The relationship between live attendance and uncertainty of outcome in European football

Erasmus University Rotterdam

Erasmus School of Economics

Master thesis

Supervisors: Dr. G. D. Granic & Prof. dr. H. Bleichrodt

The aim of this master thesis is to investigate the relationship between live attendance at sports events and the uncertainty of match outcome. The two main theories that are discussed are the uncertainty of outcome hypothesis (UOH) and the reference-dependent preferences with loss aversion. The UOH states that people who visit sports events prefer to see tight matches. The reference-dependent preference with loss aversion model states that people who visit sports events prefer to see a match with a more certain outcome. To see which theory applies to European football an empirical research using a Tobit model is done with data of the UEFA Champions League. As the UEFA Champions League consists out of a group stage and a knock-out stage, two separate analysis are done. The results show that there is no evidence for either theory for supporters visiting matches in the group stage. The results show that supporters visiting knock-out stage matches prefer to see tight matches and this evidence thus supports the UOH.

Name: Jurgen Vos

Exam number: 386446

Rotterdam, 28/8/2018

Acknowledgements

I want to thank my master thesis supervisor prof. dr. Han Bleichrodt for providing me with helpful feedback in the challenging process of writing this master thesis. Prof. Bleichtrodt steered me into the right direction throughout the thesis and was always helpful when I was struggling with something.

I also want to thank my master thesis supervisor dr. Georg Granic who was able to take over the supervision on short notice and provided me with helpful feedback in the last stage of the writing process. The clear feedback and new insights helped me finalizing the master thesis.

Table of contents

Introduction

Recently, the rumours regarding the start of a European Super Football League are increasing. This European Super League would contain the best football teams of Europe like Real Madrid CF, Juventus, Paris Saint Germain and many more teams (Platt, sd). The European Super League would be a competition comparable to the NBA and NFL, the National Basketball Association and the National Football League in the United States, where week in week out the best teams are competing to ultimately be crowned as champion at the end of the season. One of the reasons for the top clubs of Europe to create this new European Super League is the significant financial rewards (Gouka, 2017). The top clubs involved would be able to benefit from huge television rights deals. The television right deal of the NFL, for example, currently makes $4.95 billion a year of the United States rights (Platt, sd). If one compares this to the $1.57 billion a year of the global UEFA Champions League rights, one might see possibilities to increase this as football has a far greater appeal all around the world than American football has. The global TV audience for the UEFA Champions League final in 2016 was estimated around 350 million viewers, compared to the 150 million viewers of the Super Bowl in 2016 (Reekie & Grez, 2016).

Former manager of Arsenal FC, Arsène Wenger, who has managed one of Europe's biggest clubs for twenty-two years, is convinced that with the introduction of the European Super League a drastic change within the football world is about to happen and that within a few years we will be able to witness the European Football League arise (Hytner, 2018). According to Wenger, the contrast between the big clubs and the smaller clubs is too big. He argues that people aren't interested anymore in watching big teams play smaller teams, but that people are interested when the big teams are playing each other. These games are usually thrillers, as the teams are high quality. As in these games there is not a clear superior team, the match outcome is uncertain. An example of this is that Bayern Munich and Real Madrid faced each other twenty-six times in the last 40 years, from which Bayern won 11 times, drew 3 times and lost 12 times[1]. From these results one can see that these two big teams are equal.

[1] Real Madrid's record against Bayern Munich https://www.worldfootball.net/teams/real-madrid/bayern-muenchen/11/

Thus, if what Wenger argues is true, people would prefer matches with an uncertain outcome over matches with a more certain outcome. This would be in line with the uncertainty of outcome hypothesis which argues that people prefer tight matches over matches with a superior team facing an inferior team. One can speak of matches with a high uncertainty of outcome when the probability of winning the match is around the same value as the probability of losing the match. The games between Bayern Munich and Real Madrid are a prime example of a tight match as over the past decades they approximately won the same amount of games against each other. The first time the uncertainty of outcome hypothesis (UOH) was mentioned was by Rottenberg (1956) who stated that a team with a 55 percent probability of winning a match will attract more supporters to the stadium than a team with an 80 percent probability of winning a match. For over half a century, the uncertainty of outcome hypothesis has been accepted as a correct description of the outcome of consumer choices, without theoretic basis.

However, recent studies found evidence that wasn't in line with the uncertainty of outcome hypothesis. The results found by Buraimo and Simmons (2008) showed that an increase in uncertainty of outcome is associated with reduced gate attendance in the two biggest football leagues of Europe, the English Premier League and the Spanish La Liga, which is the opposite of the UOH. The theory behind this is discussed by Coates et al. (2014) who argued that people have reference-dependent preferences with loss aversion. According to them, fans simply want to see the home team win. The higher the probability of winning the home game, the higher the attendance. The reference-dependent preference with loss aversion theory by Coates et al. also explains why fans visit games with a very low probability of winning the game, as the expected utility fans derive from the thrill of an upset exceeds the expected utility fans derive from a loss.

As the theories are complete opposites of each other, and recent studies have found more evidence supporting the theory of reference-dependent preferences with loss aversion compared to the uncertainty of outcome hypothesis, an empirical research will be performed to see if the behaviour of people are following any of these theories and if so, which one. The dataset used for the empirical research is from the UEFA Champions League over the period 2011 till 2017 containing six seasons. The UEFA Champions League is arguably one of the most important leagues in football and the results might be useful to see whether a European Super

League can be successful, as such a league would contain a lot of matches with a high uncertainty of outcome. The research question is as follows:

Does an increase in uncertainty of outcome increase or decrease the match attendance for clubs playing in the UEFA Champions League?

The evidence from the empirical research using data of the UEFA Champions League knock-out stage supports the uncertainty of outcome hypothesis. The attendance in stadiums is significantly higher when the uncertainty of outcome increases, ceteris paribus. The results from the group stage data is insignificant and thus doesn't support any theory. As the knock-out stage is dominated by the big clubs, and the tight matches attract more supporters, a European Super League might be successful with the number of matches with high uncertainty of outcome it has to offer.

First, the uncertainty of outcome hypothesis and the reference-dependent preferences model with loss aversion will be explained. Secondly, the literature about the relationship between uncertainty of outcome and attendance at football games are discussed. After that an empirical research is presented using data of the UEFA Champions League over the seasons 2011-12 till 2016-17 and the results are discussed afterwards. Finally, a conclusion is formed containing limitations of this research.

Theory

A Model of Attendance under Uncertainty

The model of reference-dependent preferences (Coates et al., 2014) states that the outcome of a sporting event can be represented by a binary variable *y*. If the home team wins the game $y = 1$ and if the home team loses the game $y = 0$. From attending a game, individuals derive two types of utility, which are the consumption utility and the gain-loss utility. The first type of utility, the consumption utility, corresponds with the standard utility of consumer theory. The second type of utility, the gain-loss utility, is the utility that depends on what happens on the field compared to the consumer's reference point (Koszegi and Rabin, 2006). The individual's reference point, in this case, will be the expectation he has about the outcome of the sports event. This will be noted as $E(y = 1) = p^r$. If the individual attends a game where the home team wins $(y = 1)$ he will get consumption utility, U^W, as well as gain-loss utility

from experiencing the win conditional on the reference point p^r. The marginal impact of a positive deviation from the reference point is $\alpha > 0$. The total utility of attending a sports event where the home team wins is $U^W + \alpha(1 - p^r)$.

On the other hand, it is also possible that the home team loses the game $(y = 0)$. In this case, the individual also gets consumption utility and gain-loss utility. The consumption utility then will be U^L and the gain-loss utility will be determined by the difference between the sensation of a loss compared to the reference point. The negative deviation from the reference point is $\beta > 0$. The total utility of attending a game where the home team loses is $U^L + \beta(0 - p^r)$. The following figure (Figure 1) clearly shows what the utility looks like taking game outcomes and reference points into account.

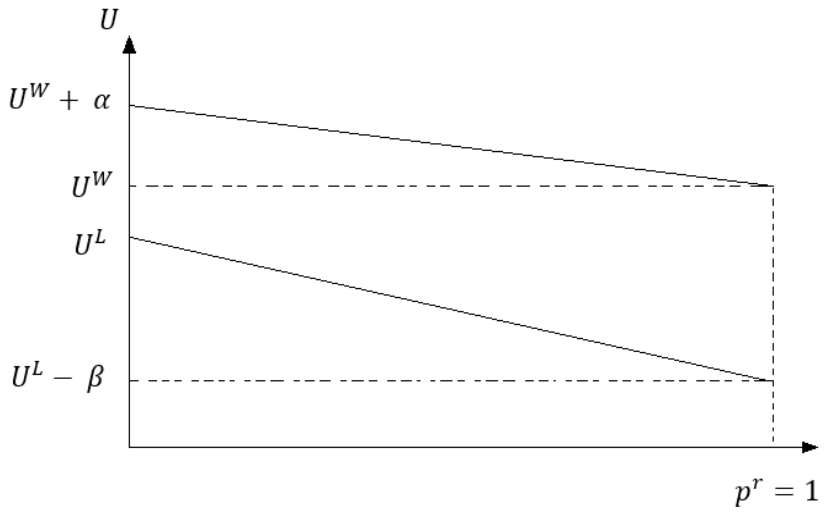

Figure 1

The maximum total utility an individual can get is when the individual does not expect a win for the home team at all, so when his reference point is equal to 0, but the home team does unexpectedly win. His total utility then will be equal to $U^W + \alpha$. If the individual went to a game of the home team and fully expected them to win and they did, his total utility will be equal to U^W. The total utility derived by a home loss will be lowest when the individual went to a game of the home team and was very certain the home team would win the game $(p^r = 1)$, but they lost the game. The total utility, in that case, will be equal to $U^W - \beta$. If the individual already expected the home team not to win the game $(p^r = 0)$ and they indeed did not, the total utility would be equal to U^L.

As stated before, the outcome of sports events is uncertain. Coates et al. assume that consumers who are trying to decide whether to attend a game form a reference point that is equal to the probability that the home team will win the game. Under this assumption the expected utility from attending a game is as follows:

1) $E(y) = p * 1 + (1 - p) * 0 = p^r$
2) $E(U) = p[U^W + \alpha(1 - p)] + (1 - p)[U^L + \beta(0 - p)]$
3) $E(U) = (pU^W + p\alpha - p^2\alpha) + (U^L - p\beta) - (pU^L - p^2\beta)$
4) $E(U) = p^2(\beta - \alpha) + p[(U^W - U^L) - (\beta - \alpha)] + U^L$

Equation (1) describes the expected outcome of a game. If one replaces the utility functions for the win and the loss, which are the one and the zero in equation (1), the equation (2) will follow, which is a utility function. If one writes this equation down and isolates the p's, equation (3) and ultimately equation (4) follow. Equation (4) shows that the expected utility from attending a game is a quadratic function of the probability of the home team winning the game. This utility function takes the game outcome uncertainty into account and enables generating utility by watching the home team win very unexpectedly or generating disutility from an unexpected loss by the home team.

With the expected utility from attending a game in mind, individuals now must decide whether they are going to attend a game. So, the choice to attend a game is binary, you go to the stadium or you do not. If an individual decides to not attend a game, he gets utility *v*. This utility level can be interpreted as the reservation utility of not attending a game. The value of *v* is uniformly distributed across the population and will be between $[\underline{v}, \overline{v}]$. In case that the expected utility of the individual is bigger than the reservation utility he will attend the game. In case that it's lower than *v* he will decide not to attend the game. People with a low reservation utility, which means a low *v*, are fans of the home team. In many cases, the expected utility will exceed the reservation utility and therefore they will decide to attend the game. The other consumers, the ones with a higher reservation utility will be pickier in deciding whether to attend the game as it is less likely that the expected utility exceeds the reservation utility.

If there was no reference-dependent utility the equation would look as follows:

5) $E(U) = (U^W - U^L)p + U^L$

This function above is an increasing function of the probability that the home team wins the game. If the probability of winning a game is equal to zero, the individual will get an expected utility level equal to U^L. If the probability of the home team winning the game increases the expected utility will also increase no matter what because Coates et al. assume that the utility of winning a game is bigger than the utility of losing a game. This would mean that if the home team is expected to win the game it will have higher attendance than when the home team is not expected to win the game. This would also imply that teams that are better would have higher attendance than teams that perform worse. Another case implied by the model of attendance under uncertainty is that of a fan of the sport itself. This individual does not support the home team, so the standard consumption utility, in this case, would be equal for seeing a win or loss for the home team which means $U^W = U^L$ and the gain-loss utility is not important $(\beta - \alpha) = 0$. The utility derived from attending a game will simply be $U^W = U^L$ for this fan.

Uncertainty of Outcome Hypothesis

In the context of the model of attendance under uncertainty, the Uncertainty of Outcome Hypothesis (UOH) will be motivated. The UOH is a concave relationship between the gate revenue and the home win probability with a maximum achieved between (0.5, 1). Rottenberg (1956) states that a team that wins around the 80 percent of its games will attract less attendance than a team that wins 55 percent of its games. This is considered the classic UOH (see figure 2).

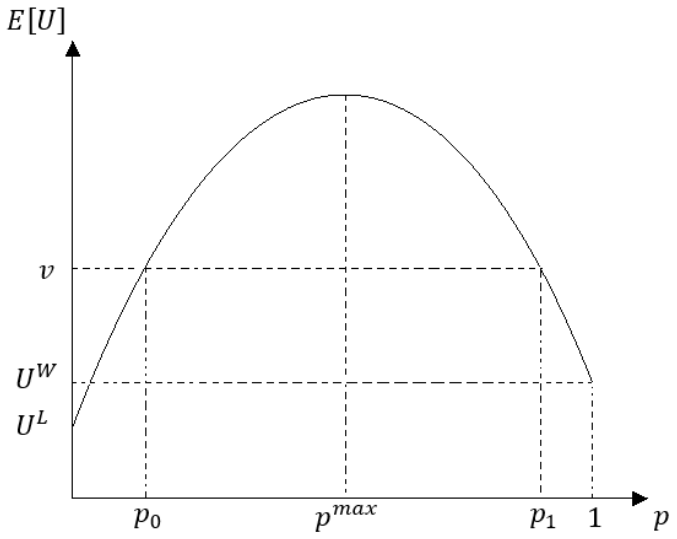

Figure 2 - UOH

Figure 2 is showing the consumer decision making consistent with the uncertainty of outcome hypothesis. The expected utility of the consumer is concave in p and the maximum expected utility is at p^{max}. The consumer will decide to go the game if the expected utility of attending the game is bigger than the reservation utility which is equal to v, just like in the model under uncertainty described before. In the case of the uncertainty of outcome hypothesis this will be between the probability of winning the game p^0 and p^1. The outcomes of these games are relatively more uncertain. The classic UOH requires that $U^W - U^L < \alpha - \beta$. This indicates that the fans' preferences for home wins are dominated by their preferences for tighter games.

Reference-Dependent Preferences and Loss Aversion

The UOH is not the only relationship between the probability that the home team will win the game and the expected utility consistent with the model. If α and β are both positive, and $\beta > \alpha$, the marginal utility from game outcomes that deviate from the reference point when the home team is expected to lose is larger than the marginal utility from game outcomes when the home is expected to win in absolute terms. This is called loss aversion where the marginal utility of a loss is bigger than the marginal utility of a gain. The UOH is not consistent with the presence of loss aversion. This is because the expected utility function is not concave when $\beta > \alpha$.

There are some differences in consumer decisions under loss aversion compared to the UOH. In figure 3 the expected utility function under the assumption of loss aversion and game attendance decisions made by the fans is shown. The shape of the utility function can be explained as follows. As mentioned before in the model of attendance under uncertainty the utility consists of two parts, the consumption utility, and the gain-loss utility. The consumption utility, $pU^W + (1-p)U^L$ increases with p. This means that the higher p, the higher the probability the home wins the game. The gain-loss utility, $(\alpha - \beta)p(1-p)$ first decreases with p at a decreasing rate until $p = 1/2$, and then it increases with p at an increasing rate. If p is smaller than $\frac{U^W - U^L}{2(\beta-\alpha)}$ the negative impact of the gain-loss utility dominates, if p is bigger than $\frac{U^W - U^L}{2(\beta-\alpha)}$ the positive impact of the consumption utility dominates. Another important factor is the reservation utility v. In general, the value of v is low enough that there is, like in figure 3, a range of declining attendance as the probability of a win of the home team increases. If this is the case in practice, data should be able to identify loss aversion. If the

value of v is sufficiently large and exceeds the value of U^L the attendance will rise only when p is relatively large so that the expected utility from attending a game exceeds the reservation utility. If this is the case, game attendance might not react to changes in at least a range of the expected home probability and thus have a flat section, because v exceeds U^L for a longer amount of time.

In figure 3 it is shown that under reference-dependent preferences and loss aversion, the game attendance at games with a relatively certain outcome generates higher expected utility than games with a relatively uncertain outcome. This is due to the loss aversion which makes fans less interested in games where the marginal utility of seeing an unexpected loss exceeds the marginal utility of seeing an unexpected win. The model of reference-dependent preferences and loss aversion is also able to explain the fans interest in seeing upsets. An upset takes place when a team is expected to lose the game, with other words the probability of the home team winning the game is small, but they win the game. This outcome generates a relatively large gain-loss utility. The fans are keeping this in mind when they visit games with a very small p and this explains the convexity of the expected utility function under loss aversion. The UOH on the other hand is not able to explain the fans' interest in seeing upsets.

It is important to know that the relationship between game outcomes and game attendance requires reference-dependent preferences and loss aversion, and not risk aversion. A fan who would be risk-averse would have a standard concave and increasing expected utility function and always would get more utility from a game with a higher p. A consumer with loss aversion can get more expected utility from a game with p = 0.1 compared to a game with p = 0.2.

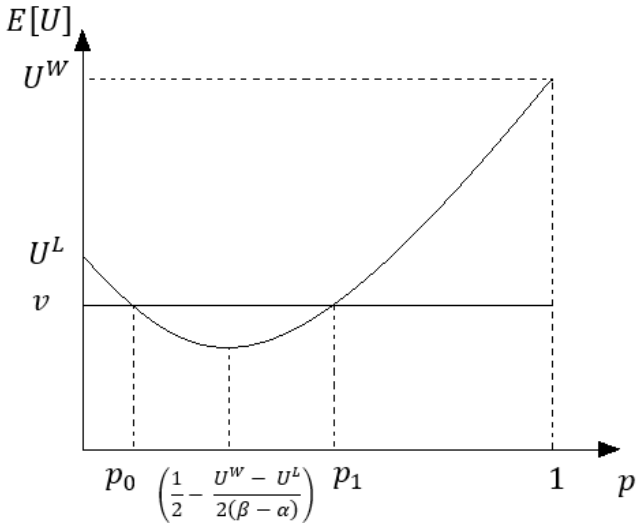

Figure 3 - Loss Aversion

Literature Review

Before the empirical research will be performed to analyse the relationship between the uncertainty of game outcome and game attendance, the literature on this topic will be reviewed. As in this research, the dataset that will be used is about the UEFA Champions League, the literature that will be reviewed will be about football as well.

English Premier League

One of the most famous and most appreciated football leagues in the world is the English Premier League. Buraimo and Simmons (2008) investigated whether football fans of the clubs in the highest division of England really valued the outcome of uncertainty. The reasons for Buraimo and Simmons to focus on the English Premier League (EPL) were that the EPL is a global brand with games viewed by people all around the world, and the EPL consists of a pool of players of a considerable high quality, which proves the quality of the league. Based on revenues and fan interest, the EPL is together with Spain's La Liga, the world's most interesting national football league. Due to the existence of capacity constraints in most EPL games, Buraimo and Simmons (2008) chose to perform a Tobit estimation to be able to censor the sold-out games. The measure they use to calculate the outcome uncertainty is the absolute difference between home and away win probabilities. As there are three different probabilities, this measure assumes that the probability of a draw is constant. According to them, it is a reasonable assumption as in their dataset the probability of drawing is 0.27 with a standard deviation of 0.03. They do deal with the slight variation in the draw probability by adopting a measure of outcome uncertainty, the Theil measure. The Theil measure explicitly takes variations of all three probabilities into account. An increase in the Theil measure is associated with increased outcome uncertainty.

The findings of their research are that the variables *probability of home win* and *probability of home win squared* are negative on the level term and positive on the quadratic term. This indicates that there is a U-shaped relationship between the probability of winning the home game and the attendance level, which is in line with the loss aversion theory. The turning point of this quadratic relationship is at a probability of 0.35 in their sample. Also, the Theil measure has a negative coefficient, which means that fans do not prefer close contests. What Buraimo and Simmons thus found is that an increase in uncertainty of outcome is associated with a

lower attendance rate. This could be due to the fact the supporters are almost all supporters of the home team who rather see their team win than lose. They also found that fans rather pay to watch a game with a very high probability of winning or a very low probability of winning. The attraction of a low probability win could be compared to "David versus Goliath" effect.

La Liga

Buraimo and Simmons (2009) also looked at the relationship between game attendance and the uncertainty of outcome in the Spanish La Liga. In their study, they focused on the supporters who visit the games and the supporters who watched the game on television. The findings of the behaviour of the supporters who watched the game in the stadium will be discussed here. Buraimo and Simmons stated that stadium fans tend to be loyal supporters compared to supporters who watch the match on television. Lots of supporters have a season ticket which enables them to visit every home game that season. Visiting every home game in a season is a sign of loyalty towards the club.

Buraimo and Simmons had a few reasons to choose La Liga as competition for the empirical research. The first reason is that the gate attendances in Spain are rarely constrained by capacity. For the research, this means they don't have to censor the data due to sold-out games. In the research of the English Premier League, they used the Tobit regression, as the stadiums in England are constrained by capacity. In this research, they chose the Prais-Winsten panel regression in which error terms are contemporaneously correlated across panels, in this case, the home teams. This is important because there is likely to be further habit persistence among fans in their attendance at successive home matches. The second reason to pick La Liga for the empirical study is that the league is dominated by two teams, FC Barcelona and Real Madrid CF. Due to their domination, there'll be a substantial number of games where a large-market team plays against a small-market team, and thus the outcome of uncertainty would be low. The substantial variation in outcome uncertainty in La Liga allows them to test for responses of the two types of audience. They used data of the seasons 2003/04 till 2006/07 in La Liga. To indicate whether a match is close they decided to use the absolute difference between home and away win probability. However, this assumes that the probability of drawing is equal. Therefore, they introduce the Theil measure, which considers

that all three probabilities may vary. In case the audience values outcome uncertainty the coefficient of the Theil measure will be positive.

Buraimo and Simmons found that the *absolute probability difference* has a significant, and positive coefficient. This means that if the difference between home and away win probability increases the attendance of the home team also increases. The coefficient of the Theil measure is negative and significant. This means the audience does not value outcome uncertainty. The results are contrary to the outcome of uncertainty hypothesis and in line with the loss aversion theory.

German Bundesliga

Pawlowski and Anders (2012) looked at the importance of uncertainty of outcome in the highest division of German football, the Bundesliga. As mentioned in their research that Buraimo and Simmons saw that fans in the EPL prefer to see their team play a much inferior team, Pawlowski and Anders wondered why matches against the high-quality 'big' clubs like Manchester United in the EPL and Bayern Munich in the Bundesliga would still sell out. As Czarnitzki and Stadtmann (2002) argued that fans rather care about the reputation of clubs than the thrill of uncertainty of outcome, the reputation of a club can be calculated in different ways and interpreted differently by people. The aim of the research of Pawlowski and Anders is first to see whether spectators in the Bundesliga prefer the home team to win a game against an inferior team more than a tight match. Secondly, they look at the away teams' brand strength. Thirdly, they focus on the perception of fans on the European clubs' competitions, where they control for both the uncertainty of championship outcome and probability of qualifying for the UEFA Champions League. The dataset they use is over the season 2015/16. In their model, they control for variables like the cost of a ticket, past performance, matchday in the season, and whether it rained during the game. To take the short-term uncertainty of outcome into account they introduce the Theil measure to their model. In this measure, the home team's winning probability, the away team's winning probability, and the draw probability are based on unbiased odds. In case the coefficient of the Theil measure is positive, the people visiting the game will prefer tight games. If the coefficient is negative, the people visiting the game prefer to see the home team play an inferior team. Pawlowski and Anders also control for the possibility that fans might rather prefer the home team to win by introducing the dummy variable *Favourite*. This dummy

variable measures if the home team's probability to win is greater than the away team's probability to win. The last variable they introduce to their model is the variable *Brand* which considers that fans might evaluate a certain place in the standing differently for a financially strong team differently compared to a not financially strong team and that teams who performed badly might still attract spectators due to a strong brand image. Pawlowski and Anders decided to use a generalized Tobit estimator since over 20% of the matches were sold-out

Their results showed that if the home team or the away team still has the chance to become the champion of the Bundesliga the attendance would be significantly higher. Furthermore, they found that the significant negative effect of the Theil measure indicates that fans do not prefer matches that are predicted to be close in score. This suggests that the less balanced a match is, the higher the attendance would be. However, this cannot be explained by the home team playing a much more inferior team as the variable *Favourite* is only weakly significant and negative. The variable *Brand*, the brand strength of the away team, on the other hand, has a significant positive effect on attendance. In general, the negative coefficient of the Theil measure which means that people aren't interested in tight games is in line with the findings of Buraimo and Simmons. However, it is rather a strong brand that increases attendance rates than spectators preferring the home team to play a much inferior team.

Another empirical research about the relationship between the uncertainty of outcome and match attendance with use of data about the Bundesliga is the research of Benz et al. (2007). The purpose of their study is to determine the effect of match uncertainty of outcome variables on match attendance. A big difference with other literature about this topic is the regression analysis method chosen by Benz et al.. They introduced quantile regression analysis to the demand for sport, as they believe that the quantile regression analysis can provide a better view on the conditional distribution of match attendance figures. The reason for this introduction of the quantile regression analysis is that they were concerned that an exclusive focus on average effects might misguide policymakers in the presence of heterogeneity in consumer demand. By heterogeneity, they talk about the variation across different quantiles concerning the group of relevant factors for match attendance. As example, they talk about the possibility that a consumer's utility from celebrating the home team's victory is increasing in the number of spectators. If this would be the case, the probability of winning of the home

team would have a greater influence on attendance demand for higher quantiles. Their data contains information on more than 1,200 matches in the seasons 1999/2000 until 2003/04.

Benz et al. found two important facts. Firstly, they found that uncertainty of outcome is only a second-order influence factor for attendance demand. The factor that appeared to be much more important is a team's reputation, based on its performances of the previous years and current ranking. This is in line with the findings of Czarnitzki and Stadtmann (2002). Secondly, they investigated whether uncertainty of outcome is relevant for all quantiles in the distribution of consumer demand. They found that match uncertainty of outcome is consistently to affect high-demand matches only. If one would try to increase the match attendance, redistributing the television-earnings might in some cases work. When one would decide to redistribute the television-earnings weak teams are able to spend more money on player talent. Being able to spend more doesn't necessarily mean that the team gets better significantly. For a team to get stronger depends on the efficiency of their talents investments. If such teams would succeed in improving their quality, it would also be in favour of the strong teams. As the game between the "high-demand" team and the improved team will get a higher outcome of match uncertainty.

Authors	**Football league**	**Uncertainty measure**	**Model**	**Results**	**Support**
Buraimo and Simmons (2008)	English Premier League	Betting odds	Tobit regression	Prob. of win <0 Prob. of win^2 >0 Theil <0	Loss aversion
Buraimo and Simmons (2009)	Spanish La Liga	Betting odds	Prais-Winsten panel regression	Absolute prob. difference >0 Theil <0	Loss aversion
Pawlowski and Anders (2012)	German Bundesliga	Betting odds	Tobit regression	Theil <0 *Brand* >0	Loss aversion (cannot be explained by home team preference)
Benz et al. (2009)	German Bundesliga	Betting odds	Quantile regression	Prob. of win >0 Prob. of win^2 <0	Uncertainty of Outcome Hypothesis (UOH)

Data

As the aim of this research is to test whether loss aversion, described in the theory section, can be applied in football as well, in this research data will be used from the UEFA Champions League over the season 2011-2012 until 2016-17. First, there will be a brief explanation on what the UEFA Champions League is and how it works. Secondly, the data collection will be discussed, and descriptive statistics of the data will be presented.

Champions League – tournament design

The UEFA Champions League is a football tournament in Europe. The champions of various competitions in Europe can compete in this tournament, as well as clubs who have finished in a spot in their competition which granted them a place in the UEFA Champions League. The number of places per country is based on its UEFA ranking for club competitions. The countries' clubs' coefficients rankings are based on the results of each association's clubs in the previous five seasons of the UEFA Champions League and UEFA Europa League. The ranking of the country determines the number of places allocated to this country. This means that if clubs from a certain country are performing very well over the course of the last five years, this country will have a high number of places allocated. The three highest noted countries get four tickets to enter the UEFA Champions League. It is possible for a country to end up with five clubs in the UEFA Champions League. This can happen if a club that has not ended in the top four of this countries' competition managed to win the UEFA Europa League. If this is not the case, the maximum number of clubs of a country in the UEFA Champions League is four. The places to enter the UEFA Champions League aren't all the same. Some places will directly enable clubs to participate in the group stage of the tournament while other places only enable the club to play playoffs. In this study the playoff games are not considered. One of the reasons not to do this is that these playoff games are held pre-season. During pre-season, there are multiple factors that might influence the behaviour of people going to a football game. For instance, lots of people go on holiday during the summer holiday which could lead to lower attendance rates, while usually they would go support the home team.

After the play-offs there are 32 clubs left. These 32 clubs will be divided over eight different groups. To ensure the differences between the eight groups are not too big they make four pools of eight based on their ranking. The number one in a group will be randomly drawn from pool one, number two of the group from pool two and so on for group 3 and group 4. Clubs from the same country cannot be drawn into the same group. After 6 rounds the teams all played each other at their home ground. The eight group winners and the eight runners-up of the group stage qualify for the round of 16. The clubs that finish third in the group will move into the round of 32 of the UEFA Europa League. The clubs that finish last in their groups will be eliminated from the European tournament.

After the group stage, the last 16 clubs will participate in a knock-out stage. The knock-out stage consists out of four rounds. In these rounds, there will be drawn pairs who both play a game at their home ground. The rounds are last 16, quarterfinal, semi-final, and final. The club that will proceed to the next round is the club that has the best result on aggregate. In case a team draws, the away team's goals have a higher weight compared to the home team's goals. For example, the team that plays at home first and draws with 2-2 will have to win or at least draw the away game with 3 goals to 3 to progress to the next round. In case it would end up in a 2-2 after 90 minutes, they will have to go to overtime which lasts for 30 minutes and if it's still equal after that, a penalty shoot-out will decide who'll proceed to the next round. In this research the final will be neglected as the final of the UEFA Champions League is played on neutral ground. This ground is set before the tournament starts.

Champions League – data collection

The data gathered for this study are about the seasons from 2011-12 until 2016-17, which are six seasons in total. The attendance data from each game are collected from the website www.worldfootball.net. The data about the capacity of the stadiums are collected from this website as well. As the UEFA have strict safety regulations some stadiums cannot make use of their full capacity. For example, certain standing ranks are not allowed in UEFA Champions League. Football stadiums like the Westfalen stadium, Allianz Arena and Veltins Arena therefore have adjusted capacities. The attendance level is the attendance level in total, no difference is made between home and away supporters, nor has there been considered if people needed a season ticket or if the tickets were bought on general sale. The betting odds were collected from the website www.oddsportal.com. The odds on the website are decimal

odds. The number of goals scored per game by each home team, the number of goals scored per game by each away team, the date of the games and in which round the games were played are also gathered via www.oddsportal.com. The ranks of the home teams and away teams per game are collected via the official website of the UEFA www.uefa.com. The data collected of the UEFA Champions League is split up in the Group Stage section and the Knock-out Stage section. In the following table, the descriptive statistics for the group stage can be found.

Dependent variable: Attendance/Capacity			
Variable	Minimum	Maximum	Mean
Attendance	0	96,290	39,242.47
Capacity	12,500	99,534	49,867.04
Attendance/Capacity	0	1.072146	0.7916756
Probability of winning home team	0.1359252	0.938388	0.5999818
Probability of winning home team-squared	0.0184757	0.880572	0.3896746
Ranking home team	1	179	32.30729
Ranking away team	1	179	32.30729
Goals scored home team	0	16	3.479167
Goals conceded home team	0	24	3.654514
Goals scored away team	0	19	3.649306
Goals conceded away team	0	22	3.465278
Observations	576	576	576

Table 1 - Descriptive Statistics Group Stages

In the seasons 2011-12 till 2016-17, the lowest attendance level in this period was equal to 0 visitors. The reason behind this is that the home team(s) with this attendance level was/were punished by the UEFA, because of the excessive behaviour of the home team supporters in a previously played game in any European tournament of the UEFA[2]. The teams that were punished in this period are CSKA Moscow, Legia Warsaw and Dynamo Kyiv. As the minimum capacity level is equal to zero, the minimum attendance level percentage is also equal to 0%. The maximum attendance level ratio for a game in the group stages in the UEFA Champions league over the period 2011-12 until 2016-17 is equal to 1.07. This means that some clubs sold more tickets than the actual capacity of their stadiums. The average attendance percentage is

[2] UEFA Champions League or UEFA Europa League

equal to 0.79 for a game in the group stage of the UEFA Champions League during the period from 2011-12 till 2016-2017. In total there are 71 different teams that participate in the UEFA Champions League group stage from the season 2011-12 until 2016-17, representing 27 different countries. The minimum amount of goals scored and conceded for the home team as well as the away team is equal to 0. This is because for this research the goals scored and conceded by the home team and away team are looked at prior the match. At the beginning of every season the goals start at 0. The maximum amount of goals scored by the home team in one season prior to a group stage game is 16 goals. The maximum amount of goals conceded by the home team in one season prior to a group stage match is 24 goals. The maximum amount of goals scored by the away team in one season prior to a group stage match is 19 goals. The maximum amount of goals conceded by the away team in one season prior to a group stage match is 22 goals. In total there are 576 group stage matches played.

In the following table, the descriptive statistics of the knock-out stage are presented.

Dependent variable: Attendance/Capacity			
Variable	Minimum	Maximum	Mean
Attendance	15,000	96,636	55,172.36
Capacity	18,524	99,534	59,633.52
Attendance/Capacity	0.470116	1.001891	0.9262175
Probability of winning home team	0.2825114	0.8881936	0.6047337
Probability of winning home team-squared	0.0798127	0.7888879	0.3830409
Ranking home team	1	84	13.86905
Ranking away team	1	84	13.86905
Goals scored home team	4	39	14.70238
Goals conceded home team	1	16	6.410714
Goals scored away team	4	39	14.70238
Goals conceded away team	1	16	6.410714
Observations	168	168	168

Table 2 - Descriptive Statistics Knock-Out Stages

The minimum attendance level of all the games in the knock-out stage of the UEFA Champions League during 2011/12 till 2016/17 is 15,000 people. This means that in contrary to the group stage there were no clubs penalized for any excessive behaviour of their fans and thus was every game played in front of home supporters. In comparison with the group stage, both the mean of the attendance level and the mean of the capacity level are noticeably higher. The reason behind this is that the higher ranked clubs in Europe usually proceed to the knock-out stage, and these clubs have generally bigger stadiums. The maximum attendance/capacity exceeds 1 in the knock-out phase, which means that there were more people in the stadium than the capacity of the stadium allows. There were 38 different teams participating in the knock-out phase of the UEFA Champions League over the seasons 2011-12 till 2016-17 and they were representing 15 different countries. The minimum amount of goals scored in one season by a team that played in at least one round of the UCL knock-out stage is 4. This includes the goals scored in the group stage of the season. The maximum amount of goals scored in one season in the UCL by a team that participated in at least one round of the knock-out stage of the UEFA Champions League is 39. The minimum amount of goals conceded by a team in one season of the UCL that played at least one match in the knock-out stage is 1. The maximum amount of goals conceded by a team in one season in the UCL that at least played one match in the knock-out stage is 16. In total there have been played 168 knock-out stage matches over the seasons 2011-12 till 2016-17.

Methodology

To be able to see whether loss aversion applies to football fans or that the football fans value uncertainty of outcome for the match, a few points must be discussed before analysing the data. First, some problems need to be tackled to make the dataset ready to be analysed. After the problems are tackled, the model that is going to be used will be discussed. Last, the econometric model will be presented.

Tied Games

One of the main differences between the dataset used in this thesis and the dataset of Coates et al. is that in football there exists a relatively big probability to tie games. Coates et al. used a dataset of the Major League Baseball, which is the highest baseball division in the United States. In this competition, a tied game is very rare. The higher probability for tied games

complicates the model because it adds another outcome to a game. The biggest question that arises is whether people interpret the tie as a good or a bad result. First, two scenarios will be described to see how a tie can be interpreted as a positive result as well as a bad result. Afterwards, the solution will be presented on how to decide whether fans will be happy or disappointed with a tied match.

The first scenario would be from the point-of-view of the home-team supporter whose team is the better team. When this team is playing against a team which is worse, the supporter would expect the team to win the game. If the home-team would end up tying this game, the utility derived from the tie would be close to the utility of losing. As in this case, the home team will only get one point instead of three points. This will feel like a loss of two points.

The second scenario would be from the point-of-view of the home-team supporter whose team is the underdog. If the probability of winning a game is low, a win would make the home-team supporters feel ecstatic, as it wouldn't be expected at all. Ending a game tied will almost have the same effect. Being able to get one point against the better team will feel like a euphoria. Therefore, the utility of tying the game will be close to the utility of winning a game. The point earned by the home-team will feel like a bonus point as the expectation was that this game no points would be earned.

As one can see, a tied game can feel like a win and like a loss. This depends on the fact if the away team is better or worse in the eyes of the supporters. In case the away team is better than the home team a tie will provide a victorious feeling. In case the away team is worse than the home team the tie will provide a losing feeling. To decide which team is favourite to win the game the UEFA coefficients ranking for clubs will be used. The UEFA coefficients ranking for clubs is based on the results of clubs competing in the UEFA Champions League and UEFA Europa League. The clubs' coefficients are determined on the sum of all points won in the previous five years of that season. This way the performances over the previous seasons are also considered to give a better view over the general performances in Europe by the clubs involved. In this research, the higher ranked club will be the favourite team to win the game.

Relative probabilities

After tackling the 'possibility to tie'-problem, the probabilities to win a match can be calculated. As stated in the Data section the odds collected are decimal odds. These decimal

odds need to be transformed into probabilities. In this case, one must be careful with transforming these decimal odds in probabilities as bookmakers do also charge a price for providing the possibility to offer a bet. Using their odds will not give a fair representation of the probabilities. Therefore, the price of providing the possibility to offer a bet must be taken out of the equation. To do this the method of Kuypers (2000) will be adopted. The first step to getting closer to the relative probabilities is that for each of the odds for winning, tying and losing the home game the probabilities will be calculated. This can be achieved by using the following formula:

$$6) \quad \frac{1}{odds\ win} + \frac{1}{odds\ tie} + \frac{1}{odds\ loss} = x$$

Using decimal odds of a bookmaker will result in value x, which will be a value that is a little over 1. However, the total sum of these probabilities should be equal to 1. Therefore, the difference between these two values is called the over-roundness of the book. By eliminating the over-roundness of the book, the relative probabilities will be calculated. The total sum of the relative probabilities for the win, tie, and loss is equal to 1. By using the following formula, the relative probabilities can be calculated for win, tie, and loss.

$$7) \quad \frac{1}{x(odds\ win)} + \frac{1}{x(odds\ tie)} + \frac{1}{x(odds\ loss)} = 1$$

To make the method described above clear an example will be presented using a game between Liverpool and FC Porto with decimal odds for a home win, draw and away win.

	Home win	**Draw**	**Away win**
Liverpool vs. FC Porto	1.44	4.28	6.39

Using formula (6) will result in a table with probabilities. These probabilities will consider the over-roundness of the book, which is the price one must pay to the bookmaker for being able to place a bet. The table with the probabilities will look as follows:

	Home win	**Draw**	**Away win**
Liverpool vs. FC Porto	0.69444	0.23364	0.15649

As one knows these probabilities are not correct yet, because when these three probabilities are being add up the total sum is equal to 1.08458. To get the relative probabilities formula (7) will be used to correct for the over-roundness of the book. Entering the numbers in the formula would give the following formula:

$$\frac{1}{1.08458(1.44)} + \frac{1}{1.08458(4.28)} + \frac{1}{1.08458(6.39)} = 1$$

The relative probabilities of the different game outcomes can be found in the following table:

	Home win	**Draw**	**Away win**
Liverpool vs. FC Porto	0.64029	0.21542	0.14429

As discussed above, the model is not able to deal with the ties. In case the home team is favourite to win the game, a tie will feel like a loss and therefore also will count as a loss. The probability of tying the game will be added on the probability of the away team winning the game. In case the home team is the underdog the tie will feel like a win and therefore the probability of tying a game will be added to the probability of the home team winning the game.

Regression model – Fixed-Effects Tobit model

The method of analysis chosen to research whether people have loss aversion or prefer tight matches is the fixed-effects Tobit model. The dataset contains values that should be censored as some games are sold out and there are certain games where attendance exceeds capacity. The Tobit model will be able to deal with this problem. Also, there are some clubs that got a punishment of the UEFA for excessive behaviour of the supporters which resulted in matches with 0 attendance. These games will also be censored. The fixed effects for the group stage dataset will be team-fixed effects, season fixed effects, month fixed effects and day fixed effects. The fixed effects for the knock-out stage are almost similar to the ones for the group stage. The only small difference is that the month fixed effects will be replaced by round fixed effects.

Econometric model

The structural econometric model that will be used to analyse the behaviour of the supporter or the football clubs in the UEFA Champions League over the period 2011 till 2017 is as follows:

8) $$AttendanceCapacity_{ijt} = \beta_0 + \theta p_{ijt} + \gamma p_{ijt}^2 + \beta_1 GoalsScoredHomeTeam_{ijt} + \beta_2 GoalsConcededHomeTeam_{ijt} + \beta_3 GoalsScoredAwayTeam_{ijt} + \beta_4 GoalsConcededAwayTeam_{ijt} + \beta_5 HomeRank_{ijt} + \beta_6 AwayRank_{ijt} + \beta_7 HomeTeam_{ijt} + \beta_8 AwayTeam_{ijt} + \beta_9 Season_{ijt} + \beta_{10} Month(Round)_{ijt} + \beta_{11} Day_{ijt} + \varepsilon_{ijt}$$

The dependent variable in this model is Attendance/Capacity which describes the fraction of attendance on capacity. The reason to take this variable as the dependent variable is that the aim of the research is to see whether people are more likely to visit tight matches or have loss aversion. The value of attendance/capacity is clear to see how many people were attending the game given the capacity of a stadium. As the capacity of stadiums differ between different football teams, the value of attendance/capacity allows to compare the values between the different football clubs.

The variables Goals scored home team, Goals conceded home team, Goals scored away team, Goals conceded away team, are independent variables that gives information about the amount of goals scored by the teams facing each other prior to the game played. This is the knowledge of the supporter who decides to go to the football match assuming that this person didn't buy the ticket a long time in advance. The goals scored and conceded are only about that particular season and all teams start at 0 goals scored and conceded prior to the first match in the group stage being played of that season.

The independent variables Home Rank and Away Rank describe the ranking of the teams participating in the UEFA Champions League. The ranking is based on the European results of the clubs over the previous five years prior to the current season. This variable is used to give an indication on the quality of the teams. Higher ranked teams will have better players than low ranked teams.

The variables Home team and Away team describe team related characteristics. The variables Season, Month and Day are describing time-characteristics. The possible days of playing Champions League games are Tuesday and Wednesday and the months that are relevant for this research are October, November and December. The variable Round is the replacement of the variable Month for the knock-out stage as it describes which round of the knock-out

stage the game takes place. The possible rounds are last 16, quarter finals and semi-finals. The key parameters of this model are the ones on the variables reflecting the probability of winning p.

The key parameters in this model, θ and γ, are functions of the parameters in the behavioural model we have described in chapter *Theory* (Coates et al, 2014). The parameter γ represents $\frac{(\beta-\alpha)}{(\overline{v}-\underline{v})}$ and shows if there are any reference-dependent preferences and loss aversion. The parameter θ represents $\frac{((U^W-U^L)-(\beta-\alpha))}{(\overline{v}-\underline{v})}$ and tells something about the relationship between the uncertainty of game outcome and the attendance level of the game. The values of these parameters in the model will be able to tell us something more about the relationship between expected game outcomes and attendance rates. The estimates of these parameters in the Random-Effects Tobit model will be used to see what the relationship is between the uncertainty of game outcome and the attendance rates in UEFA Champions League games over the period 2011-2017.

In case $\gamma > 0$ it implies that $\beta > \alpha$. As seen in the chapter Theory, if $\beta > \alpha$ the marginal consumer has loss aversion. If $\gamma = 0$, the marginal consumer does not have reference-dependent preferences. When $\gamma < 0$, it implies that $\beta - \alpha < 0$. When this is the case the marginal consumer gets more absolute utility from an unexpected win than an unexpected loss. When $\theta > 0$ it implies that the consumption utility of going to a game is bigger than the gain-loss utility derived from the game. When $\theta < 0$ it means that the marginal consumer gains more consumption utility from a home win than from a home loss, but the impact of the gain-loss utility is bigger than the consumption utility.

When these two parameters are combined, it is possible to make a distinction between the two theories, the UOH, and loss aversion. For loss aversion, it is necessary that fans have loss aversion and reference-dependent preferences indicated by the parameters $\gamma > 0$ and $\theta < 0$. For the UOH it is necessary that $\gamma < 0$ and $\theta > 0$, which implies that $\beta - \alpha < 0 \ \leq U^W - U^L$.

Results

The results presented below are about the group stage and knock-out stage over the seasons 2011-12 till 2016-17. To obtain robust variance estimates a clustered sandwich estimator is used. The clustered sandwich estimator specifies that the standard errors allow for intragroup correlation. This means that the observations are independent across groups, but not necessarily within groups.[3] The chosen cluster in this analysis is *Capacity*, which is the independent unit of observation. As described in the Methodology a Tobit regression model is used. The lower limit for the Tobit model is 0, as it censors the games that are punished by the UEFA, and thus the people were not able to buy a ticket for this game. The upper limit of the Tobit model is 1. These are the games that sold out.

Results - Group stage

The results of the UEFA Champions League group stage over the period 2011-17 are reported in the following table. The table contains different model specifications. Model I contains only the variables of interest, probability of winning and probability of winning-squared, and time-fixed effects and team-fixed effects. Model II contains the previous named variables and the amount of goals scored and conceded by the home and away team up to the game in that

Dependent variable: Attendance/Capacity						
	Model I		Model II		Model III	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Probability of winning	0.6925057	0.028	0.6929812	0.025	0.6702817	0.027
Probability of winning-squared	-0.3896155	0.113	-0.3889348	0.106	-0.3858898	0.103
Home team goals scored	--		0.000057	0.985	0.0014389	0.614
Home team goals conceded	--		-0.0019293	0.574	-0.0015383	0.655
Away team goals scored	--		0.0028887	0.441	0.0025746	0.504
Away team goals conceded	--		-0.0025327	0.404	-0.0020779	0.492
Ranking home team	--		--		0.0022035	0.010
Ranking away team	--		--		0.0001799	0.764
Month, day indicators	Yes		Yes		Yes	
Team, season indicators	Yes		Yes		Yes	
Observations	541		541		541	
F-test	3917.94	0.000	1884.47	0.000	5750.28	0.000

Table 3 - Censored Attendance Regression Results for group stage of the UCL with 541 uncensored observations, 5 left-uncensored observations, and 30 right-censored observations.

[3] https://www.stata.com/manuals13/xtvce_options.pdf

specific season. Model III completes the model by also taking the ranking of the home and away team into account. The estimated parameter of probability of winning is significant in Model I at a significance level of 5%. However, the probability of winning-squared is not significant at a 10% significance level. Therefore, we're not able to say something about the behaviour of football fans visiting a Champions League group stage. In model II and model III, the probability of winning is still significant at a 5% significance level. The probability of winning-squared still stays insignificant at a 10% level.

Results – Knock-out stage

The results reported in the following table describe the results of the UEFA Champions League knock-out stage over the period 2011-17. The model right-censored 23 observations of the 168 total observations. These 23 right-censored observations were equal to 1 or above, which means that the match was sold out. The following table also contains different model specifications. Model I describes the model containing the variables of interest *Probability of winning* and *Probability of winning*-squared and is controlled for the time-fixed effects and team-fixed effects. As one can see both variables, *probability of winning* and its squared version, are significant at a 1% significance level. The sign of the level term is positive, and the sign of the quadratic term is negative. This indicates that there is an inverted U-shaped relationship between the *Probability of winning* and the dependent variable *Attendance/Capacity*. Introducing additional variables to the model doesn't change the significance of the *Probability of winning* and *Probability of winning*-squared, it's also significant at a 1% significance level in model II and model III.

If one looks at model II, one can see that the variables *Home team goals conceded*, *Away team goals scored*, and *Away team goals conceded* are significant at a 1% significance level. An increase in the total amount of goals conceded prior to a match for the home team, as well for the away team, has a negative effect on the Attendance/Capacity, ceteris paribus. Conceding a higher amount of goals usually means that the team doesn't defend very well. If somebody supports the home team one wouldn't like to see their team defending bad. As every goal conceded in a home match in the knock-out stage can be crucial in progressing to the next round one would like to see their team defend well. In case the away team has a

weak defence, one would get less interested in the game as it could be too easy for the home team to score against a weak defence.

Dependent variable: Attendance/Capacity						
	Model I		Model II		Model III	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Probability of winning	0.0686941	0.000	0.0473071	0.000	0.0257332	0.000
Probability of winning-squared	-0.1486847	0.000	-0.1386013	0.000	-0.1200139	0.000
Home team goals scored	--		-0.0000116	0.708	0.0000986	0.001
Home team goals conceded	--		-0.002739	0.000	-0.002434	0.000
Away team goals scored	--		0.0014509	0.000	0.0015674	0.000
Away team goals conceded	--		-0.0015273	0.000	-0.0011706	0.000
Ranking home team	--		--		0.000364	0.000
Ranking away team	--		--		0.0004026	0.000
Round, day indicators	Yes		Yes		Yes	
Team, season indicators	Yes		Yes		Yes	
Observations	145		145		145	
F-test	2.7e+07	0.000	2.4e+06	0.000	3.0e+06	0.000

Table 4 – Censored Attendance Regression Results for knock-out stage of the UCL with 145 uncensored observations, and 23 right-censored observations

An increase in the amount of goals scored by the away team prior to the game increases the attendance/capacity, ceteris paribus. Reason for this could be that people are curious to see whether the home team can hold up against an offensive team. Goals scored and conceded by the two different teams says something about the quality of play by the teams. A team that scores a lot of goals and concedes few goals will be considered as high quality while the other way around will mean that the team lacks quality. The results suggest that people prefer to see high quality football being played.

In model III the ranking of the home team and the away team is also considered. An increase in the ranking of the home team, as well as the ranking of the away team, increases the attendance/capacity at a 1% significance level, ceteris paribus. An increase in ranking means being further away of being ranked as number 1 of Europe. Therefore, having a lower ranking as team increases the attendance/capacity ratio, ceteris paribus. The reason behind this is that lower ranked teams don't easily progress to the knock-out stage of the UEFA Champions League when one considers their performances the previous years, and thus their supporters are interested in attending a match in the stadium of these lower ranked teams in the knock-

out stage. As one saw, a lower ranking of the away team also increases the attendance to capacity-ratio, ceteris paribus. Reason for this could be that due to the lower ranking the fans supporting the away team are more willing to travel abroad to visit the game.

Another reason could be that the ranking of the team is in some way misleading. Some teams are considered as important clubs in Europe based on their history. Decades ago this team might have been one of the most important teams, but nowadays they don't perform that well in Europe anymore. Good examples of these kind of teams are AC Milan and Ajax, who have won the UEFA Champions League (previously Europa Cup I) years or decades ago, but nowadays aren't able to progress to the UEFA Champions League knock-out stage. However, it can also be the other way around. The relatively new top teams that, out of the blue, play in the UEFA Champions League with a strong squad, but are still ranked low influence this. In this case one should think about Manchester City and RB Leipzig. These teams are playing for league titles, but don't have much history of winning titles, especially not on European level, which lowers their ranking.

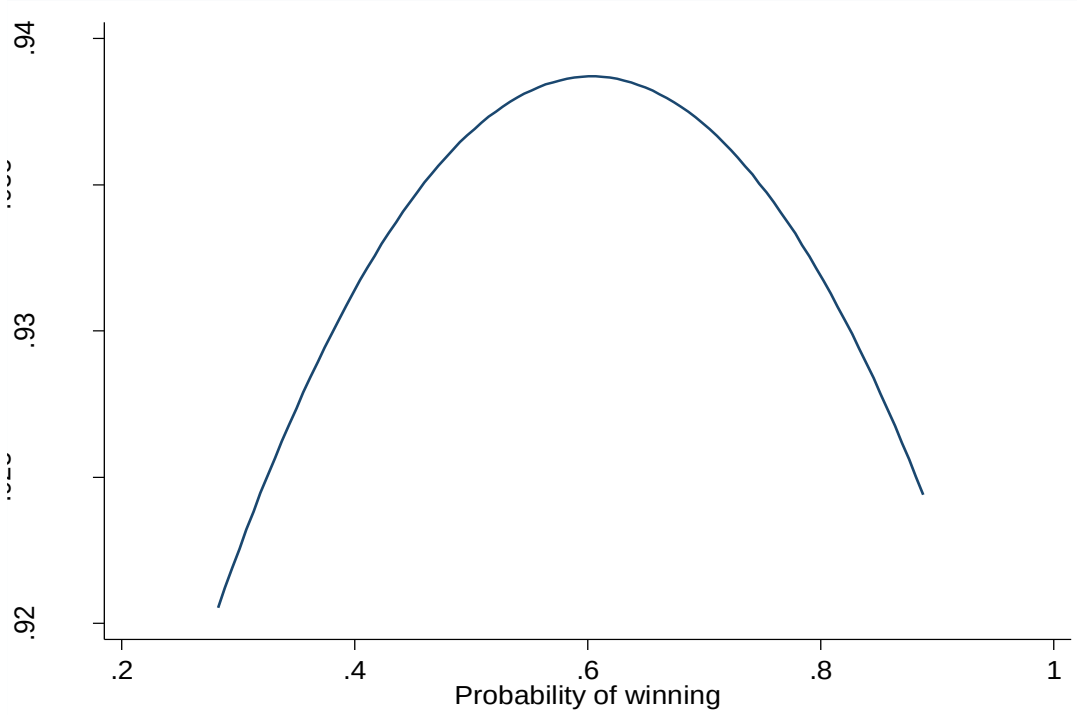

Figure 4 - Relationship between Attendance/Capacity and Probability of winning (Model III knock-out stage)

Conclusion

For a long time, the uncertainty of outcome hypothesis has been accepted by researchers explaining the behaviour of supporters visiting a sports match. However, lately more evidence has been found that people are behaving differently. The theoretical basis of Coates et al. (2014) and various empirical researches show that people are behaving according reference-dependent preferences with loss aversion. In this research the relationship between live attendance and the uncertainty of match outcome is investigated using one of the biggest football tournaments in the world, the UEFA Champions League.

The two models used are discussed in Rottenberg (1956) and Coates et al. (2014) discussing the uncertainty of outcome hypothesis and the theory of reference-dependent preferences with loss aversion. The uncertainty of outcome hypothesis states that people prefer to see tight matches over matches with superior teams playing inferior teams. The relationship between live attendance and uncertainty of outcome is concave and has a maximum between (0.5, 1). The theory of reference-dependent preferences with loss aversion states that people prefer matches with a more certain outcome over tight matches, as home team supporters rather see their team win than lose. Visiting a game where they expect to win but end up losing gives a bigger marginal utility loss compared to the marginal gain of utility visiting a game where they expect to lose but end up winning.

Literature regarding the relationship between live attendance and the uncertainty of match outcome has found results supporting both theories, however more empirical evidence has been found for the theory of reference-dependent preferences with loss aversion. Buraimo and Simmons (2008 & 2009) found evidence that in the English Premier League and in the Spanish La Liga, people behave according to the theory of reference-dependent preference with loss aversion. In their analysis they used different methods to estimate the probability of winning. In the empirical research using the data of the English Premier League they used the level and quadratic term of the probability of winning and the Theil measure. In the empirical research using the data of the Spanish La Liga they used the absolute difference between the probability of winning and the probability of losing as an indicator of how tight a match is, as well as the Theil measure. Benz et al. (2009) found evidence in favour of the uncertainty of outcome hypothesis using data of the German Bundesliga using the level and quadratic term

of the probability of winning as indicator of uncertainty of match outcome. Pawlowski and Anders (2012) also researched the relationship between live attendance and uncertainty of match outcome using data of the German Bundesliga, however they found support for loss aversion which could not be explained by home team preference.

To investigate the relationship between live attendance and uncertainty of match outcome in football the data of the UEFA Champions League over the period 2011-12 till 2016-17 has been used. As this tournament consists out of two different stages, the results are based on the group stage and the knock-out stage. The results for the group stage do not confirm nor reject any theory due to the insignificance of the key variable *Probability of winning-squared*. As this variable says something about the uncertainty of outcome of the game, it must have a significant effect.

The results of the knock-out stage are significant and the key variables *Probability of winning* and *Probability of winning-squared* show that an increase in uncertainty of the outcome increases the attendance-to-capacity-ratio, ceteris paribus. The results of the knock-out stage for these key variables are in line with the Uncertainty of Outcome hypothesis (Rottenberg, 1956). Other variables considered in extended models are goals scored by the teams prior to the game in that current season, goals conceded by the teams prior to the game in that current season and the ranking of the teams. Model III containing all these variables named above showed that all these variables had a significant effect on the attendance/capacity-ratio. The model is also controlled for time-fixed effects and team-fixed effects.

The findings of this research are interesting as the UEFA Champions League is a completely different competition than a regular domestic competition. In previous researches to the relationship between live attendance and uncertainty of outcome data of domestic leagues was used, but never those of tournaments where elimination of clubs is a possibility. In most of these researches evidence was found that people behave according to the theory of reference-dependent preferences. However, in the UEFA Champions League the possibility of elimination exists, especially in the knock-out stage. As each round in the knock-out stage consists out of only two games, from which one is played at home and one away, losing over these games combined means you're eliminated. In a domestic league clubs can get relegated at the end of a season, but not eliminated during the season. The different setup of the UEFA Champions League compared to a domestic league might be a reason for the different support

of theories. Therefore, it would be interesting to see further investigation in the relationship between live attendance and uncertainty of outcome using datasets of domestic cups to look whether there is evidence of people preferring tight matches in cup tournaments. This would also be a point of interest for the European Super League. The knock-out stage in the UEFA Champions League contains a lot of clubs that would participate in the European Super League and as the results showed the supporters prefer tight matches in the knock-out stage. However, if this is due to the tournament setup, a European Super League could maybe be not as successful as previously thought.

One of the limitations of this research is the collection of attendance data. The attendance data is from the UEFA which is representable of how many were attending the match in the stadium, however no distinction has been made between home and away supporters nor is it known whether the game is part of their season ticket or that they had to purchase it as a single ticket. One could solve this by cooperating with the participating football clubs and get access to their data. This way one can focus for example on single tickets sold for the match.

Another limitation is the way to deal with the probability of tying a game. In the UEFA Champions League a tie can be an important result for a team. In the group stage that one point can make a difference between progressing to the knock-out stage or going to the Europa League. In the knock-out stage a tie can also be of great importance especially it the first leg of the round is won by a team. A tie in the second leg will guarantee a spot in the next round for the team that won in the first leg. Therefore, a tie can't be neglected and do fans possibly go to the stadium hoping for a tied result. In this research the tie has been added to the probability of winning in case the team was lower ranked than the opponent. A possible more accurate way to deal with the tie is to look at each game individually to see what a tie would mean for the home team. Another possibility to deal with tying a game is to look at the absolute difference between the probability of winning and the probability of losing. The bigger the absolute difference, the lower the uncertainty of outcome.

Finally, a limitation is that clubs nor players are looked at as brands in this research. Especially on the level of the UEFA Champions League players and clubs can be looked at as brands which might attract people to visit a game because of the away team and thus doesn't relate to the uncertainty of match outcome. Pawlowski and Anders (2012) found that their variable Brand had a significant positive significant effect on the attendance level. Players like Ronaldo and

Messi might have that same effect on attendance levels as people might want to see these players play against the team they support. Introducing a variable Brand for players or clubs can be interesting for further research.

References

Benz, M.-A., Brandes, L., & Franck, E. (2009). Do soccer associations really spend on a good thing? Empirical evidence on heterogeneity in the consumer response to match uncertainty of outcome. *Contemporary Economic Policy Vol. 27, No. 2*, 216-235.

Buraimo, B., & Simmons, R. (2008). Do Sports Fans Really Value Uncertainty of Outcome? Evidence from the English Premier League. *International Journal of Sport Finance, 3*, 146-155.

Buraimo, B., & Simmons, R. (2009). A tale of two audiences: Spectators, television viewers and outcome uncertainty in Spanish football. *Journal of Economics and Business 61*, 326-338.

Coates, D., Humphreys, B. R., & Zhou, L. (2014). Reference-dependent preferences, loss aversion, and live game attendance. *Economic Inquiry Vol. 52, No. 3*, 959-973.

Czarnitzki, D., & Stadtmann, G. (2002). Uncertainty of outcome versus reputation: Empirical evidence for the First German Football Division. *Empirical Economics 27*, 101-112.

Gouka, M. (2017, September 15). *Waarom een Super League onvermijdelijk is*. Opgehaald van AD: https://www.ad.nl/champions-league-en-europa-league/waarom-een-super-league-onvermijdelijk-is~a771b669/

Hytner, D. (2018, May 10). *Arsène Wenger: European super league will force Premier League midweek*. Opgehaald van The Guardian: https://www.theguardian.com/football/2018/may/10/arsene-wenger-european-super-league-premier-league-arsenal

Köszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics, Vol. 71, No. 4*, 1133-1165.

Kuypers, T. (2000). Information and Efficiency: an empirical study of a fixed odds betting market. *Applied Economics, 32*, 1353-1363.

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

Platt, O. (sd). *European Super League: What would a breakaway elite competition look like?* Opgehaald van GOAL: http://www.goal.com/en/news/european-super-league-teams-format-guide/1tknnzskezzc21d9mkejsukf9o

Reekie, H., & Grez, M. (2016, May 26). *Champions League final vs. the Super Bowl*. Opgehaald van CNN: https://edition.cnn.com/2016/05/26/football/champions-league-vs-the-super-bowl-real-madrid-atletico-madrid/index.html

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