Reference-dependent preferences in the case of Allsvenskan

Outcome uncertainty and live-game attendance

Gustav Johansson

Gustav Hallén

Fall semester 2017

January 17th, 2018

Bachelor Thesis, 15 ECTS

Department of Economics, Lund University

Supervisor: Erik Wengström

Abstract

This thesis studies how different preferences for uncertainty affects live game attendance in the Swedish premier football league Allsvenskan. A framework proposed by previous literature concerning this relationship is highlighted. In the presence of reference dependency, two types of preferences is of interest. On one hand, there is the Uncertainty of Outcome Hypothesis (UOH), stating that uncertainty drives attendance. On the other, there are loss averse preferences, predicting higher attendance on games where the outcome is certain. To examine what preferences holds for Allsvenskan, a panel data regression is used to investigate the relationship between the level of uncertainty and attendance. The probability of a home team win, derived from the odds, is used as a proxy for the uncertainty present in the game. Our panel includes data on twelve teams that have participated in Allsvenskan during the seasons 2014-2016. The dataset is created by combining game-by-game statistics from the official webpage of Allsvenskan, with odds data compiled from 20+ bookmakers. Additionally, other sources are used to obtain data on control variables which rounds out the set. Ultimately we find the UOH to be present in the aggregate spectator of Allsvenskan. Also, Attendance is found to peak at a point where the probability to win is 61,5%.

Keywords: Reference-dependent preferences, Uncertainty of Outcome Hypothesis, Loss Aversion, Football, Allsvenskan.

Acknowledgement

We wish to sincerely thank our supervisor Erik Wengström for guidance and interest in helping us. The assistance offered has been an invaluable resource in the creation of this thesis.

Table of Contents

1. Introduction

Uncertainty is a fact of life. Some enjoy it, while others don't. Those who do, get a thrill out of observing an activity where the outcome is not known, while others crave the security of certainty. Individuals are faced with this struggle every day. In essence, there is some level of uncertainty involved in every economic transaction that's conducted. This uncertainty is present in both small decisions concerning everyday purchases, such as a cup of coffee, and in decisions of major importance, such as where to invest your money. How we deal with these decisions comes down to personal preferences and characteristics. It is therefore of importance to understand the preferences that drive our decisions in the presence of uncertainty. For this reason, the topic of uncertainty and how it affects us has been a hot topic in academic circles for years. The interest in decisions under uncertainty has seen application in a multitude of fields, among them sports. The decision to attend a sport event comes with uncertainty as the outcome of the event is unknown. The application of behavioural economics to sports is beneficial as there are comprehensive data on both the level of uncertainty and on different outcomes. Sport is also relatable to a vast majority of people which can simplify a discussion that at other times may be abstract or difficult to understand. How people make decisions under uncertainty has long been examined but its relation to sports had only recently been given a comprehensive theoretical model to explain the process.

Rottenberg (1956) posited the idea that some sports fans prefer uncertain outcomes offered by two evenly matched teams, rather than a landslide victory of a major team or an underdog beating the odds. This has since become known as the Uncertainty of Outcome hypothesis (UOH). In essence, the hypothesis brings to light the importance of competition between teams, as a team that is far above the level of the others will burn itself out with its unexciting dominance. Up until recently the UOH has been without any real theoretical basis. However, with Coates et al. (2014) and their study on Major League Baseball[1], the idea behind the hypothesis has been extended, creating a theoretical basis. The framework includes an extension that highlights loss averse preferences and its effect on attendance, which the UOH previously hasn't able to explain. As a fan that is loss averse will experience more utility

[1] Commonly known as the MLB, which will be the notation used in the following text.

from a match where the outcome is certain, loss averse preferences predicts higher attendance on games with less uncertainty of outcome. These theories stand in contrast to each other. Determining which one of them is the most realistic is of importance for stakeholders within leagues in order to maximize attendance, profits and fan interest. Attendance figures in Europe is increasing but at the same time the outcome seem to be more certain as fewer teams compete for the league trophies. This may point to a shift within the literature of sport economics and that the predictions of the UOH no longer can be taken for granted as the absolute truth.

Expanding on the research made by Coates et al. (2014), we wish to apply the model on the Swedish football league Allsvenskan.[2] This as a way of broadening the observed effects of different attitudes towards uncertainty, by applying it to a unexamined league and country. Our reasoning for this stems from the observed difference in preferences between different sports, leagues and cultures. Ultimately, we wish to answer the question *"What preferences concerning uncertainty affect the live-game attendance in Allsvenskan?"* Answering this question may offer important insights on fan preferences and the demand in Allsvenskan, which in turn can be used by team managers and league organizers.

This is specifically done by examine the three seasons of Allsvenskan spanning from 2014 to 2016 using a panel data regression. Data on attendance was collected and combined with odds data which has been recalculated into the probability of the home team winning. This probability is used as a measurement of uncertainty, enabling the analysis on how attendance is affected by uncertainty. Additionally, we have collected data on other variables that has had a proven significant effect on live-game attendance, such as characteristics of the home and the away team.

[2] For the purposes of this thesis, "football" will be used instead of the word "soccer" as soccer is mostly used from an American point of view.

2. Theory

This section describes the theory behind the study of spectator preferences and their effects on attendance. Firstly, the background of the theoretical framework is introduced. Secondly, the basics of the model is explained, and the idea of reference-dependent preferences is explored. The final subsection includes different types of spectator preferences most relevant for this study, including some special cases.

2.1 Background of the theoretical framework

This thesis is based on the framework developed by Coates et al. (2014). Prior to its publication, studies into uncertainty and sports were still conducted, but they were severely lacking in some ways. First, the UOH lacked a straightforward theoretical basis. Secondly, theories couldn't offer explanations for what the UOH would consider deviant behaviour, as games where the outcome was almost certain still had an appeal to spectators. As will be made clear later in our thesis, Coates et al. (2014) found a way to explain preferences that weren't in line with the UOH. It is because of this explanatory power that it's used as a base in our study.

Coates et al. (2014) describe that their model is based on the model by Card and Dahl (2009), which is in turn was based on the gain-loss framework originally developed by Kőszegi and Rabin (2006). The framework developed by Card and Dahl (2009) looked at how anticipation of a certain outcome in an NFL[3] game may lead to familial violence if this anticipation is not met. Coates et al. (2014) removed the violent aspects of the model while retaining the focus on expectations of the outcome. Particularly, the study examined how expectations of an outcome may affect the live-game attendance. We intend to follow this approach, although shifting the attention from the MLB to that of Allsvenskan.

The model states that consumers experience two kinds of utilities upon attending a game: Consumption Utility and Gain-Loss Utility. The former corresponds to the intrinsic utility a consumer obtains by attending the event. One form or another of consumption utility will as such be obtained from the event, regardless of outcome (Kőszegi & Rabin, 2006).

[3] Stands for National Football League, the American Football league present in the United States of America.

Gain-Loss Utility occurs when there is uncertainty in the outcome of an event, and an individual stand to gain utility from a particular outcome depending on its relation to an initial expectation (Kőszegi & Rabin, 2006). Coates et al. (2014) assumes that individuals form their reference point accordingly to the objective probability that the home team will win. In accordance to Coates et al. (2014) we derive this probability from the odds of the home team winning, expressed as the percentage likelihood of this outcome, later referred to as the reference point, p.

There is however an important difference of note between Allsvenskan and the MLB. The possibility of a draw in football adds an outcome not present in the MLB, and complicates our application of the model. To remedy this, we assume that a draw for the home team will be viewed as a loss. This is a bold and not entirely realistic assumption but a necessary one for the model to make sense. Our reasoning is that the home team has an advantage by playing at their home field. A draw therefore means being deprived of the full award of winning, hence it being viewed as a form of defeat. Following this line of thinking, the away team are at a disadvantage when playing away and a draw from their point of view can be viewed as a win. It could be argued that for a strong team, a draw against a weaker opponent would be viewed negatively while the opposite may be true for the weaker team. To model this, a multitude of assumptions must be made, which would complicate the model. The potential gains of this complication would not be sufficient to cover the costs in terms of time and additional complexity. We feel that the use of the original model would be jeopardized and the added complexity would be beyond the scope of this thesis. Furthermore, when calculating the probability of a home team win based on the odds, we've included the likelihood of a draw to obtain a probability between 0-1. Therefore, a draw outcome has been taken into consideration in this case (see equation 6).

2.2 Basics of the Model

In the model by Coates et al. (2014), Consumption Utility is divided into U_w and U_l, the first representing a win and the later a loss. Gain-Loss Utility is based on the initial expectation in the form of the individual's reference point, p, in relation to the actual outcome. The marginal impact of a positive deviation from the reference point is referred to as α. When $\alpha > 0$, the utility of a home team win (y=1) becomes

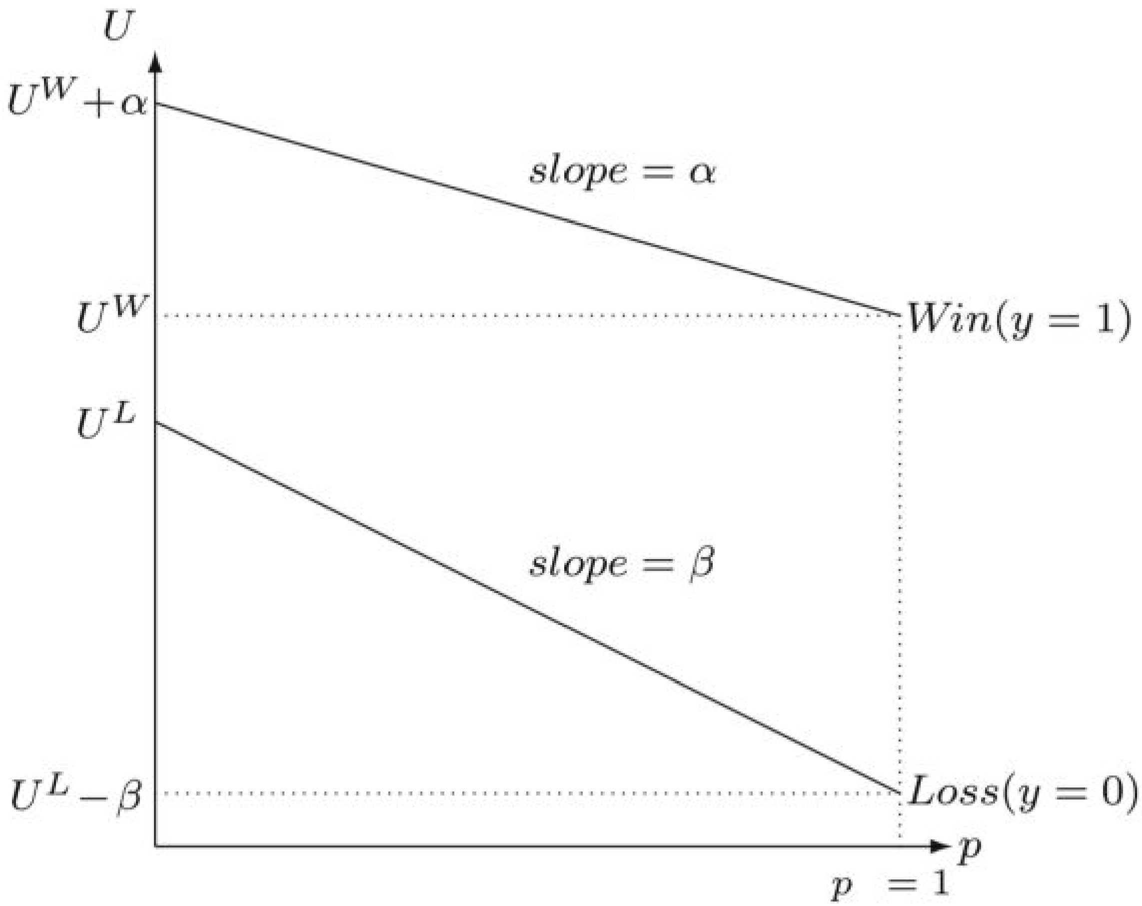
$U_w + \alpha(y - p) = U_w + \alpha(1 - p)$.

The marginal impact of a negative deviation from the reference point is referred to as β. When $\beta > 0$, the utility of a home loss (y=0) becomes

$U_l + \beta(y - p) = U_l + \beta(0 - p)$

Under the assumption that $U_w > U_l$, implying that the utility gained from a win is greater than the utility of a loss one can graphically present the relationship between the reference point, game outcome and the total utility gained from attending a game (figure 1). Total utility is displayed on the y-axis while the reference point is presented on the x-axis. Since the reference point is denoted as a probability it spans from 0-1. Where 0 implies that fans have no expectations of a home team win and 1 implies a fully expected win. Maximum utility of attending a game is shown as $U_w + \alpha$, interpreted as a home team win when there was an expected zero-probability to win. On the opposite side of the spectrum, minimum utility attending a game is given by $U_l - \beta$. In this case, the fans fully expected the home team to win but the outcome was a loss. Having low expectations and expecting a loss garners more utility, as shown by the figure. Starting out in the intercepts on the y-axis, an individual experiences a decline in utility as the expectations of the outcome increase. The marginal impacts of α and β determines the final outcome by capturing the reference-dependent preferences.

Figure 1. Reference point, outcome of game and total utility (Coates et al, 2014)[4]



As the outcome of the game is uncertain, an individual can't know how much utility the game will bring. Under the assumption made by Coates et al. (2014), the probability of a win is the objective probability of the outcome implied by the odds. An individual can derive the expected utility from attending a game by multiplying the probability of a win with the utility of a win and then add the probability of a loss times the utility of a loss:

(1) $E[U] = p[U_w + \alpha(1-p)] + (1-p)[U_l + \beta(0-p)]$

Here the utility of consumption is clearly observable as the expected utility of a win and the expected utility of a loss. To make the nature of gain-loss utility even clearer, equation 1 is rearranged as done by Coates et al. (2014), resulting in the following quadratic function:

(2) $E[U] = (\beta - \alpha)p^2 + [(U_w - U_l) - (\beta - \alpha)]p + U_l$

[4] Somewhat modified. P is used instead of Pr on the x-axis.

Equation two builds on the prior equation by showing that the expected utility of attending a game is a quadratic function of the home team probability to win. This equation is a key to understand the different preferences later discussed and further a vital part of our econometric application. As will be made clear, the marginal impact of a deviation from the reference point α, β as well as the consumption utilities U_w and U_l is of importance when deriving types of preferences.

There is an alternative utility which must be considered by those contemplating whether to attend a match, called the reservation utility (ν). It is generated by choosing not to attend and is the counterpart to expected utility of any given match. If $\nu > E(U)$, doing something else is considered more valuable than to attend and the individual have no incentives to attend the game. This assumes that every individual maximizes utility. An important and intuitive note should be stated about the nature of reservation utility: it is not the same for every individual. In accordance with Coates et al. (2014), we assume that ν is evenly distributed over $(\overline{\nu} - \underline{\nu})$. The most determined fans will have a low reservation utility $\underline{\nu}$ while others not so determined will have a high reservation utility $\overline{\nu}$. Those with a high reservation utility will also need a higher expected utility in order to attend a game, vice versa.

2.3 Types of Spectator Preferences

2.3.1 UOH and the presence of Reference-Dependent Preferences

As mentioned before the UOH supports the idea that fans prefer uncertain outcomes. That is when the home team is equally expected to lose as they are to win. In a context according to the model developed by Coates et al. (2014), UOH generates an expected utility function that is concave in p and reaches a maximum somewhere around $p = 0,5$. From equation (2) we can see that a concave function must satisfy the condition $(\beta - \alpha) < 0$. This states that the marginal impact of a positive deviation from the reference point is greater than the marginal impact of a negative deviation from the reference point. Put differently, an unexpected win results in higher utility than a unexpected loss. As stated earlier, the utility of a win is assumed to be greater than the utility of a loss ($U_w > U_l$), which means that the UOH-consistent preferences requires $(\beta - \alpha) < 0 \leq (U_w - U_l)$.

Figure 2. UOH-consistent preferences (Coates et al, 2014)

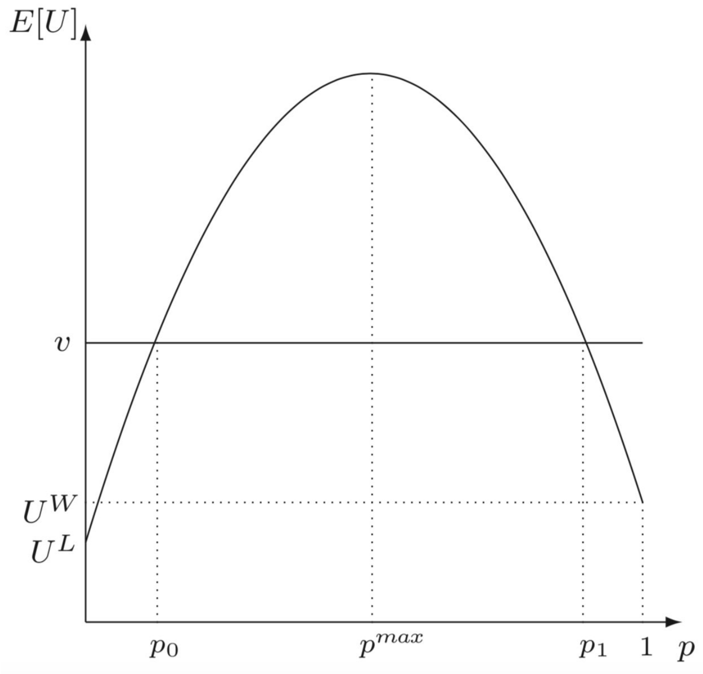


Figure 2 shows that expected utility for UOH-consistent preferences is concave and that expected utility peaks at p^{max}, which refers to a probability close to 0,5. Reservation utility is denoted by ν .As we can see there is a lower probability level (p_0) and an upper probability level (p_1) that shows when the individual decides to attend. Outside of these points an UOH-consistent individual won't attend the game as the reservation utility is larger than the expected utility of attending. In other words, this may be interpreted as a situation where the outcome is too certain.

2.3.2 Loss Aversion and the presence of Reference-Dependent Preferences

Acting almost as a polar opposite to the UOH, we have an individual which is loss averse. Those who are loss averse would rather see a game where the result is certain than one involving uncertainty. For an individual to be considered loss averse the marginal utility of an unexpected loss must be larger than that of an unexpected win, or put differently: $\beta > \alpha$. As mentioned before $\beta > \alpha$ is not consistent with the concave expected utility function that symbolizes UOH.

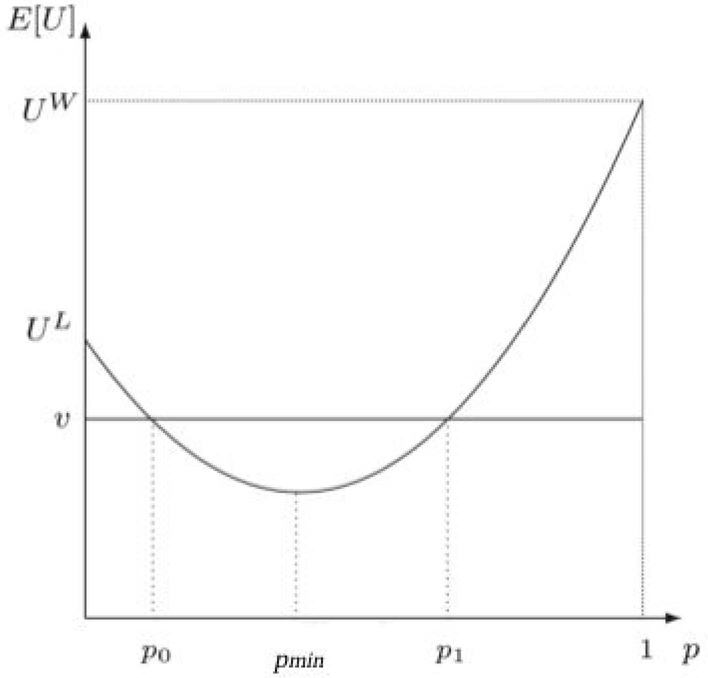
As the individual is loss averse an increase in p has two effects on expected utility. By rearranging equation (1) this intuition is put more clearly:

$$E[U] = [p\,U_w + (1-p)\,U_l] + (\alpha - \beta)\,p\,(1-p)$$

When p rises in figure 3 the expected intrinsic consumption utility $[p\,U_w + (1-p)\,U_l]$ increases while the gain-loss utility $(\alpha-\beta)\,p\,(1-p)$ first decreases to the point where $p = 0,5$ and then increases as p approaches 1. Consumption utility starts to dominate the effect of the gain-loss utility in the point $p_{\min}$.[5] When p reaches p_1, that is when $E(U)$ exceeds ν, a loss averse individual attends the game as the outcome is no longer that uncertain. Closer to $p = 1$, the expected utility of attendance grows. The interpretation of all this is fairly simple: when loss averse individuals perceive a game between two teams to be too evenly matched, due to a p that offers no clear winner, they will experience an uncertainty that is not to their liking. Because of this they will not attend.

[5] Occurs when p < (1/2 - (U_W- U_l)/(2(β-α))

Figure 3. Loss aversion-consistent preferences (Coates et al, 2014)[6]



An important note about note about the relationship between U_w and U_l is in place. If $(U_w - U_l)$ is sufficiently large and positive both figure 2 and 3 will be strictly increasing over the interval [0,1]. However, the relationship $(U_w - U_l)$ cannot be easily determined.

2.3.3 Absence of Reference-Dependent Preferences

There is a possibility that fans don't have reference-dependent preferences. Instead, they may have a more classical form of utility that is derived exclusively from the final outcome. In this case, $\alpha = \beta = 0$ and the expected utility function is

(3) $E(U) = (U_w - U_l)p + U_l$

This expected utility function increases with the probability of a home team win. A fan that experience utility exclusively from the consumption utility simply just want the home team to win. This would predict that teams with higher probability to win will have more attendance than teams with lower probability to win. This is in stark contrast to the conclusion of the

[6] Somewhat modified. Pmin is used instead of (1/2 - (U_W - U_l)/(2(**β-α**)).

UOH, where a team with a very high probability of winning will find themselves with low attendance. An individual with no reference-dependent preferences requires that $\beta - \alpha = 0$ and $U_w - U_l > 0$.

2.3.4 Special cases of preferences

An individual that is *risk averse* will, when faced with two choices with similar expected outcome, choose the one that is less risky. The presence of risk aversion would just add convexity to equation (3). Namely, the expected utility function would increase more and more as the probability of a win approaches 1.

Additional preferences may exist. Coates et al (2014) highlights the "pure fan", which refers to an individual that experience the same utility from a win as from a loss $(U_w - U_l)$ and whom who don't have reference-dependent preferences $(\beta - \alpha) = 0$. Even though Coates et al. (2014) addresses this as a "pure fan" it may also apply to an individual that is indifferent in viewing live-game football. For example, an individual that only attend in someone else's company. However, this is an expansion from our point of view. In accordance to Coates et al (2014) we assume that these special cases describe a smaller part of potential attendees at live-games and will therefore not be included in our following analysis.

3. Empirical method

This section will outline the method used to estimate preferences concerning uncertainty and its effect on live-game attendance. The first subsection highlights the econometric approach that is used to derive preferences in accordance with the theory. The second subsection contains an explanation of the data we have used. The third subsection presents possible estimation issues that have been considered.

3.1 Econometric application

To test which preferences characterizes the spectators of Allsvenskan an OLS method with panel data was used. The home team was set as panel and the corresponding round as time. Peel and Thomas (1992) put forth the hedonic idea that in a sport setting, a match can be viewed as a package. The package is determined by game specific factors as well as surrounding factors not directly related to the game, such as the weather or day of week. The specific factors related to the games are of interest. In particular, the characteristics of both teams individually and what they offer by facing each other. Following this idea, with support of equation 2 and Coates et al. (2014), the econometric function for estimating fan preferences and its effect on attendance can be determined. With the denotation i=hometeam, j=awayteam and t=time the econometric function can be denoted:

$$ln\ Attendance_{ijt} = \lambda + \theta p_{\mathrm{ijt}} + \gamma p^2_{\mathrm{ijt}} + X_{\mathrm{ijt}} + Z_{\mathrm{it}} + V_{\mathrm{jt}} + W_{\mathrm{ij}} + D_{\mathrm{j}} + D_{\mathrm{t}} + \varepsilon_{\mathrm{ijt}}$$

The natural logarithm of attendance is used to obtain the effect of the independent variables on attendance in percentage rather than an absolute number (Gujarati, 2002). In our case, using a natural logarithm also had the benefit of causing the variable to become more normally distributed, as it was right skewed prior to implementing the logarithm. In this function p stands for the probability of a home win and p^2 is the squared probability. Probability p is squared to capture the possible non-linear nature of the uncertainty generated by the probability of a win. As presented above in figure 2 and 3, preferences consistent with the UOH or loss aversion will lack linearity. Further, will p be referred to *Prwin* and p^2 to *Prwin2*. To go along with the idea of a match being viewed as a package, we have included variables that depend on one or a combination of i, j and t. As each home team differs in characteristics a fixed effect model is used to illustrate team specific attributes such as the market facing each team, reputation, history and economic standing. The same is true for each away team and a dummy variable is therefore included for each visiting team. Attendance may depend on which weekday a game is played but may also vary from season to season. To compensate two additional time dummies are included. The team specific variables that depend on time are the standings of the *Previous Season* for both the home team and the away team. *Average Ticket Price* (*ATP*) and *Geographical Riva*l are included to illustrate the effect of two specific teams facing each other. To clarify, the following more precise model will be used for the purposes of this study:

$$\ln Attendance_{ijt} = \lambda + \theta\, Prwin_{ijt} + \gamma\, Prwin2_{ijt} + \delta_1 ATP_{ijt} + \delta_2 PreviousSeasonHome_{it} + \delta_3 PreviousSeasonAway_{jt} + \delta_4 GeographicalRival_{ij} + \delta_5 Awayteam_{j} + \delta_6 DayOfweek_{t} + \delta_6 Season_{t} + \varepsilon_{ijt}$$

As will be made clear further down in this section, the interesting parameters in our study are θ and γ. Derived from equation 2 and as proposed by (Coates et al, 2014) these parameters are given by:

(4) $\gamma = (\beta - \alpha)/(\overline{v} - \underline{v})$

(5) $\theta = [(U_w - U_l) - (\beta - \alpha)]\ /(\overline{v} - \underline{v})$

Parameters from the regression model presented above can be used to test the different preferences highlighted in the theory section. To do this, the following hypotheses about how preferences concerning uncertainty affects live-game attendance in Allsvenskan will be tested (Coates et al, 2014):

H1: $\gamma > 0$ implies that $\beta > \alpha$ from equation (4). This refer to a marginal consumer that is loss averse as the marginal impact of a loss (β) is larger than the marginal impact of a win (α).

H2: $\gamma = 0$ and $\theta > 0$ implies that $\beta - \alpha = 0$ from equation (4) and that $U_w - U_l > 0$ from equation (5). This refer to a marginal consumer that does **not** have reference-dependent preferences and gets more utility of a home team win than a home team loss. If H2 is rejected, it means that there are reference dependent preferences.

H3: $\gamma < 0$ and $\theta > 0$ implies that $\beta - \alpha < 0$ from equation (4) and that $U_w - U_l > 0$ from equation (5). This refer to an UOH-consistent marginal consumer which prefer uncertain outcomes. That is when the home team is equally expected to lose as they are to win.

3.2 Data

The primary variables in our type of study is attendance statistics and odds figures for each match. To put these variables in context we have included variables that have proven to be significant predictors of attendance in the literature highlighting football.

3.2.1 The Swedish Allsvenskan

Allsvenskan is the top professional football league in the Swedish football league system. The league contains 16 clubs and the competition follows the classic Double Round-Robin format. Meaning that teams are faced against each other twice during a season; once at home and once away. This sums up to a total of 30 matches per team and season, equally divided between 15 home games and 15 away games. Teams receive 3 points for a win, 1 point for a draw and no points for a loss. Teams are ranked by their taken points during the season and the one that ends up with the most wins the league (Allsvenskan, 2017). The winner of Allsvenskan qualifies for the Champions League, while the second and third placed teams qualifies for the Europa League. Teams placed 15-16 are relegated to a lower division while the team placed 14 must qualify to stay (Svensk fotboll, 2017).

3.2.2 Attendance data

Attendance data was collected from game-by-game statistics presented by the official website of Allsvenskan (Allsvenskan, n.d). In order to achieve a balanced panel, we have chosen to only consider teams that have been part of the three seasons, spanning from 2014 to 2016. During these years, a total of 20 teams have been part of Allsvenskan and 12 teams have participated in all three seasons. Table 1 provides an overview of the attendance data and teams considered in in this thesis. Data on stadium capacity is also provided in the same table. Official data on this have been collected from The Swedish Football Association and can be found on the respective team website on svenskfotboll.se.

Table 1. Attendance statistics Allsvenskan 2014-2016

Club		2014	2015	2016	Total
AIK	Total	190 833	221 144	166 846	578 823
	Mean	17 348	20 104	15 168	17 540
	Max capacity	50 000	50 000	50 000	50 000
	Observations	11	11	11	33
Häcken	Total	36 294	40 476	37 115	113 885
	Mean	3 299	3 680	3 374	3 451
	Max capacity	18 416	18 416 / 6 500	6 500	12 061
	Observations	11	11	11	33
Djurgården	Total	150 803	165 888	147 833	464 524
	Mean	13 709	15 081	13 439	14 076
	Max capacity	30 000	30 000	30 000	30 000
	Observations	11	11	11	33
Falkenberg	Total	39 541	40 429	35 041	115 011
	Mean	3 595	3 675	3 186	3 485
	Max capacity	4 000	4 000	4 000	4 000
	Observations	11	11	11	33
Gefle	Total	41 627	46 100	35 614	123 341
	Mean	3 784	4 191	3 238	3 738
	Max capacity	6 703	6 703 / 6 432	6 432	6 540
	Observations	11	11	11	33
Göteborg	Total	127 443	155 141	131 045	413 629
	Mean	11 586	14 104	11 913	12 534
	Max capacity	18 416	18 416	18 416	18 416
	Observations	11	11	11	33
Helsingborg	Total	92 338	91 006	80 828	264 172
	Mean	8 394	8 273	7 348	8 005
	Max capacity	16 000	16 000	16 000	16 000
	Observations	11	11	11	33
Elsborg	Total	93 420	103 410	81 196	278 026
	Mean	8 493	9 401	7 381	8 425
	Max capacity	16 200	16 200	16 200	16 200
	Observations	11	11	11	33
Kalmar	Total	69 400	68 751	63 557	201 708
	Mean	6 309	6 250	5 778	6 112
	Max capacity	12 000	12 000	12 000	12 000
	Observations	11	11	11	33
Malmö	Total	165 294	194 984	195 326	555 604
	Mean	15 027	17 726	17 757	16 836
	Max capacity	24 000	24 000	24 000	24 000
	Observations	11	11	11	33
Norrköping	Total	66 822	105 827	114 354	287 003
	Mean	6 075	9 621	10 396	8 697
	Max capacity	15 734	15 734	15 734	15 734
	Observations	11	11	11	33
Örebro	Total	76 349	77 981	82 328	236 658
	Mean	6 941	7 089	7 484	7 171
	Max capacity	12 645	12 645	12 645	12 645
	Observations	11	11	11	33
Total	Total	1 150 164	1 311 137	1 171 083	3 632 384
	Mean	8 713	9 933	8 872	9 173
	Max capacity	-	-	-	-
	Observations	132	132	132	396

3.2.3 Odds and Probabilities

To calculate the probability of the home team winning we have used betting odds. This since there are robust empirical evidence that betting odds are the best publicly available source when predicting probability forecasts for sports (Strumbelj, 2014). The odds used for the

2014, 2015 and 2016 season are all collected from betexplorer.com, which creates aggregated odds based on 20+ bookmakers. Betexplorer reports the latest pre-game odds available from the different bookmakers, often reported minutes before kickoff.

Converting these odds into their corresponding probability was done using the basic normalization method. This method has been largely applied within the literature (Franck et al., 2010). In football there are three possible outcomes ε, i.e. win, draw and loss. Therefore, there are three different decimal odds to bet on; O_w , O_d and O_l . In order to make profit, the bookmakers set unfair odds. This can can be seen as the sum of the inverse odds (booksum) $\Sigma \frac{1}{O\varepsilon}$ usually ends up larger than 1 due to the so-called overround or bookmaker margin. Because of this, the inverse of the odds cannot directly be interpreted as probabilities. To cope with this, the basic normalization method assumes that the overround is equally divided on the three different outcomes. Therefore, we derive the probability of the home team winning by the formula:

$$(6)\ Prwin = \frac{1/Ow}{(1/Ow + 1/Od + 1/Ol)}$$

Even though the method has been used frequently in the literature there are some arguing that there may be better ways to forecast probabilities. Mainly, this is based on the findings of the favourite longshot-bias within football and other sports (Cain et al., 2000; Vlastakis et al., 2009). Basically, the phenomenon says that bettors tend to overvalue long shots and undervalue favourites and that bookmakers sets odds accordingly. Strumbelj (2014) shows that Shin probabilities may be a more accurate way to forecast probabilities. Due to the complex calculations and the small differences in probabilities comparing to the basic normalization method the authors however states that the basic method is a strong alternative and sometimes preferable. Because of this, and since Shin probabilities are seldom used in the literature, we have chosen to use the traditional and well-accepted basic normalization method.

3.2.4 Control variables

Geographical rival is a dummy variable defined in this study as when one team faces an opponent considered to be a geographical neighbor. A match will be considered to be played between rivals if the teams are from the same city or from two cities in close proximity. Due to our definition of a geographical rival, every team in the original dataset had at least one geographical rival. As teams were dropped since they didn't played in Allsvenskan the whole period, a few remaining observations lost their geographical rival. Elfsborg, Gefle and Falkenberg have two geographical rivals each as their closest neighboring city contains two teams. All other teams have one or none geographical rival. This variable is supposed to take care of the well-known derby effect that often drives up attendance.

The outcome of the previous season for both the Home and the Away team was used to illustrate the effects of performance on attendance. UEFA (2017) reports a clear link within European football between attendance and the previous seasons league position. On average, moving one position up the league increases attendance by 3% while moving one position down decreases attendance by 3%. The previous season is used instead of the form of the current season to avoid possible collinearity. This is due to the current season already being taken into account by the use of the odds. The previous season offers a measurement of the team´s form in a more historic way than the recent performance included in the odds.

Average ticket price (ATP) is derived for each team and every season by using the data on total ticket revenue for each season. This data is derived from the report presented by EY (2017) and the reports presented by Deloitte (2015, 2016), in which the financial standing in Allsvenskan for each season is summarized. To calculate ATP, we divided total ticket revenue with total attendance. As the public available ticket revenue includes ticket revenues obtained from cup matches we have included attendance from these matches to get total attendance during the season. [7]

[7] Cup matches in our data set included Champions League, Europa League and Svenska Cupen.

3.3 Estimation issues

Panel data is a combination of time series data and cross-sectional data, leading to a number of challenges when used (Gujarati, 2002). The problems that may infect cross-sectional data (e.g. heteroscedasticity) and time series data (e.g. autocorrelation) must be taken care of. These problems are also present in the Gauss-Markov theorem, which states under which conditions the linear OLS regression is the best linear unbiased estimator (BLUE) with the lowest variance (Wooldridge, 2012). For the Gauss-Markov theorem to hold the following assumptions must be in place: Non-heteroscedastic error terms, no autocorrelation in the error terms, exogenous explanatory variables and finally, all data must be stationary (Wooldridge, 2012).

Heteroscedastic error terms are found when the variance in the error terms are not constant. The consequence of this is incorrect standard errors and as a result, test such as t-test and F-tests return faulty conclusions. Estimations would therefore end up as ambiguous (Wooldridge, 2012). To test for heteroskedasticity we conducted a modified Wald test which pointed towards heteroskedasticity, for test results see Appendix[8]. The presence of heteroskedasticity was corrected using White's standard errors. Of note is that the use of the natural logarithm of the dependent variable can act as a remedy for heteroscedasticity as well (Gujarati, 2002). However, heteroscedasticity was not completely removed as evident by the test, calling for the use of the standard errors.

Autocorrelated error terms has the same effect on standard errors as heteroscedasticity and occurs when error terms are correlated with others at different points in time (Wooldridge, 2012). The Wooldridge test conducted for autocorrelation showed no autocorrelation to be present.

An exogenous explanatory variable is a variable that is not affected by other variables in the system. In contrast, an endogenous variable is a variable that is influenced by other variables within the system. Endogenous variables cause issues as estimations won't be right on average (Gujarati, 2002). In studies on demand, a possible endogeneity problem may occur between price and demand. There is a possibility that attendance is affected by ATP and that

[8] All test results can be found in the Appendix section.

ATP is affected by attendance. For example, profit-maximizing teams in areas with greater population will set higher prices due to their market size. Such a problem would generally need an instrument variable, i.e. a variable that affects ATP but not attendance (Gujarati, 2002). Stationary data is equivalent to when expected value and variance is constant over time for a panel (Gujarati, 2002). An important note regarding the Gauss-Markov theorem is that while it is very important it is not necessary for it to hold completely. A regression can still be of value and return accurate estimations without it being BLUE (Wooldridge, 2012).

Additional problems that need to be taken care of when using panel data is whether the error term ε may be correlated with the regressors. If ε is assumed to be uncorrelated with the regressors the random effects model is preferable and if the contrary holds true, i.e. ε is correlated with the regressors the fixed effects model is preferable (Gujarati, 2002). As we assumed the latter and to cope for team specific characteristics we used the fixed effects model in our study. A Hausman-test was also conducted and pointed to the use of the fixed effect model being correct, for test results see appendix. The fixed effect model controls for time-invariant differences in observable and unobservable factors that varies between teams. The fixed effect model reduces the threat of omitted variable bias and the possible endogeneity bias described (Dranove, 2012). This is only true under the assumption that unobservable factors, for example respective teams market size and supporter culture, not vary over time. Given the time spectra of the study, this seems as a plausible assumption.

Except for the OLS model, other econometric models have been used in the literature concerning drivers of attendance. A sometimes-occurring model is the censored Tobit (Buraimo & Simmons 2008, Welki & Zlatoper, 1994). This model differs from the OLS as it has different goals and follows a maximum likelihood estimation. This model may be appropriate when a significant part of the games are sell-outs, in Buraimo & Simmons (2008) 58 % of the observations was designated to being censored and in Welki & Zlatoper (1994) the share was 61%. The reasoning behind this stems from the assumption that the observed attendance is the true effective demand. If a large amount of the games are sell-outs one could expect the true effective demand to exceed the observed demand. If this is true, OLS is known to yield biased estimates (Greene, 1997). As only 12 observations (3%) in our data set did sell-out we found no need for using the censored Tobit model. As the Tobit relies on stricter assumptions than the OLS it would only complicate the interpretations of the study.

Finally, one potential issue would be misspecification. Misspecification would be if one were to choose a linear model when the correct option would have been a non-linear model (Dougherty, 2011). Testing for this issue was done using a Ramsey RESET test. The test confirmed our model to be correctly specified as a linear model.

4. Previous research

The literature on variables that affect live-game attendance in sports events are plentiful. A smaller part of this literature is focused on the nature of reference dependency and the UOH. However, as measures of uncertainty often are included in these studies, conclusions of preferences can be drawn even though the particular study´s main focus may be something else. This section serves to offer a glimpse into some findings made by other authors particularly focusing on football. The observed preferences differ among leagues and countries, which makes the study of Allsvenskan even more interesting as the literature on the league combined with uncertainty is uncharted, at least to our knowledge.

In 1988, Peel & Thomas estimated the probability to win by using the odds to investigate the UOH in the top four divisions of English football. In this paper they only included a linear measure of probability to win, i.e. not the squared probability. They found the measurement to be significant in some respect for the four divisions of English football they studied. The positive sign on home win probability pointed towards preferences according to H2. However, when the same authors examined the same league but at a different time, now together with the squared probability, they found support for H1 (Peel & Thomas, 1992). Both studies used the ordinary least squares (OLS) model in order to estimate the parameters.

More than a decade later, Forrest et al. (2005) looked at the top three divisions of English football. Instead of using odds they derived probability ratios by comparing the underlying quality of the teams and home team advantage. By using an OLS and both linear and quadratic measures of probability to win, they found results that strengthens the idea of loss averse preferences (H1) and that close competition may decrease attendance. The same pattern found by Peel & Thomas (1992).

Jena and Reilly (2016) used a very similar approach to that of Coates et al. (2014) when looking at the Irish league, The League of Ireland. In contrast to Coates et al. (2014) which used a censored Tobit, Jena and Reilly used the OLS. By obtaining the probabilities of a home team victory from posted odds, the UOH was found to apply. H3 is therefore applicable to the Irish league as γ was below zero while the opposite was true for θ.

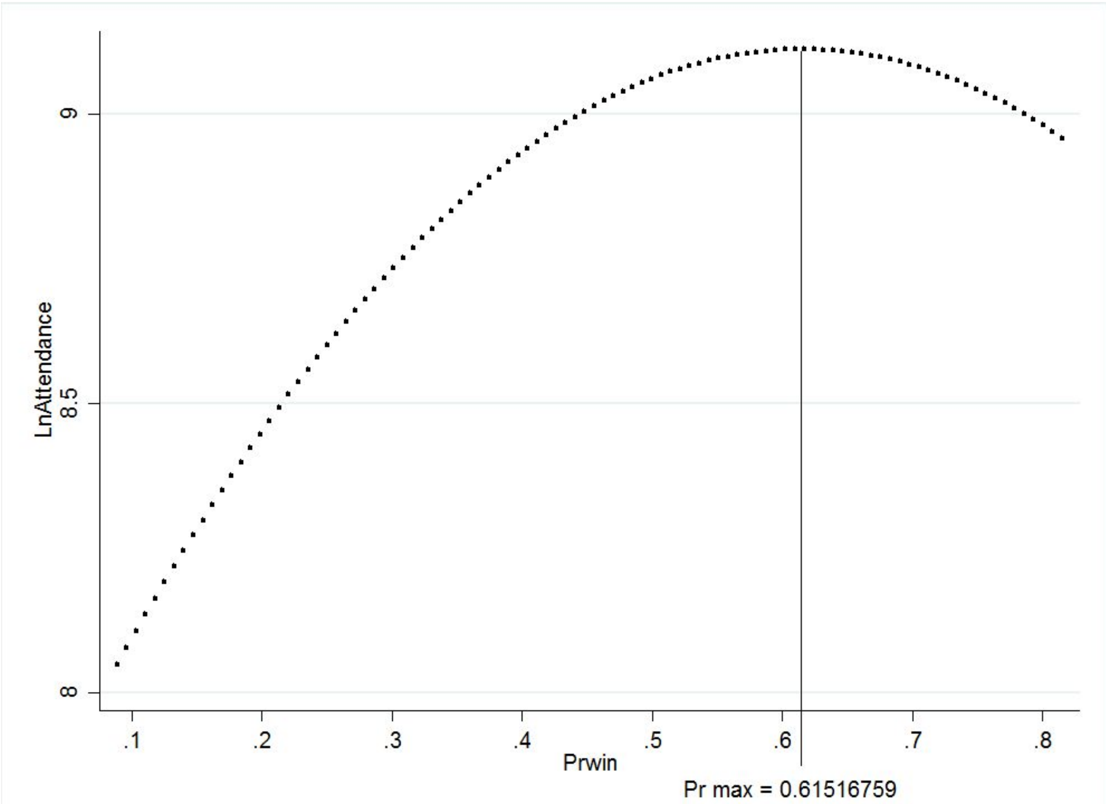
In the same year Schreyer et al. (2016) published their findings on game uncertainty in the German Bundesliga. They studied how uncertainty affects attendance by taking a closer look at the behavior of season pass holders. Uncertainty was proxied by probability to win as derived from betting odds. Using a series a different regressions, including an OLS as well as a quadratic, they found support for the UOH (which they called GOU for Game Outcome Uncertainty). This support was not only for those with a season pass but a general finding for all attendance within the league.

By using a collection of uncertainty variables not derived from the odds, such as position in relation to opponent and chance of being in first place, Madalozzo & Berber Villar, (2009) studied the Brazilian football league by using an OLS in panel form in search of the effects of uncertainty. They found no significance for their p-values. A similar situation is reported in Czarnitzki and Stadtmann (2002), where the focus is on the German Bundesliga. Using the probability of a home team win as well as the probability squared derived from the odds, they also found insignificant results. Ultimately this leads to the conclusion that uncertainty has a minor effect in these particular cases.

5. Result

The estimated results of the fixed effect OLS model can be found in table 2. Summary statistics of the variables is found in table 3. Both coefficients of interest, *Prwin* and *Prwin2*, is significant on a 5% significance level. The signs of the coefficients show clear support for H3, i.e. the UOH. This suggest that fans in Allsvenskan have reference dependent preferences, that they prefer tight games and experience a concave utility function as the probability of a win goes from 0 to 1. Figure 4 shows the relationship between the natural logarithm of attendance, *Prwin* and *Prwin2*.

Figure 4. Quadratic relationship between LnAttendance, Prwin and Prwin2



As predicted by the UOH, our plotted line suggests an inverse u-shaped relationship were more uncertain outcomes generate higher attendance in the setting of Allsvenskan. The graph reaches its maximum in *Prwin*=0,615 and thereafter suggests a decreasing attendance as the probability to win approaches 1. This also indicates that highest attendance is given when the home team is slightly more likely to win than to lose or end in a draw. The support of the UOH is in line with the findings of Jena & Reilly (2016) were the Irish League was examined. In this setting a turning point of 0,64 was found to apply. Our findings differ from the results of Coates et al. (2014), which this study is based on. Coates et al. (2014) found support for H1 and loss averse reference-dependent preferences which is in line with the findings within the divisions of English football.

Table 2. Estimation results from the fixed effect OLS model using robust standard errors. Dependent variable: LnAttendance.

Variable	Estimates	Variable	Estimates
Prwin	1.667** (.684)	Dow	
Prwin2	-1.499** (.676)	Sunday	‡
Previous_season_home	.009 (.008)	Monday	-.022 (.031)
Previous_season_away	.003 (.003)	Tuesday	-.076 (.138)
Geographical_rival	.337*** (.054)	Wednesday	.006 (.038)
ATP	-.002 (.001)	Thursday	-.119* (.061)
		Friday	.056 (.050)
Awayteam		Saturday	.031 (.036)
AIK	‡		
BK Häcken	-.433*** (.041)	Year	
Djurgårdens IF	-.029 (.062)	2014	‡
Falkenbergs FF	-.396*** (.104)	2015	.140*** (.039)
Gefle IF	-.482*** (.082)	2016	.024 (.045)
Helsingborgs IF	-.284*** (.050)		
IF Elfsborg	-.133 (.078)	cons	8.769*** (0.241)
IFK Göteborg	.079 (.089)		
IFK Norrköping	-.252*** (.041)		
Kalmar FF	-.389*** (.054)		
Malmö FF	.099** (.038)		
Örebro SK	-.412*** (.069)		
Observations	396		
Number of teams	12		
Observations per team	33		
R-sq within	0.5588		

Note: Robust standard errors within parentheses; ***, ** and * denotes significance at the 0.1, 0.05 and 0.01 levels; ‡ indicates the reference category.

Table 3. Summary statistics of the variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Prwin	396	0.455	(.159)	0.089	0.815
Prwin2	396	0.232	(.150)	0.008	0.664
Previous_season_home	396	7.833	(4.310)	0	14
Previous_season_away	396	7.833	(4.310)	0	14
Geographical_rival	396	0.091	(.288)	0	1
ATP	396	123.778	(36.757)	55	193

Note: The variables Previous_season_home and Previous_season_away is coded as 0 if the corresponding team played in a lower division the previous season, 1 if they were placed 14th, 2 if they were placed 13th and so on.

There were also other variables that was found to have a significant impact on attendance. The variable Geographical_rival show strong significance on attendance. The coefficient 0,3372 implies that when facing a geographical rival, attendance increases with 33,72%. This is in line with the well-known derby effect presented in the literature. Our definition of a geographical rival is somewhat broader than the usual definition. This highlights the fact that in the case of Allsvenskan, the derby effect not only exists between teams from the same city but also between teams from cities in close proximity.

Some time variables had a significant effect on attendance. One day of the week differed from the reference category Sunday with a marginal significance, namely Thursday. The negative coefficient of Thursday states that attendance decreases by 11,93% comparing to the attendance figures on games played on Sundays. However, this effect was only significant on a 10% significance level. The season of 2015 stands out as its significantly different from the reference category 2014. This may have different explanations but generally the attendance figures of Allsvenskan suffers during years when national championships are played (Fotbollskanalen, 2016). Championships on a national scale take place every other year as was the case during 2014 and 2016. Another possible explanation may be that the popular and attendance-strong team Hammarby entered the scene 2015 after five years of struggling in lower divisions. The reasoning behind this stems from the strong supporter rivalry within the league and that the other supporters may have been triggered by the entrance of Hammarby. This indicates a reality where the teams not only compete with their success on the pitch but also on the grandstand.

Looking at the dummy variable of *Awayteam,* Malmö FF is the only team that has a significant positive effect in relation to the reference team AIK. This points towards that Malmö FF is the team that drives the most attendance as it is playing away. IFK Göteborg does also show a positive sign compared to AIK but as the estimates is insignificant one could not draw statistical conclusions of this. This is however interesting results as by the end of the 2016 season Malmö FF was first, IFK Göteborg second and AIK in third place in the

Marathon table of Allsvenskan, a table displaying total points of all teams that has participated in Allsvenskan since 1924 (Allsvenskan, 2016). Djurgården and Elfsborg show negative but insignificant differences compared to AIK which points to their similarities in driving up attendance as they play away. The eight remaining teams is significantly different compared to AIK. All of them show negative coefficients, meaning that they in total drive less attendance then AIK when playing away. Malmö FF has been considered a top tier team in Allsvenskan for some time now and the 2017 title was the fourth in five years. This recent success may be the reason for Malmö FF having a more substantial effect on attendance as a visiting team than other top tier teams. Finally, it is important to note that Malmö FF, IFK Göteborg, AIK and Djurgården all hale from large cities in a Swedish context. The larger populations of these cities likely add to their influence on attendance in relation to others, as well as causing them to display similar effects to each other, with Malmö FF pulling ahead for previously stated reasons. More unexpectedly is that Elfsborg show similarities with these teams as they origin from a smaller city.

Somewhat unexpectedly, the control variables ATP, Previous_season_home and Previous_season_away show insignificant estimates on LnAttendance. In a variety of sports, ATP is commonly found to have a negative impact on attendance (Borland & Macdonald, 2003). The insignificance of ATP points to the price not being a determining factor in whether to attend a game or not in Allsvenskan. Previous_season_home was expected to have a significant positive impact on attendance. This as fans can buy seasonal tickets before the start of the season and our expectations was that a higher ranking the previous year would imply more tickets being sold the following year. The insignificance of the previous home season may be due to the difference in purchasing a season ticket and actually attending. If the success of the previous season doesn't carry into the present season, holders of seasonal tickets may choose to not attend as the utility of staying home is greater (Dobson & Goddard, 2011).

6. Conclusion

The investigation into fan preferences in Allsvenskan during the period of seasons 2014-2016, with odds as a measurement of uncertainty, concludes that the UOH applies. This offers the answer to our original question: *What preferences concerning uncertainty affect the live-game attendance in Allsvenskan?*. To clarify, this means that the general consumer of Allsvenskan prefer games in which the outcome is uncertain. This conclusion has been based on the theoretical framework developed by Coates et al. (2014) in which fan preferences have been given a theoretical basis, something that was lacking in previous literature.

Our results show that highest attendance is given when the home teams probability to win is 61,5%. Coates et al. (2014, p.963) proposes that the UOH is consistent with a concave expected utility function that reaches its maximum *"in a value greater than or equal to 0,5 and substantially less than 1"*. Due the fact that we studied a sport with three outcomes and that Coates et al. (2014) studied a sport with two outcomes, the interpretations are somewhat different. It should be noted that when the home team's probability to win is around 61,5% the corresponding probabilities to a draw is in general 23 % and to an away win 17 % in our dataset. This clearly doesn´t refer to a situation where both teams are more or less as likely to win. A reminder of our discussion in the theory section is in place. Due to the advantage of being a home team a draw for the home team can be seen as a loss and being deprived of the full award of winning. Due to the disadvantage of being the away team a draw for the away team can be seen in the light of a win. This as the away team not only brought with them a point, but also took important points from their opponents when they had the advantage. This reasoning clearly evens out some of the difference between the teams winning probabilities and making the interpretations of the UOH more compatible within a sport with three possible outcomes. Even though we find support for uncertainty driving up attendance, we also conclude that fans in Allsvenskan put substantially more value in wins than losses as maximum attendance is found in a point where the home team is considered to be in favor.

Our results fall in line with the conclusion draw by Jena & Reilly (2016) in the Irish league, as well as the research done by Schreyer et al. (2016) on the German Bundesliga. However, it is important to again note that Schreyer et al. (2016) use a very different method from ours making direct comparisons somewhat difficult. The results contradict the findings within the English divisions (Peel & Thomas, 1992; Forrest et al., 2005) that found support for loss aversion but also the studies of the Brazilian league (Madalozzo & Berber Villar, 2009) and the Bundesliga (Czarnitzki & Stadtmann, 2002). This emphasizes the importance of cultural differences and that fan preferences varies within leagues and by time. As previously stated in section four, there is to our knowledge no other studies conducted on Allsvenskan except for this one, making any comparisons impossible.

Due to our findings, which are in line with the Uncertainty of Outcome Hypothesis, one may argue that the recent “financial fair play” regulations of the European football may be motivated. These regulations were implemented by UEFA by the end of the 2012 season with the two principal goals of rebalancing competition and to ensure long-term financial stability of European club football. Particularly, the financial fair play regulation includes limitations on club’s deficits and the opportunities to acquire external funding (Vöpel, 2011). As the regulations affects all European clubs the smoothening of competition is expected both between leagues as a whole but also between clubs within leagues. However, as proposed by Vöpel (2011), additional redistributional regulations may be needed in order to solve the lack of competition found within leagues.

As sporting successes drives the financial conditions of a club, it can attract better players and managerial staff, resulting in an increased probability of further success. This is highly applicable in Allsvenskan due to the domination of Malmö FF within recent years. To maintain the positive attendance trend within Allsvenskan competitive balance is necessary and the Swedish football federation need to take this into account when trying to develop the league. An interesting development can be found in the Danish Superliga where a new league system, including playoffs, recently have been implemented. In theory, this will increase the overall excitement as there is more to play for. However, attendance figures has dropped since the change and arrangements like these have to be carefully thought out. This thesis also points towards that fan preferences varies between leagues and cultures, which means that an arrangement like this can’t directly be copied from another league or culture. To achieve desired results, it has to be based on the preferences present in the concerned league. This points to an important note: conclusions made of one league may not be applicable to other or even similar leagues. Acquiring a general result that can benefit all of football would be ideal but the required resources to study this would reach extreme numbers. Another issue concerning research regarding uncertainty in sports is that there is no collectively accepted model. Standardisation would offer the benefit of more reasonable comparisons. As the model in Coates et al. (2014) is fairly new, it is our hope that it will become accepted as a standardized framework, possibly enabling general conclusions to be made.

It should also be noted that league organizers and teams may have different goals. League organizers is probably more likely to show interest in a competitive balance as they are responsible for the prosperity of the league. On the contrary, a team's ultimate goal is probably sporting success. However, this is more likely to happen with a crowd backing them up. Meaning that there are incentives for teams to strive for competitive balance as well.

Except for the need of more research into Allsvenskan to enable comparisons, future research into this matter would be wise to include additional teams and seasons to obtain more general results. As we dropped teams that didn't played in Allsvenskan the three seasons, there is a possibility that the results would have been different when including them. Our feeling is though that the results would only differ slightly and the added complexity would be outside the boundaries set by this thesis being at a bachelor level. The inclusion of additional seasons would also be beneficial as one could study changes in preferences over the years and possibly use it to find trends or patterns. With this, conclusions about future preferences could be drawn. One way to include more teams that could potentially reap further benefits is to include Superettan, the league below Allsvenskan. Doing so would bridge the gap between the two tiers and account for teams transitioning from one to the other, making the relative number of teams that are left out substantially smaller. It is however important to note that this would broaden the study beyond our scope and place less emphasis on Allsvenskan.

References

Borland, J. & Macdonald, R. (2003). *Demand for sport.* Oxford review of economic policy. Vol. 19, No. 3, pp 478-502.

Buraimo, B. & Simmons, R. (2008). *Do Sports Fans Really Value Uncertainty of Outcome? Evidence from the English Premier League.* International Journal of Sport Finance, 2008, No. 3, pp. 146-155.

Cain, M, Law, D & Peel, D (2000), The favourite-longshot bias and market efficiency in UK football betting, *Scottish Journal Of Political Economy*, 47, 1, pp. 25.

Card, D. & Dahl, G. (2009). *Family violence and football: the effect of unexpected emotional cues on violent behavior.* NBER Working Papers 15497, National Bureau of Economic Research, 2009.

Coates, D, Humphrey, B & Zhou, L. (2014). *Reference-dependent preferences, loss aversion, and live game attendance.* Economic inquiry, Vol. 52, No. 3, pp. 989-973.

Czarnitzki, D, & Stadtmann, G (2002). *Uncertainty of outcome versus reputation: Empirical evidence for the First German Football Division*, Empirical Economics, Vol. 27, No. 1, pp. 101-112.

Dougherty, C (2011). *Introduction To Econometrics*, Oxford : Oxford University Press, cop. 2011.

Dobson, S.M, Goddard, J.A (2011). *The Economics of football***,** New york: Cambridge University Press

Forrest, D. Beaumont, J. Goddard J & Simmons, R. (2005). *Home Advantage and the Debate About Competitive balance in Professional Sports Leagues.* Journal of sports science. Vol. 23, No. 4, pp 439-435.

Franck, E. Verbeek, E & Nüesch S. (2010). *Prediction accuracy of different market structures - bookmakers versus a betting exchange.* International Journal of Forecasting. Vol. 26, No. 3, pp. 448-459.

Greene, W.H. (1997). *Econometric Analysis* , Upper Saddle River, New Jersey :Prentice-Hall

Gujarati, DN (2002). *Basic Econometrics*, Boston : McGrawHill

Jena, F, & Reilly, B (2016). *Testing the uncertainty outcome hypothesis using data from second tier soccer in Ireland*, Applied Economics Letters, Vol 23, No.18, pp. 1257.

Kőszegi, B & Rabin, M (2006). *A Model of Reference-Dependent Preferences*, The Quarterly Journal of Economics, Vol. 121, No. 4, pp. 1133–1165.

Madalozzo, R, & Berber Villar, R (2009). *Brazilian Football What Brings Fans to the Game?*, Journal Of Sports Economics, Vol. 10, No. 6, pp. 639.

Peel, D, & Thomas, D (1988). *Outcome Uncertainty and the Demand for Football: an Analysis of Match Attendance in the English Football League*, Scottish Journal Of Political Economy, Vol. 35, No. 3, pp. 242-249.

Peel, D. & Thomas, D (1992). *The demand for football: some evidence on outcome uncertainty* . Empirical Economics, Vol. 17, No. 2, pp. 323-331.

Rottenberg, S. (1956). *The Baseball Players' Labor Market.* Journal of Political Economy, Vol. 64, No. 3, pp. 242–258.

Schreyer, D, Schmidt, S, & 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, pp. 192-217.

Strumbelj, E. (2014). *On determining probability forecasts from betting odds.* International journal of forecasting, Vol. 30, No. 4, pp. 934-943.

Vlastakis, N, Dotsis, G, & Markellos, R (2009). *How efficient is the European football betting market? Evidence from arbitrage and trading strategies*, Journal Of Forecasting, 28, 5, pp. 426-444.

Vöpel, H. (2011). *Do we really need financial fair play in European Club Football?* An Economic Analysis, CESIfo DICE Report.

Welki, A & Zlatoper, T (1994). *US Professional Football: The Demand for Game-Day Attendance in 1991*, Managerial And Decision Economics, 5, pp. 489.

Wooldridge, J.M. (2012). *Introductory Econometrics : A Modern Approach*, Mason, OH : South Western, Cengage Learning.

Online references:

Allsvenskan (2017). *Om oss*, Available Online:
https://www.allsvenskan.se/om-oss/

Allsvenskan (n.d) *Arkiv 2016, 2015, 2014,* Available online:
https://www.allsvenskan.se/tabell/arkiv-2016/
https://www.allsvenskan.se/tabell/arkiv-2015/
https://www.allsvenskan.se/tabell/arkiv-2014/

Allsvenskan (2016), *Maratontabellen 1924-2016*, Available Online:
https://www.allsvenskan.se/maratontabell/

Deloitte (2015) Penningligan - *Årlig genomgång av svensk fotbollsekonomi säsongen 2014,* Available online:

https://www2.deloitte.com/content/dam/Deloitte/se/Documents/about-deloitte/Penningligan%20Allsvenskan_2015_19-.pdf

Deloitte (2016) *Penningligan - En inblick i den svenska elitfotbollens ekonomi säsongen 2015*, Available online:
https://www2.deloitte.com/content/dam/Deloitte/se/Documents/about-deloitte/penningligan-sasongen-2015.pdf

Dranove, D. (2012) *Practical Regression: Fixed Effects Models*. Case #7-112-005 Kelogg School of Management. Available online:
Practical Regression: Fixed Effects Models - Northwestern University

EY (2017) *Hur mår svensk elitfotboll? en analys av den finansiella ställningen i Allsvenskan,* Available online:
http://www.ey.com/Publication/vwLUAssets/Hur_mar_svensk_elitfotboll/$FILE/EY_%20Sport%20Business_Fotboll_LR_single_2017.pdf

Fotbollskanalen (2016) *Allsvenskans stora publiksuccé - nära rekordsiffror,* Available online:
https://www.fotbollskanalen.se/allsvenskan/allsvenskans-stora-publiksucce---nara-rekordsiffror/

UEFA (2017) *The European Club Footballing Landscape - Club Licensing Benchmarking Report Financial Year 2015.* Available Online:
http://www.uefa.com/MultimediaFiles/Download/OfficialDocument/uefaorg/Finance/02/42/27/91/2422791_DOWNLOAD.pdf

Svensk fotboll (2017). *Allsvenskan - men*, Available Online:
http://svenskfotboll.se/in-english/domestic-football/allsvenskan-men/

Appendix

Hausman test:

```
. hausman fe re

                 b = consistent under Ho and Ha; obtained from xtreg
      B = inconsistent under Ha, efficient under Ho; obtained from xtreg

   Test:  Ho:  difference in coefficients not systematic

               chi2(25) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                        =      266.40
               Prob>chi2 =      0.0000
               (V_b-V_B is not positive definite)
```

Authors note: “H0: Difference in coefficients not systematic” is rejected which points to differences in coefficients being systemic i.e. there being a fixed effect.

Modified Wald test:

```
. xttest3

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (12)  =      64.43
Prob>chi2 =      0.0000
```

Authors note: H0 points to the existence of homoscedasticity. Due to its rejection, heteroscedasticity is shown to be present.

Wooldridge test:

```
Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation
    F(  1,      11) =      3.992
           Prob > F =      0.0710
```

Authors note: H0 states there to be no autocorrelation. We cannot reject HO, therefore no autocorrelation is present.

Ramsey RESET-test:

```
. xtreg lnattendance fit_2 Prwin Prwin2 ATP prev_season_home prev_season_away geographical_riv
> ear, fe robust

Fixed-effects (within) regression               Number of obs      =       396
Group variable: Hometeam                        Number of groups   =        12

R-sq:                                           Obs per group:
     within  = 0.5614                                          min =        33
     between = 0.9649                                          avg =      33.0
     overall = 0.4771                                          max =        33

                                                F(11,11)           =         .
corr(u_i, Xb)  = -0.9367                        Prob > F           =         .

                             (Std. Err. adjusted for 12 clusters in Hometeam)
```

lnattendance	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
fit_2	-.0588883	.1167822	-0.50	0.624	-.3159241	.1981475
Prwin	3.578891	3.968387	0.90	0.386	-5.15547	12.31325
Prwin2	-3.27336	3.767673	-0.87	0.404	-11.56595	5.019232
ATP	-.0036443	.0044189	-0.82	0.427	-.0133701	.0060816
prev_season_home	.0193058	.0228036	0.85	0.415	-.0308846	.0694961
prev_season_away	.0056078	.005922	0.95	0.364	-.0074264	.018642
geographical_rival	.6961107	.7516295	0.93	0.374	-.9582146	2.350436
Awayteam						
BK Häcken	-.886932	.8908234	-1.00	0.341	-2.847621	1.073757
Djurgårdens IF	-.0584217	.0763629	-0.77	0.460	-.2264953	.109652
Falkenbergs FF	-.7996916	.7761727	-1.03	0.325	-2.508036	.908653
Gefle IF	-.9761743	.9831268	-0.99	0.342	-3.140022	1.187673
Helsingborgs IF	-.578724	.5506184	-1.05	0.316	-1.790627	.6331789
IF Elfsborg	-.2720475	.2314043	-1.18	0.265	-.781365	.23727
IFK Göteborg	.1612462	.2220411	0.73	0.483	-.327463	.6499554
IFK Norrköping	-.5144845	.5075745	-1.01	0.333	-1.631648	.6026794
Kalmar FF	-.7940086	.7890529	-1.01	0.336	-2.530702	.9426852
Malmö FF	.2068181	.2315264	0.89	0.391	-.302768	.7164043
Örebro SK	-.840224	.8444887	-0.99	0.341	-2.698931	1.018483
dow						
1	-.0423587	.0546442	-0.78	0.455	-.1626297	.0779123
2	-.1431897	.0794668	-1.80	0.099	-.3180949	.0317154
3	.0157458	.0403295	0.39	0.704	-.0730188	.1045104
4	-.2449652	.2769405	-0.88	0.395	-.8545072	.3645768
5	.1111905	.1307835	0.85	0.413	-.1766619	.399043
6	.0590556	.0548635	1.08	0.305	-.061698	.1798093
year						
2015	.2902417	.3073887	0.94	0.365	-.3863162	.9667996
2016	.0504603	.0824313	0.61	0.553	-.1309698	.2318905
_cons	13.29328	8.987381	1.48	0.167	-6.487816	33.07437
sigma_u	1.2804683					
sigma_e	.21897741					
rho	.97158538	(fraction of variance due to u_i)				

.

Authors note: Of importance is the variable “fit_2” which is the squared fitted values. Due to its insignificance, misspecification can be ruled out.