set of teams, the effects of *Championship* and *UEFA Champions League* as well as *Relegation* are significantly different from zero. Note that the coefficient for *Relegation* is negative, which probably has to do with the selection of the subset of teams that contains clubs that only rarely compete for avoidance of relegation. The effect of *UEFA Europa League* is insignificant. We separate the effect of the play-offs in column (2) with the inclusion of the interaction terms *D*UEFA Champions League* and *D*UEFA Europa League*. For the UEFA Champions League and the UEFA Europa League, the interaction terms are significant. This suggests that these play-offs did matter for regular-league match-attendance. We consider a subset of eight seasons in column (3), i.e. four seasons prior the introduction of the play-offs and four seasons after the introduction. Now, the effect of the interaction term related to UEFA Champions League is insignificant, but the interaction term for UEFA Europa League remains statistically significant. In general, these results suggest that seasonal uncertainty regarding the qualification for UEFA competitions gained importance for attendants after the introduction of the play-offs. This might be the result of an increased suspense towards the end of the season. In that sense, it relates to the notion as proposed by Ely, Frankel and Kamenica (2015) and as discussed by Pawlowski, Nalbantis and Coates (2018).

Table 2: Impact of play-offs for European football

Seasons	2001/02-2015/16		2001/02-2015/16		2001/02-2008/09	
	(1)		(2)		(3)	
Championship	0.205	(0.075) ***	0.233	(0.075) ***	0.024	(0.081)
UEFA Champions League	0.134	(0.043) ***	0.022	(0.052)	0.170	(0.082) **
D*Champions League			0.280	(0.074) ***	0.031	(0.092)
UEFA Europa League	-0.000	(0.038)	-0.098	(0.054) *	-0.204	(0.055) ***
D*UEFA Europa League			0.149	(0.062) **	0.360	(0.078) ***
Relegation	-0.123	(0.059) **	-0.120	(0.059) **	-0.042	(0.076)
-Loglikelihood	342.20		330.98		159.54	
Observations	2294		2294		1223	
Censored observations	978		978		557	

Note: Tobit regression with log(attendance) as dependent variable and with the upper limit set at 95% of the stadium capacity. Estimates are based on a balanced panel of nine teams that were active in the *Eredivisie* throughout the sample period. The nine clubs are Ajax, AZ, FC Groningen, FC Twente, FC Utrecht, Feyenoord, PSV, SC Heerenveen and Vitesse. All specifications contain all other variables included in column (2) of Table 1 (not reported). Standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.10

Figure 2: Model predictions of mean attendance with and without play-offs
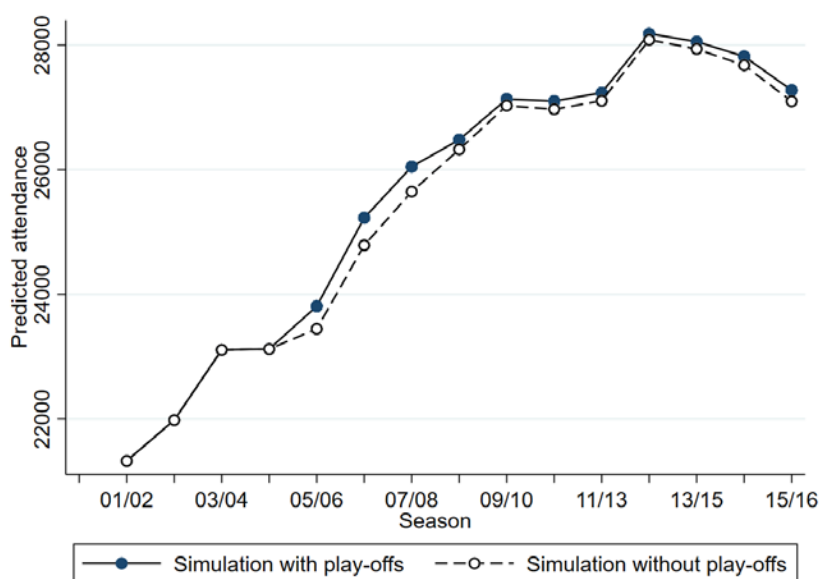

Note: Play-offs were introduced in the season 2005/06. Simulations are based on the results obtained from column (2) in Table 2.

We continue with an evaluation of the overall impact of the play-offs and use the results of column (2) of Table 2 to obtain model predictions. For each observation, we obtain a fitted value, taking

into account the upper limit of 95% of capacity. Then, we also calculate a value without taking the interaction terms into account, as if there would have never been play-offs. The results are shown in Figure 2, where we plot the mean predicted values by season for both situations with and without play-offs. Obviously, for the first four seasons there is no difference. From season 2005/06 onward, after the introduction of the play-offs, we observe very small differences. This suggests that the play-offs have made a rather stable contribution to the variation of within-season stadium attendance during regular league matches. It also suggests that between-season variation, and in particular the increase in the average attendance over time (as observed in both Figure 1 for the full sample and in Figure 2 for the balanced panel) cannot be attributed to the play-offs. From Figure 2 it is also clear that the effect of the play-offs on stadium attendance is bigger in the first three years when there were also play-offs for the UEFA Champions League (on average 393 attendants per match). In later years, the effect is based only on the play-offs for the UEFA Europa League (on average 126 attendants per match).

Figure 3: Scatterplot of attendance during the play-offs and regular league matches
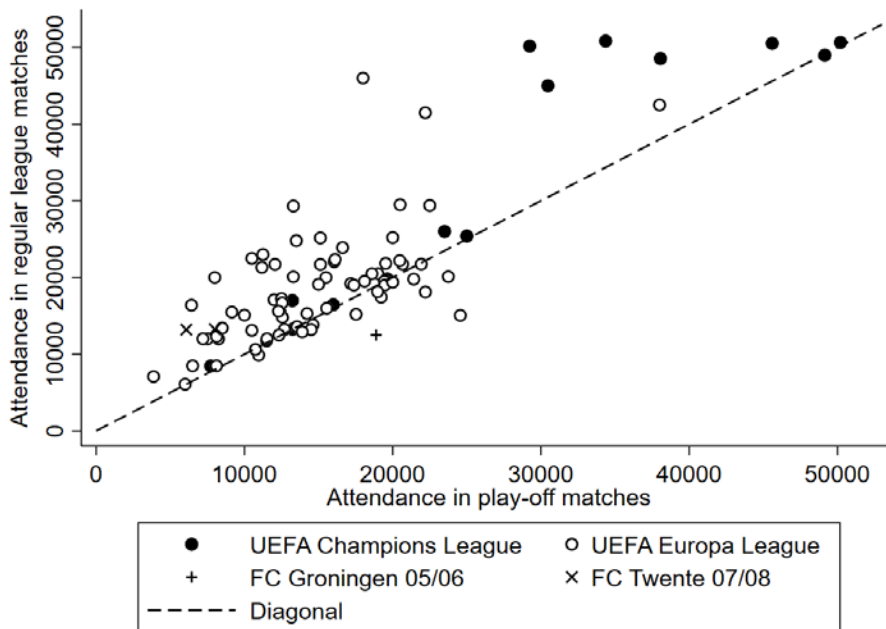
Of course, the end-of-season play-off matches themselves also attract attendance. This additional attendance should be considered in an evaluation of the effectiveness of the play-offs on stadium attendance. In Figure 3, we plot the attendance during the play-off matches against the attendance

during the same match in the regular league. The special cases for FC Groningen and FC Twente are indicated, because these play-off matches were played in a different stadium than the regular league match. Observations above the diagonal indicate that attendance was higher during the regular league match. Although there are many observations close to the diagonal, the majority lies within the upper left part of the figure. While match uncertainty, the quality of the competing teams as well as seasonal uncertainty should be relatively high for the play-off matches, fans seem to be less likely to attend them, compared to regular league matches. Thus, this we cannot explain with the behavioral economic principles applied above. Instead, a potential reason could be that season ticket holders, in general, have to pay extra for these play-offs.

Besides this modest interest in the play-off matches by fans, there are some other drawbacks of the system. The play-off matches are generally seen as an unwanted extension of the season by clubs, managers and players. This holds especially in the seasons when there is a UEFA European Championship or a FIFA World Cup coming up, for which preparations start immediately after the season. Furthermore, it can be argued that it is not always the strongest team that wins the play-offs and, thus earns a ticket for European football in next season. Koning (2007) found that this probably is the case. For Dutch football in general, it would be better if the best teams represent the Netherlands in European competitions. They have the best perspective to survive the group stage and maybe a few rounds in these competitions, earning points for the UEFA rankings for club competitions. Furthermore, such international matches are valuable for the development of young and talented players that might be selected for national teams. In general, these players play for the better teams.

6   Discussion and conclusion

We investigate how outcome uncertainty, loss aversion and team quality affect the variation in stadium attendance for sixteen seasons of the highest tier of Dutch professional football. Using different measures of match uncertainty, we find evidence that contradicts the classical UOH. Instead, our results suggest that fans have reference-dependent preferences with loss aversion. Although we cannot rule out any preference for uncertain outcomes, the results suggest that loss aversion dominates. However, no home win preference is found. Furthermore, team characteristics are important determinants of stadium attendance. In particular, the quality of the team as measured by bookmaker odds and the cumulative surprise measured as the difference between actual points

and expected points have positive effects on stadium attendance. Furthermore, our analysis reveals that the variable we introduce to reflect match-expectation outperforms the home win probability as indicator of match uncertainty. This stresses the importance to account for the probability of a draw in attendance demand models within football.

As to seasonal uncertainty, we find significant results for uncertainty related to the championship victory, the qualification for the UEFA Champions League and seasonal significance related to relegation. This is in line with the UOH. Furthermore, we find that the introduction of play-offs for European matches in 2005/06 has had a positive effect on match attendance. However, the magnitude of the effect on regular league matches is small. Also, fans seem to be less interested to attend the play-off matches than regular league matches against the same opponent, despite high uncertainty regarding the winner.

In general, stadium attendance in the Dutch *Eredivisie* is high. This suggests that there is no urgent need to adjust the current policy of clubs and the sports body. Furthermore, given the small marginal effect of the introduction of the play-offs, one might question the effectiveness of any policy measure related to the aspects investigated in this study. There seems to be a vast majority of fans that simply decides to attend any match, irrespective of the conditions, competing teams and the season. As long as this majority keeps returning to the stadium, there does not seem to be any problem. However, in recent years attendance has dropped somewhat. If this continues in the near future, clubs may feel the need for a policy change. Then, it is useful to understand why people decide (not to) attend the stadium, so that one can formulate clear policy advice. Our results may serve as a basis.

References

Bartling, B., Brandes, L., and Schunk, D. (2015) Expectations as Reference Points: Field Evidence from Professional Soccer. *Management Science*, 61, 2646-2661.

Bojke, C. (2007) The impact of post-season play-off systems on the attendance at regular season games, in: J. Albert and R. Koning (eds.), *Statistical Thinking in Sports*, Chapman and Hall/CRC, 179-202.

Borland, J., and Macdonald, R. (2003) Demand for Sport. *Oxford Review of Economic Policy*, 19, 478–502.

Budzinski, O., and Pawlowski, T. (2017) The behavioral economics of competitive balance: theories, findings, and implications. *International Journal of Sport Finance*, 12(2), 109-122.

Buraimo. B. and Simmons. R. (2015) Uncertainty of outcome or star quality? Television audience demand for English Premier League football. *International Journal of the Economics of Business*, 22, 449-469.

Cairns, J. A., Jennett, N., and Sloane, P. J. (1986) The economics of professional team sports: A survey of theory and evidence. *Journal of Economic Studies*, 13, 3–80.

Coates, D., and Humphreys, B. R. (2012) Game attendance and outcome uncertainty in the National Hockey League. *Journal of Sports Economics*, 13(4), 364-377.

Coates, D., Humphreys, B. R., and Zhou, L. (2014) Reference-dependent preferences, loss aversion, and live game attendance. *Economic Inquiry*, 52(3), 959–973.

Cox, A. (2015) Spectator demand, uncertainty of results, and public interest: evidence from the English Premier League. *Journal of Sports Economics*, 19(1), 3-30.

Czarnitzki, D., and Stadtmann, G. (2002) Uncertainty of outcome versus reputation: Empirical evidence for the first German football division. *Empirical Economics*, 27, 101–112.

Ely, J., Frankel, A., and Kamenica, E. (2015) Suspense and surprise. *Journal of Political Economy*, 123(1), 215-260.

Forrest, D., Simmons, R., and Buraimo, B. (2005) Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy*, 52, 641–61.

Greene, W. (2004a) The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects, *Econometrics Journal*, 7, 98-119.

Greene, W. (2004b) Fixed effects and bias due to the incidental parameters problem in the Tobit model, *Econometric Reviews*, 23 (2), 125-147.

Humphreys, B. R., and Zhou, L. (2015) The Louis-Schmelling paradox and the league standing effect reconsidered. *Journal of Sports Economics*, 16(8), 835-852.

Jennett, N. (1984) Attendances, uncertainty of outcome and policy in Scottish League Football. *Scottish Journal of Political Economy*, 31, 176–198.

Kahneman, D., and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263-291.

Koning, R. H. (2007) Post-season play and league design in Dutch soccer, in: P. Rodriquez, S. Kesenne and J. Garcia (eds.), *Governance and Competition in Professional Sports Leagues*, Oviedo: Ediciones de la Universidad de Oviedo, 191-215.

Koszegi, B., and Rabin, M. (2006) A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4), 1133-1165.

Martins, A. M., Cró, S. (2016) The demand for football in Portugal: new insights on outcome uncertainty. *Journal of Sports Economics*, 19(4), 473-497.

Neale, W. C. (1964) The peculiar economics of professional sports: A contribution to the theory of the firm in sporting competition and in market competition. *The Quarterly Journal of Economics*, 78, 1–14.

Pawlowski, T., and Anders, C. (2012) Stadium attendance in German professional football - The (un)importance of uncertainty of outcome reconsidered. *Applied Economics Letters*, 19, 1553–1556.

Pawlowski, T., Nalbantis, G., and Coates, D. (2018) Perceived game uncertainty, suspense and the demand for sport. *Economic Inquiry*, 56(1), 173-192.

Pawlowski, T. (2013) Testing the uncertainty of outcome hypothesis in European professional football: A stated preference approach. *Journal of Sports Economics*, 14(4), 341–367.

Rottenberg, S. (1956) The baseball players' labor market. *Journal of Political Economy*, 64, 242-258.

Schreyer, D., Schmidt, S. L., and Torgler, B. (2016) Against all odds? Exploring the role of game outcome uncertainty in season ticket holders' stadium attendance demand. *Journal of Economic Psychology*, 56, 192-217.

Van Ours, J.C. and M.A. van Tuijl (2016) In-season head-coach dismissals and the performance of professional football teams, *Economic Inquiry*, 54 (1), 591-604.

Appendix A: Further details on our data

Our data on match date, competing teams, match results, stadium capacities and attendance figures are collected from various internet sources, sports magazines and newspaper archives. The attendance figures are typically a reported number, from which it is not possible to separate season-ticket holders and pay-at-the-gate attendance. Furthermore, some reported data are rounded to e.g. 100. The betting odds are from bookmaker agencies William Hill (97 percent), Ladbrokes (2 percent) and others (1 percent). Data on the general weather conditions on match days refer to De Bilt, a municipality in the center of the Netherlands where the KNMI, the Royal Dutch Meteorological Institute is located. From the data collected, we create the variables presented and discussed in section 3. For many variables the definition is straightforward, but some variables require more explanation or they are used only in appendix B:

Theil-index: $\tilde{T} = \sum \pi \ln\left(\frac{1}{\pi}\right)$ with π being either the home win probability, the probability for a draw or the away win probability. The index is increasing when probabilities become more equal, with a value close to zero in case of a high probability for one of the outcomes and $\ln(3)$ in case of equal probabilities.

PPG: Points-per-game measure, calculated as: $PPG = |HA + P_h - P_a|$ in which, h denotes the home team, a indicates the away team and s refers to the season. The points-per-game match uncertainty measure PPG is the absolute value of home advantage (HA) plus the number of points per game of the home team (P_h) minus the number of point per game of the away team (P_a). Home advantage is measured as the difference between the points per game won by all home teams and all away teams in previous season. The point per game values of the home and away team are measured as the number of points per game obtained in the current season. With this measure, match uncertainty increases with decreasing values.

Seasonal significance: Either "Championship Significance", "UEFA Champions League Significance", "UEFA Europa League Significance" or "Relegation Significance". Significance is defined as: $\sigma = \frac{1}{N - n}$ if $P_i + R_i > T$; $\sigma = 0$ if $P_i < T$; $\sigma = 0$ if $T \leq P_i + R_i$ in which i denotes the match, c indicates the club and s refers to the season. Then, significance σ is given by the reciprocal of the total number of matches in the season (N) minus the number of matches already played (n) prior to match i. This is only the case if the potential number of points for a specific team ($P_i + R_i$) is larger or equal to the total number of points needed to obtain the predefined total number of points. This total is constant throughout the season and represented by T which should be larger than the current number of points ($P_i < T$) prior to a match. In case the potential number of points is smaller than the total number of points, significance becomes zero.

Weekday: Dummy with value one if a match was played on a Monday, Tuesday, Wednesday or Thursday. Value is zero in case of a national holiday. National holidays include Easter, Queen's Birth Day / King's Birth Day, Ascension Day, Pentecost and Christmas. In the Netherlands, there is a national holiday on Sunday and Monday during Easter and on Sunday and Monday during

Pentecost. Furthermore, there is a national holiday on the 26th of December, the second day of Christmas.

Temperature: Daily mean temperature measured in 0.1 degrees Celsius; divided by 10 to make interpretation easier.

Precipitation: Daily precipitation, amount in 0.1 mm; divided by 10 to make interpretation easier.

In table A1, descriptive statistics are presented. Since we need previous season's bookmaker odds for the construction of the team quality and cumulative surprise variables, and we lack odds prior to the season 2000/01, this season is left out of the analysis. Thus, we are left with 15 seasons with 18 clubs and 17 home matches per club, resulting in 4,590 observations. For 186 observations, we found a reported number of attendance that was (slightly) higher than our documented stadium capacity. For these cases, we assumed attendance to be the same as capacity. During two matches, no fans were allowed to attend as a punishment by the Dutch football association. These two matches are left out of the analysis. Furthermore, bookmaker odds are missing for two matches, which results in 4,586 observations for the match uncertainty measures that are based on these bookmaker odds. A total of 1,896 observations (41 percent) is censored at the upper limit of 95 percent of capacity. The right hand panel of Table A1 contains descriptive statistics for a subsample of nine teams that were active in the *Eredivisie* throughout the sample period. This balanced panel contains 2,294 observations, with 978 observations (43 percent) censored at 95 percent of capacity. We find a rather high standard deviation and difference between the minimum and the maximum value for attendance, which can easily be explained by the fact that all clubs are grouped together for these values. After controlling for capacity in the attendance rate variable, the differences are less severe.

The team quality variable cannot become negative, by definition, and the mean value is close to 47 for the full sample of clubs. This is approximately the number of points that teams typically obtain if they finish in the middle of the league, thus representing average quality. We also find that the mean values for the Theil-index and PPG are rather high (low) in comparison to their minimum and maximum values, which is less the case for the Home win probability and Match-expectation. Finally, the mean values for seasonal uncertainty are rather low, which can be explained by the fact that seasonal uncertainty develops exponentially. For the first match the value is equal to 0.03 in all cases. For a given match later in the season, the value is equal to 0.5 when only two matches are remaining, while it only reaches unity in case only one match is left to win. Furthermore, seasonal uncertainty is zero in case a team is not in contention anymore for a certain end-of-season achievement. Either, because at some point in the season it has already obtained a sufficient number of points, or it is not able to obtain a sufficient number of points anymore.

Table A1: Descriptive statistics for all teams and for the balanced panel of 9 clubs

	All teams					Balanced panel of 9 teams				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Attendance	4,588	18,191	12,086	2,258	53,052	2,295	26,128	12,307	6,057	53,052
Attendance Rate	4,588	0.88	0.12	0.30	1	2,295	0.89	0.11	0.42	1
Log Attendance	4,588	9.61	0.65	7.72	10.88	2,295	10.06	0.48	8.71	10.88
Log Capacity	4,588	9.74	0.64	8.17	10.89	2,295	10.18	0.48	9.10	10.89
Log attendance previous season	4,588	9.59	0.64	8.03	10.83	2,295	10.04	0.49	8.87	10.84
Log opponent att. prev. season	4,588	9.59	0.64	8.03	10.83	2,295	9.57	0.63	8.03	10.84
Quality Team previous 34 matches	4,588	46.96	11.76	23.52	76.95	2,295	55.25	10.40	35.71	76.95
Cumulative surprise (prev 34 m.)	4,588	0.13	7.20	-20.54	23,47	2,295	2.29	7.19	-20.54	23.47
Theil	4,586	0.98	0.13	0.44	1.1	2,294	0.93	0.15	0.44	1.1
PPG	4,588	0.83	0.63	0	3.94	2,295	1.01	0.68	0	3.94
Home Win Probability	4,586	0.46	0.18	0.06	0.88	2,294	0.55	0.17	0.11	0.88
Home Win Probability-squared	4,586	0.24	0.17	0	0.78	2,294	0.33	0.18	0.01	0.78
Match-Expectation	4,586	1.63	0.50	0.31	2.73	2,294	1.87	0.46	0.53	2.73
Match-Expectation-squared	4,586	2.91	1.63	0.10	7.44	2,294	3.73	1.65	0.28	7.44
Promoted	4,588	0.11	0.31	0	1	2,295	0	0	0	0
Rank Difference (absolute)	4,588	6.01	4.10	0	17	2,295	6.05	4.13	0	17
Championship	4,588	0.03	0.05	0	1	2,295	0.04	0.07	0	1
Champions League	4,588	0.04	0.07	0	1	2,295	0.05	0.09	0	1
UEFA Europa League	4,588	0.06	0.08	0	1	2,295	0.06	0.09	0	1
Relegation	4,588	0.05	0.09	0	1	2,295	0.03	0.05	0	1
Derby	4,588	0.09	0.29	0	1	2,295	0.06	0.24	0	1
Weekday	4,588	0.07	0.25	0	1	2,295	0.06	0.24	0	1
Temperature	4,588	0.89	0.56	-1.21	2.64	2,295	0.89	0.56	-1.21	2.64
Precipitation	4,588	0.23	0.42	0	4.00	2,295	0.23	0.43	0	4.00
July/August	4,588	0.09	0.29	0	1	2,295	0.09	0.29	0	1
September	4,588	0.10	0.30	0	1	2,295	0.10	0.30	0	1
October	4,588	0.11	0.31	0	1	2,295	0.11	0.31	0	1
November	4,588	0.11	0.31	0	1	2,295	0.11	0.31	0	1
December	4,588	0.11	0.31	0	1	2,295	0.11	0.31	0	1
January	4,588	0.07	0.26	0	1	2,295	0.07	0.25	0	1
February	4,588	0.13	0.34	0	1	2,295	0.13	0.34	0	1
March	4,588	0.11	0.32	0	1	2,295	0.12	0.32	0	1
April	4,588	0.12	0.32	0	1	2,295	0.12	0.33	0	1
May/June	4,588	0.05	0.22	0	1	2,295	0.05	0.22	0	1

Note: the balanced panel consist of nine clubs that are active in the *Eredivisie* in all seasons considered. These clubs are Ajax, AZ, FC Groningen, FC Twente, FC Utrecht, Feyenoord, PSV, SC Heerenveen and Vitesse. Since bookmaker odds are missing for two matches (one for the balanced panel), *Theil*, *Home win probability* and *Match-Expectation* contain two (one) less observation than the other variables.

In Figure A1 we plot the relevant number of points necessary to obtain an end-of-season achievement. The vertical reference-line marks the introduction of the end-of-season play-off scheme. The line for relegation remains flat for the entire sample period. For the UEFA Europa League, a stable number of points is necessary to qualify until the introduction of the play-offs. After the introduction, again a stable number of points is necessary for qualification, now for the

end-of-season play-offs. The line representing Championship exhibits a higher degree of variability, although it remains somewhere around 80 points.

Figure A1: Points needed to obtain some end-of-season achievement
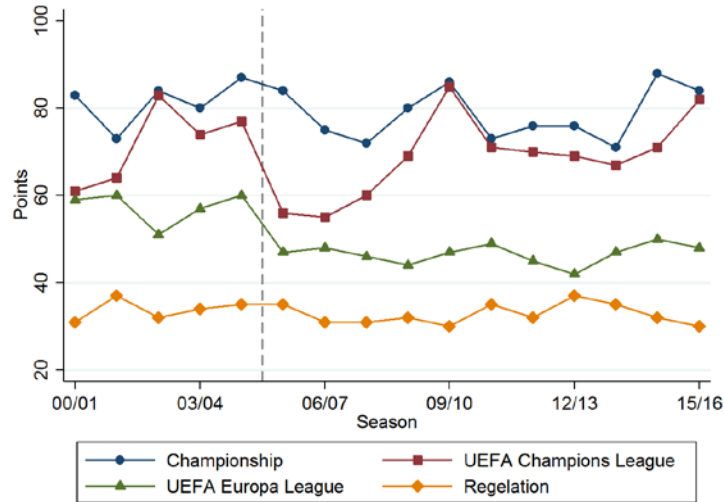
No stable pattern is found for the UEFA Champions League. This can be explained by changes in the number of direct tickets for this tournament, as well as the introduction of the play-offs during the sample period. Only one extra team, besides the champion, was allowed to qualify from the season 2001/02 onwards, as a result of the rather poor performances of the Dutch teams in Europe in the preceding seasons. However, the introduction of the play-off scheme in the 2005/06 season made it possible for multiple teams to obtain this qualification in an end-of-season competition. Teams ranked second up to fifth were competing for this second entry ticket during the 2005/06 – 2007/08 seasons. The play-offs were abandoned after the 2007/08 season, after which the runner up in the league directly qualified for the Champions League.[9] Again, we observe a rather flat line from that moment onwards, with two exceptions, in which both the champion as well as the runner-up obtained more than 80 points, which is rather high. In general, taking the changes in rules and regulations regarding qualifications into account, we find that the necessary number of points is rather stable over seasons. Especially for relegation and the UEFA Europa League (before and

---

[9] Only once out of three times, the runner-up won the play-offs for the lucrative Champions League (Ajax in the 2006/07 season). In the other two seasons, the club ranked fourth in the regular league won the ticket. Note that, when the play-offs for the UEFA Champions League were canceled, the runner-up in the league qualified for a start-of-season preliminary round of the UEFA Champions League, but for convenience, we do not make such a distinction here.

28

after the introduction of play-offs), which suggests that the assumption of a known final table is comparable to expectations being based on previous years' results.

Appendix B: Alternative baseline results

Table B1: Baseline results with Theil-index and Points per game

	(1)		(2)	
Stadium capacity (log)	0.634	(0.013) ***	0.633	(0.013) ***
Av. attendance previous season (log)	0.251	(0.014) ***	0.253	(0.014) ***
Opponent av. attend. prev. season (log)	0.081	(0.005) ***	0.075	(0.005) ***
Team quality (previous 34 matches)	0.201	(0.072) ***	0.281	(0.070) ***
Cumulative surprise (previous 34 matches)	0.529	(0.049) ***	0.635	(0.046) ***
Promoted	0.020	(0.010) *	0.028	(0.010) ***
Pre-match difference in rank (absolute)	-0.003	(0.001) ***	-0.000	(0.001)
Theil	-0.173	(0.036) ***		
PPG			-0.002	(0.007)
Championship	0.246	(0.090) ***	0.263	(0.092) ***
UEFA Champions League	0.163	(0.053) ***	0.166	(0.054) ***
UEFA Europa League	0.036	(0.038)	0.035	(0.038)
Relegation	0.118	(0.039) ***	0.118	(0.039) ***
Derby	0.046	(0.010) ***	0.048	(0.010) ***
Weekday	-0.045	(0.011) ***	-0.046	(0.011) ***
Temperature	0.015	(0.008) *	0.015	(0.008) *
Precipitation	-0.004	(0.007)	-0.003	(0.007)
July/August	-0.085	(0.018) ***	-0.081	(0.019) ***
September	-0.057	(0.017) ***	-0.057	(0.017) ***
October	-0.025	(0.015)	-0.025	(0.015)
November	-0.013	(0.014)	-0.013	(0.014)
December	-0.015	(0.014)	-0.014	(0.014)
February	-0.006	(0.013)	-0.006	(0.013)
March	0.020	(0.014)	0.020	(0.014)
April	0.046	(0.015) ***	0.048	(0.015) ***
May/June	0.001	(0.018)	0.002	(0.019)
-Loglikelihood	377.27		366.11	
Observations	4,586		4,588	
Censored observations	1,896		1,897	

Note: Tobit regression with log(attendance) as dependent variable and with the upper limit set at 95% of the stadium capacity. The estimates are based on 4,586 observations (1,896 censored). Standard errors in parentheses. ****p<0.01, **p<0.05, *p<0.10*

Both the *Theil-index* and the *PPG* measure are frequently used as measures of match uncertainty in attendance demand models to test the UOH. The *Theil-index* measures increased uncertainty with increasing values. The negative significant coefficient in column (1) provides evidence against the hypothesis. The *PPG* variable measures increased uncertainty with decreasing values. Although the coefficient is negative, it is insignificant in column (2). In general, both the *Theil-index* and *PPG* are not able to explain behavioral economic decision making under reference-dependent preferences and loss aversion. Therefore, using *Home win probability* and *Match-Expectation* seems more insightful.