The wisdom of crowds and football players

Master thesis Economics and Business

Stefan Baan
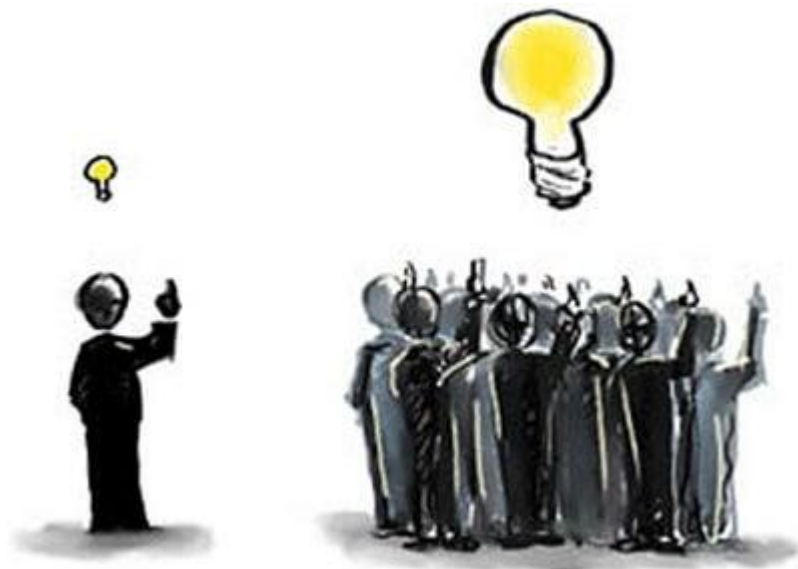
Abstract

This thesis analyzes the prediction accuracy of the wisdom of crowds on the outcomes of football matches. A simple prediction model, based on player valuations reported on the football statistics website 'www.transfermarkt.de', is drafted and used to predict the league results of the Premier League, Bundesliga and Primera División over the period 2013-2016. The resulting prediction accuracy is compared with that of four benchmark methods: ELO ratings, aggregated betting odds, a 'home team wins' model and pure chance. My analysis shows that the wisdom of crowds model is a better prediction method than pure chance and the 'home team wins' model, is able to compete with ELO ratings, but is clearly a worse forecasting method than the model based on aggregated betting odds. Furthermore, I investigated the presence of biases that potentially could influence the valuations of users on 'www.transfermarkt.de'. I find that cognitive biases are likely to influence user responses and therefore could undermine the prediction accuracy of the corresponding model. This could explain why the wisdom of crowds does not outperform aggregated betting odds. Other factors might distort the validity of user stated player valuations; further research should therefore try and focus on examining the impact of these confounders on the valuations of Transfermarkt users.

Preface

I have been told that writing a thesis could be a pitfall: something which could drive you crazy. Apart from the fact that I had to cancel some appointments this summer with an ever growing group of moaning friends, I have to point out that I actually liked it! This is no coincidence, since I genuinely appreciate the behavioral side of economics. The opportunity to dig much deeper in this relatively new field of economics, provided me with lots of positive energy to write this thesis. Besides that, over the last couple of years I learned something important about myself: I love to write. I especially love to write about topics I'm interested in.

Even though the foundations were present, none of this would have been possible without the help of my supervisor Georg Granic. I liked his personal approach and I am very thankful for the way in which he has helped me during the writing of my thesis. I have been provided with the opportunity to go my own way, something which I appreciate. I also would like to thank my parents and sisters for their support over the course of my study: they created a pleasant environment in which I have been able to perform.

I don't know what's ahead of me, but to quote one of my favorite bands Genesis:

''I know what I like and I like what I know.''

Enjoy reading!

Table of contents

1 Introduction

''Two heads are better than one'', this old proverb states that two people may be able to solve a problem that an individual can't. When this thought is extrapolated, it follows that the larger the group, the bigger the chance is a problem will be tackled. But does this also hold when prediction problems are examined? James Surowiecki thinks it does. In his book 'The Wisdom of Crowds' he explains that, under certain conditions, crowds are better forecasters than any individual could ever be. Faster market judgement, better coordination and cooperation should ensure that 'larger' groups are able to make relatively good predictions (Surowiecki, 2005).

The superiority of crowd judgments over individual predictions has been observed in many fields. Recent research shows the presence of the wisdom of crowds in the prediction of elections (Gaissmaier & Marewski, 2011; Murr, 2011; Franch, 2013), stock market fluctuations (Ray, 2006; Foutz & Jank, 2007) and sports match outcomes (Spann & Skiera, 2009; Herzog & Hertwig, 2011; Sinha, Dyer, Gimpel, & Smith, 2013). Especially the last point is important for my thesis, as it demonstrates that crowds are able to structurally outperform ranking systems and experts, whilst competing with betting odds. The universal applicability of the wisdom of crowds effect (WOC effect) has been challenged. It is well understood that for the effect to emerge four important criteria have to be met: crowd members should rely on private information (*diversity of opinion*), opinions of crowd members should be independent (*independence*) and they need to draw on local knowledge (*decentralization*). Finally, the judgments of the crowd need to be collected and aggregated into a collective wisdom (*aggregation*) (Surowiecki, 2005).

Much research has been done on the validity of those criteria. Especially the first two, diversity of opinion and independence, are found to be essential factors to ensure a crowd to be 'wise' and its prediction to be relatively accurate (Larrick, Mannes, Soll, & Krueger, 2011; Lorenz, Rauhut, Schweitzer, & Helbing, 2011). According to behavioral economists around the world, individuals' judgements and predictions are subject to biases. Even though it is widely accepted that these biases leave their mark on decisions, scientists differ in opinion regarding their implications on the wisdom of crowds (Tversky & Kahneman, 1974; Surowiecki, 2005; Davis-Stober, Budescu, Dana, & Broomell, 2015). This however, is an important issue when determining whether the predictions of a crowd could be considered 'wise' (accurate) or not (Simmons, Nelson, Galak, & Frederick, 2011).

The potential predicting power of the crowd draws the attention of economists especially to sports betting. Over the last few years, the market for betting and gambling on sports matches has indeed increased tremendously (Australian Institute of Family Studies, 2014; Jeffrey Young, 2014). The global

sports betting market is estimated to be worth around the three trillion dollars (90 percent of this amount is said to be part of the illegal gambling business; therefore final figures are hard to give) (The Daily Mail, 2015). Profits are desirable and bettors are determined to invest their money in the 'winning' bet. Every individual is free to choose their preferred strategy, but one thing remains certain: they all want to place the best bet. It is therefore interesting to investigate whether a prediction model, based on the wisdom of crowds, could compete with (or even outperform) other prediction methods.

Several studies show that crowds in certain cases indeed are better able to predict the outcomes of sports matches, than other forecasting methods do. Herzog and Hertwig (2011) for example, show in their paper that crowds can pretty accurately predict outcomes of sports matches by mere recognition, compared to other prediction measurements. In this research the crowd is represented by subjects without full knowledge about the factor of interest (no processing of all available information). However, Herzog and Hertwig show that this ignorance in fact is bliss, and ensures better predictions. This less-is-more effect is confirmed by several empirical findings (Goldstein & Gigerenzer, 2002; Smithson, 2010; Gaissmaier & Marewski, 2011). However, a paper of Pachur and Biele (2007), concerning predictions of the European Football Championship, shows that this partial ignorance of the crowd only beats chance. The accuracy of the predictions of the crowd is not found to be better than the prediction of experts. The analysis of different sports forecasting methods of Spann and Skiera (2009) on the other hand, displays contradictory results. Their study of German football matches shows that the accuracy of aggregated individual expectations and betting odds are comparable and they were found to be significantly better predictors than experts.

Conflicting results and contradictory conclusions are the order of the day. It is however of great interest to many people to know which prediction method is the most accurate (Vlastakis, Dotsis, & Markellos, 2009). Because of that, this thesis will examine whether following the prediction of the crowd is a wise decision. Or in other words: does the wisdom of crowds hold when predicting sports matches? Is the crowd really that wise? Besides answering this main research question, I will analyze whether the four criteria that are necessary for the WOC effect to emerge are violated because of the presence of cognitive biases.

To find out whether the wisdom of crowds holds, three of the largest football competitions in Europe will be examined. I will develop a prediction model by using data from 'www.transfermarkt.de': a football statistics website containing player valuations stated by registered users. This model, based on the 'wisdom of crowds' (WOC model), will be compared to other prediction methods (anchors based on pure chance, 'home team wins', ELO ratings and aggregated betting odds). Afterwards I will

take a look at the relative prediction accuracy of my model by analyzing whether the four criteria of Surowiecki are violated or not, because of cognitive biases of the Transfermarkt valuators.

The remainder of this thesis is organized as follows. In the next Chapter the theoretical framework is discussed. I will elaborate on 'www.transfermarkt.de', Surowiecki's wisdom of crowds, the valuation of football players and some cognitive biases. Chapter 3 deals with the methods I use to compose and compare the WOC model. Afterwards, Chapter 4 displays the results of the comparisons between my model and the benchmark models; this is followed by an analysis of the (distorting) influence of cognitive biases on Surowiecki's four criteria; therefore on the eventual effect of the wisdom of crowds. Finally, Chapter 5 shows the discussion of this thesis, which includes the conclusion, limitations and some recommendations for further research.

2 Theoretical Framework

This chapter elaborates on all theoretical insights that are important for this thesis. First of all, the operation of the football statistics website 'www.transfermarkt.de' is explained. Besides that, Surowiecki's wisdom of crowds and its four essential criteria are dealt with. This is followed by an analysis of the way in which football players are valued by the 'crowd'. The theoretical framework is concluded by introducing several cognitive biases that are able to distort the valuation of football players by Transfermarkt users.

2.1 Transfermarkt

To answer the questions raised in the introduction, I will use data from 'www.transfermarkt.de'. This website contains a large database of statistics regarding football clubs and players and is often used in scientific research (Baur & Lehmann, 2007; Torgler & Schmidt, 2007; Bryson, Frick, & Simmons, 2013; Herm, Callsen-Bracker, & Kreis, 2014; Pavlović, Milačić, & Ljumović, 2014). One of the aspects which the site is most known (and used) for, is the player valuation system. The value of a football player is constructed with the help of the opinion of registered users. These users are able to value a player according to their own standards and expectations. If the initial market value of a player is regarded either too high or too low, Transfermarkt visitors can suggest a revision. Transfermarkt offers them the opportunity to discuss the formation of the 'market value' of the football players on different forums. Subsequently, Transfermarkt aggregates these valuations and updates the players' market value accordingly. With the help of the player market value, team values are composed (Transfermarkt, 2016). These team values consist of the aggregated market value of all players in a team's selection; this forms the basis of the WOC model in this thesis.

2.2 Surowiecki's wisdom of crowds

James Surowiecki is an avid defender of the wisdom of crowds. His book delivers many examples from everyday life to empower his thoughts. These examples are not exclusively there to prove that a large group is relatively good in judging and predicting, but also to provide evidence for the existence of several criteria that must hold for a crowd to be able to deliver relatively good predictions.

The first essential element is *diversity of opinion*. In order to add value to the crowd, an individual should have private information, or at least an 'eccentric interpretation of the facts' (Surowiecki, 2005). This diversity ensures that different perspectives and approaches are added to the group. More homogeneous groups are at risk of being less accurate in their judgements, because of the

greater impact individual biases and heuristics will have. According to Surowiecki, each person's guess, contains both information and error. Similar people with similar thinking patterns are more likely to make identical mistakes, which consequently causes systematic errors in the decision making process. These errors however, will cancel themselves out if a sufficiently large diverse subset of people is judging (Surowiecki, 2005; Larrick, Mannes, Soll, & Krueger, 2011; Simmons, Nelson, Galak, & Frederick, 2011).

The second criterion is *independence*: the opinion of the people in the crowd may not be 'determined by the opinions of others' (Surowiecki, 2005). In a way this independence has some overlap with the first criterion. Just like diversity of opinion, independence makes sure that individual´s errors don´t become correlated. Even though rationality is often referred to as the cornerstone of good judgements, it's presence is not a necessity for a crowd to perform well. Biased people do no harm to the group performance, as long as they are independent (Surowiecki, 2005).

Besides diversity of opinion and independence, Surowiecki claims the indispensableness of *decentralization*. Each group member needs to extract their expertise from their own (local) environment. Certain knowledge is only available through encountering specific events that are (mostly) exclusively accessible. If every individual's input draws on this local knowledge, more and deeper insights will arise (e.g. knowledge obtained by working for a company with a unique marketing formula); this will enrich the crowd's aggregated wisdom. Simultaneously, the input of these 'specialized' people will increase the independence (Surowiecki, 2005).

Surowiecki's final criterion states that all private judgments must be *aggregated*. After all crowd members have provided their knowledge, this input has to be processed and turned into a collective wisdom. This will lead to an aggregated wisdom that is greater than the initial individual wisdom: the wisdom of the crowd (Eysenck, 1939; Wolfers & Zitzewitz, 2004; Surowiecki, 2005; Larrick & Soll, 2006).

2.3 Valuation of football players

In order to find out if the registered Transfermarkt users measure up to Surowiecki's criteria, it is important to know how these individuals value football players. A paper of Frick (2006) collects all findings of recent research about the effects of dozens of variables on the transfer fees in European football. Among the factors found to be correlated with transfer fees are: age, remaining contract years, career games played, career goals scored, international caps and even characteristics such as the average attendance of the buying and the selling club and goal difference of the selling club in the previous season. Some research also emphasizes 'off-field' determinants, like media attention

and personal background (Tunaru & Viney, 2010; Herm, Callsen-Bracker, & Kreis, 2014). 'www.transfermarkt.de' Moderators have even posted guidelines for the player valuations on the different forums, to help the registered users in their valuation process. However, these explanatory variables are not necessarily considered by Transfermarkt users. Herm et al (2014) did some extensive research on the variables that *are* taken into account by individuals when judging football players. They found that Transfermarkt users mainly use age, precision, scoring, assertion and flexibility as benchmarks for their player valuations. Compared to the variables Frick mentioned, such valuations seem to be a bit myopic. However, when these judgments are compared to actual transfer fees, Transfermarkt users appear to be very good predictors.

2.4 Cognitive biases

People are structurally subject to their own biases when judging, deciding and predicting (Schwenk, 1988; Das & Teng, 1999; Simmons, Nelson, Galak, & Frederick, 2011). Cognitive biases form the basis for irrational and systematically distorted decision making. This decision making is affected by illogical reasoning, which is caused by people's own subjective reality (Bless, Fiedler, & Strack, 2004; Haselton, Nettle, & Murray, 2005). Cognitive biases are often referred to as the leading cause of bounded rationality and play into the hands of distorted decisions (Kahneman, Knetsch, & Thaler, 1991). Like stated before, this bounded rationality does not need to be an obstruction in this thesis, as long as Surowiecki's criteria hold. However, there are several cognitive biases that could form a potential threat to the Transfermarkt valuations.

One of the most debated cognitive biases is *anchoring.* Anchoring suggests that people, when deciding, are heavily influenced by the first piece of information they receive. This piece of information, 'the anchor', acts like a reference point: estimates depend on this anchor, whether it provides relevant information or not (Investopedia, 2016). Tversky and Kahneman (1974) state that anchoring proves to be an effective heuristic for predicting. However, they also found that individuals insufficiently adjust their estimations from predetermined (irrelevant) anchors, which in its turn could lead to systematic errors. Anchoring is often found to be so robust, that it seems difficult to eliminate or even diminish its effect on people's estimations (Wilson, Houston, Etling, & Brekke, 1996; Strack & Mussweiler, 1997; Chapman & Johnson, 1999; Mussweiler & Strack, 2000).

Another famous bias is *conformity*. People who want to conform adjust their behavior in such a way, it is normatively and socially accepted by the group they want to be part of (Cialdini & Goldstein, 2004). The effects of conformity on human behavior are not universally regarded as the same by researchers. Some say being subject to conformity leads to appropriate behavior, whilst others underline its harmfulness (Krueger & Massey, 2009; Larrick, Mannes, Soll, & Krueger, 2011). One of

the most famous experiments regarding conformity, is conducted by Solomon Arsch (1951). He gathered several students to answer a simple question. All but one of these students were part of the experiment; only one was the actual subject. All students received two cards: the first one showed a single line whilst the second one displayed three (A, B, C) labeled lines with different lengths. All participants had to say out loud which one of the three lines, in their opinion corresponded with the one on the first card. Only one of the three lines was exactly the same length, whilst the others were either obviously longer or shorter. The students who were part of the experiment answered first, after which the actual subject got the opportunity to state his answer. When the collaborating students answered 'correctly', the subject went along with the right answers in every case. However, when the question was answered 'incorrectly', as many as 36,8% of the subjects would agree with the other wrong participants. This experiment is a great example of the fact that peer pressure is able to alter people's decisions very easily.

A cognitive bias that is somewhat related to conformity, is the *bandwagon effect.* Individuals who are subject to this bias, are inclined to alter their behavior and preferences in such a way that these are in accordance with those of a growing group of people. The fact that something gets attention from an increasing amount of people, also attracts other individuals (Marsh, 1985; Biddle, 1991). Trends and hypes emerge because of this bandwagon effect (Long, Fox, & York, 2007). The bandwagon effect has been scrutinized considerably, especially in the field of politics; many papers show a significant impact of this bias on voting strategies (Straffin Jr, 1977; McAllister & Studlar, 1991; (Mehrabian, 1998). Hong and Konrad (1998) even state that some voters choose against their own preference, in favor of the winning candidate or party. The bandwagon effect could therefore be regarded as a considerably strong bias and its influence on judging should not be underestimated.

Another cognitive bias that has been a factor of interest of many researchers, is *wishful thinking.* This bias causes people to actually think and believe their hopes and desires have a bigger chance of occurring than they in reality have (Bastardi & Ross, 2011): it is described as the link between 'preference and expectation' (Babad, Hills, & O'Driscoll, 1992). Wishful thinking is detected in a wide range of situations: when predicting election outcomes (Bar-Hillel & Budescu, 1995; Krizan, Miller, & Johar, 2010), price fluctuations (Seybert & Bloomfield, 2009; Mayraz, 2011) and sports match outcomes (Babad E. , 1987; Price, 2000; Massey & Simmons, 2011). Because of the difference in attached and real probability of an event, wishful thinking can cause disproportionate inaccurate predictions and expectations.

Even though previous research shows the relative accuracy of Transfermarkt player valuations (e.g. compared to actual transfer fees), this is no prove of the usefulness of its player market values for

predicting the outcome of football matches. The biases mentioned above could form an obstacle for the Transfermarkt users to provide valuations satisfying Surowiecki's criteria. Subsequently, this could lead to distorted player market values, which could diminish the WOC effect and hamper the prediction accuracy of the model in this thesis. In the next few chapters I will discuss if the 'wisdom of the Transfermarkt users' is able to generate a useful prediction model.

3 Methods

This section starts off with a brief explanation of the used data, after which I will elaborate on the operation of the WOC model and the way in which the model is drafted (calibrated). Finally, several benchmark prediction methods are introduced that are used as a reference point for the results of the WOC model.

3.1 Data

To elicit a WOC effect, aggregated Transfermarkt team values are used. These values are updated twice a month (on 1st and 15th). Transfermarkt provides aggregated team values from November 2010 on; all years from that month on, are taken into account. The first 2,5 years (November 2010 – May 2013) are used as a benchmark to create a prediction model that is as accurate as possible (*calibration of the WOC model*). The remaining three years (August 2013 – May 2016) are used to compare the predictions of the WOC model with the actual results. In this thesis, the league games of the three largest European competitions (Premier League, Bundesliga, Primera Division) (UEFA, 2016) are analyzed. These national leagues are economically the most interesting football competitions. Accountancy firm Deloitte composes a yearly ranking list with the highest earning football clubs in the world; over the course of the last fifteen years English, Spanish and German teams feature prominently (Deloitte, 2016). By analyzing three different national leagues, cultural and league-dependent characteristics are less likely to induce skewed results, which makes the analysis more valid.

3.2 Composition of the WOC model

The WOC model predicts a match based upon the most recent team value updates on Transfermarkt. The two team values a match involves are converted into a ratio: home team market value/away team market value. Ratio's greater or equal to 1, thus indicating an equal or higher home team value, will predict a home team win. Ratio's below 1, thus indicating a lower home team value, will predict an away team win (draws are disregarded). The predictions are compared to the actual game results, after which the prediction accuracy of the calibration model is calculated.

For this thesis I decided not to adjust the team values according to the total amount of players in a team. Teams with a small selection and relatively many high-value players, have high average player values. In my model these teams could become 'undervalued', compared to teams with broad selections. However, having access to many players could be considered an advantage, because of the fact that broad selections should be better able to absorb injuries, suspensions and form crises, compared to smaller selections. The positive impact of selection size on match outcomes, is

confirmed by Peeters (2016): a broader selection can therefore be seen as an indication of a greater winning chance. Besides that, I want to keep the prediction model as simple as possible; an analysis of mere team values contributes to this goal. To create a prediction method - based on the wisdom of crowds - that is as accurate as possible, it is necessary to calibrate the model. In this phase the team values are adjusted to maximize the prediction accuracy. Matches in the Premier League, Bundesliga and Primera División over the course of the seasons 2010/2011, 2011/2012 and 2012/2013 are taken into account in the calibration phase to set up a relatively accurate prediction model. After the calibration model is composed, the seasons 2013/2014, 2014/2015 and 2015/2016 of the aforementioned competitions are predicted, according to the drafted (calibrated) WOC model.

3.3 Calibration of the WOC model

<u>**Table 1**</u>

Calibration model's prediction accuracy of the three analyzed competitions

Competition	Matches (N)	Correctly predicted	Prediction accuracy
Primera División	1051	596	56.71%
Bundesliga	828	420	50.72%
Premier League	1041	527	50.62%
Total	2920	1543	52.84%

A total of 2920 matches over the seasons 2010/2011, 2011/2012 and 2012/2013 are analyzed. The prediction accuracy of the calibration model is comparable for the English and German football leagues (50.62% and 50.72%). The Spanish competition however, appears to fit better with the drafted model (56.71%). Of all 2920 matches, 2239 games are predicted to result in a home win (76.78%), whilst the remaining 681 are predicted to result in a win for the away team (23.32%).

Various literature shows a clear home advantage in sports matches (Courneya & Carron, 1992; Harville & Smith, 1994; Nevill & Holder, 1999; Garicano, Palacios-Huerta, & Prendergast, 2005); therefore the home team market values are, according to the most accurate predictions in the benchmark years, multiplied by a certain factor.
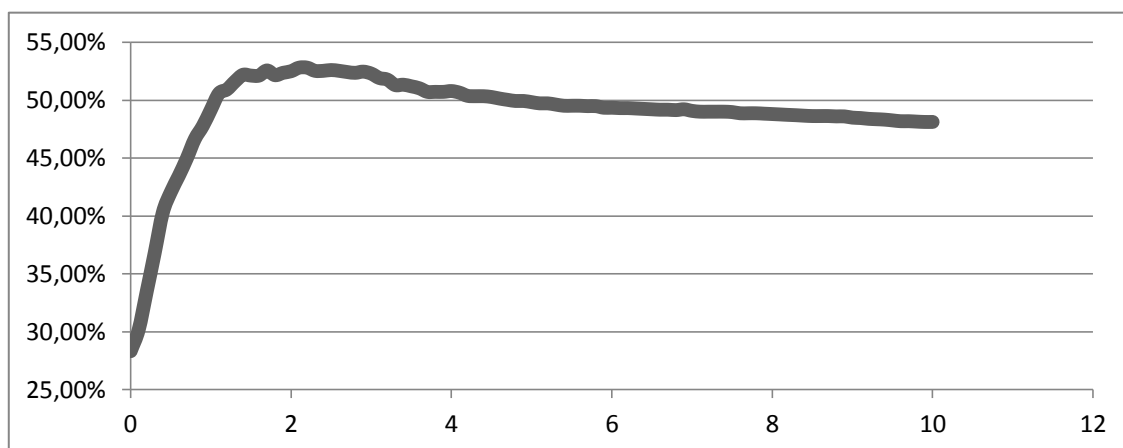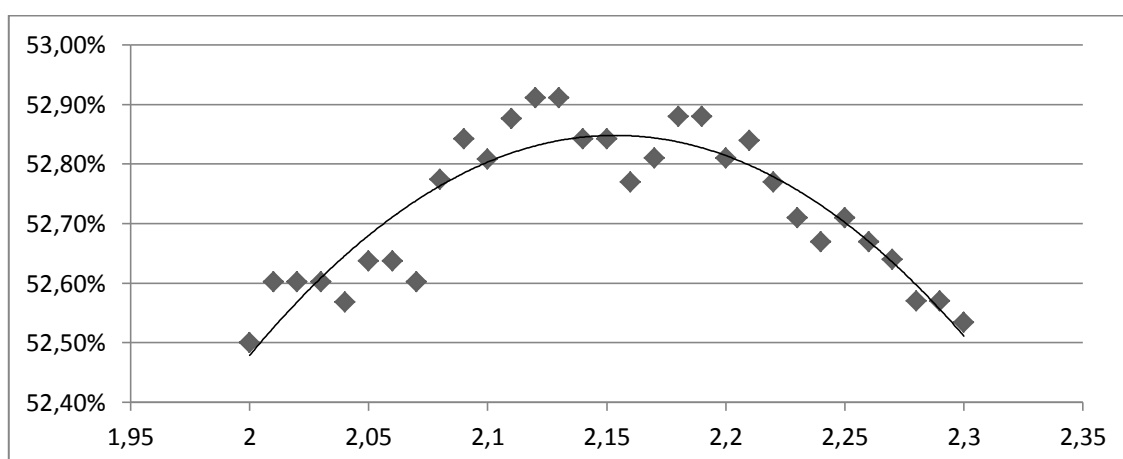
Figure 1: *Prediction accuracy for different home team value multipliers*

Figure 1 shows the distribution of the prediction accuracy for different home team value multiplications. When the home team value is multiplied by zero, the prediction accuracy of the calibration model will be the same as that of a 'simple' prediction model that always predicts an away win (after all, home team values will always be smaller (i.e. zero) than away team values). After a rapid increase, the prediction accuracy stagnates around the 52%. This is displayed by the relatively flat area in the graph between the home team value multipliers 1.3 and 3.2. After reaching this 'top', the prediction accuracy is slowly decreasing with every home team value multiplier increase, until this multiplier is sufficiently large to ensure every home team value to be greater than away team values (this is the equivalent of the 'home team wins' model). To obtain the best home team value multiplier, the area with the highest prediction accuracies is closely watched:

Note. Trend line is drafted with the help of Excel-function 'trend line'

Figure 2: *Highest prediction accuracies for different home team value multipliers with trend line*

As one can see in figure 2, the distribution of the prediction accuracy for different home team value multiplications tends to follow a parabola. The trend line confirms this notion and displays a parabola with formula y = -0.1568x^2 + 0.6754x – 0.1987 and an 'R squared' of 0.8323. This means that more than 80 percent of the variance in prediction accuracy(=y), is predictable from the home team value multiplier(= x) (Stat Trek, 2016). Even though home team value multipliers of 2.12 and 2.13 result in the highest prediction accuracy in the calibration phase (52.91% of the matches is predicted correctly), this is not necessarily the best multiplier in the actual WOC model. Because all relatively high prediction accuracies are taken into account by the trend line, this is a fitting instrument to determine the optimal home team value multiplier. The corresponding home team value multiplier of the parabola's top is used as the final multiplier in the WOC model: 2.15[1].

During the calibration phase it also became clear that, in order to acquire the best prediction method, the model should only predict either a win or a loss. Besides the fact that accuracy is improved, the model will remain as simple as possible: something which is recommended by many scientists (Armstrong, 2005; Goldstein & Gigerenzer, 2009). Furthermore, according to Ockham's Razor, a simple model should be preferred if it predicts just as good or even better than a more complex one (Jefferys, 1992; Myung & Pitt, 1997). Unnecessary complicating elements should be left out, because models including more variables eventually have a smaller chance to be a reflection of the reality.

3.4 Benchmarks

In order to say something about the usefulness of the WOC model, its results are compared to the prediction accuracy of several benchmarks. These other prediction methods differ from each other in complexity (i.e. information processing).

First of all, the simplest of all prediction methods is based on pure chance. Because all prediction methods only predict either a win or a loss, this method also is restricted to these two possible outcomes. Draws are disregarded: the model based on pure chance only predicts home and away wins. This method of forecasting does not include any information: pure chance indicates that every possible outcome has an equal chance ($\frac{1}{2}$) of occurring. Predictions based on pure chance should be perceived as a lower benchmark.

Another naïve prediction method that is used, is one that always predicts a home win. Based on the aforementioned literature about home team advantage, one could state that this method is slightly

---

[1] Calculations in Appendix

more sophisticated than the previous one based on pure chance. However, there is still very little (recent) information included in this way of forecasting.

Besides the two fairly simple prediction models, the much more advanced ELO ratings are employed to act as a benchmark. The ELO rating system calculates a club's relative strength based upon past results; its development is dependent of two factors. The match result is the first indicator of what kind of exchange of points will be made. For example, when Bayern Munich loses against Borussia Dortmund, points will be transferred from Bayern to Borussia. The new update will show a lower ELO rating for Bayern and a higher rating for Borussia (whereas the amount of points lost by Bayern, is equal to the amount of points won by Borussia). The initial ELO rating of both teams is the second determinant for the values updates. The greater the initial difference in ELO rating between the two opposing teams, the smaller the updates when the team with the higher ELO rating wins (winning team gains and losing team loses relatively little) and the bigger the updates when the team with the lower ELO rating wins (winning team gains and losing team loses relatively much). The used source for the obtained ELO ratings also considers home team advantage and goal difference (Schiefler, 2015). Just like the WOC model, the ELO ratings model will never predict a draw: in this analysis, the team with the highest ELO rating will win the match (regardless of the size of the difference). When the ELO ratings are equal for both home and away team, a coin flip will determine the prediction of the outcome. ELO ratings are updated after every single match played; it is therefore possible that this benchmark model is provided with the most recent performance information (Schiefler, 2016).

Finally, the odds of seven different betting sites are taken into account (Bet365, Betway, Betvictor, Ladbrokes, William Hill, Pinnacle Sports, and Interwetten). Bookmakers of betting sites announce the odds of the possible outcomes (home win, draw, away win) before every match. High odds indicate that bookmakers don't expect the corresponding outcome to happen: the chance of it occurring is perceived as relatively low. Low odds on the other hand, indicate a high perceived occurrence probability by the bookmakers. For example, consider the following match with the corresponding odds:

Table 1

Bet 365's betting odds of Manchester United - Bournemouth

Match	Home win	Draw	Away win
17-5-2016: Manchester United - Bournemouth	1.67	4.2	5.25

Source: http://www.football-data.co.uk/

Bookmakers of Bet365 expect Manchester United to win the match against Bournemouth; therefore odds are relatively low for a bet on a United win (1.67), whilst odds are high for a Bournemouth win (5.25). The lowest of the three published odds, indicates the result the bookmakers attach the biggest probability of occurrence to. Therefore relatively small amounts of money can be won by placing such a bet (1.67 times the deposit will be paid out when Manchester United wins). Because the lowest odds reflect the bookmakers' most expected outcome, they will also reflect the model's prediction in this analysis (Štrumbelj, 2014).

The odds of all seven betting sites are aggregated. Aggregating odds is an easy and efficient way to diminish inefficiencies of individual bookmakers (Herzog & Hertwig, 2011), which increases the model's accuracy. Aggregated betting odds form the basis of a relatively accurate prediction method, which is likely to beat most of the other benchmarks (Vlastakis, Dotsis, & Markellos, 2009; Hvattum & Arntzen, 2010). This is due to the fact that betting markets are better able to incorporate available information, compared to other benchmarks (Forrest, Goddard, & Simmons, 2005). Besides that, bookmakers are able to update their odds up to the very last moments before a match, which allows them to take the most recent available information into account (e.g. injuries, suspensions, transfers) (Hvattum & Arntzen, 2010).

According to this forecasting method, the betting odds of the seven betting sites never predict a draw (i.e. odds for a draw are always higher than for either home or away wins). When the aggregated betting odds are the same for both a home and an away win, bookmakers attach the same probability to both a home and an away win. When this is the case, a coin flip will determine the prediction of the game (i.e. 50% chance of predicting home win & 50% chance of predicting away win); otherwise the outcome with the lowest odds will reflect the predicted outcome.

After all benchmarks are analyzed, their prediction power is compared to the accuracy of the WOC model. To test for significant deviations from the other prediction methods, a binomial test is carried out.

4 Results

To start off this section, the actual game statistics of the seasons 2013/2014, 2014/2015 and 2015/2016 for all three different competitions are displayed. Subsequently, the results of the WOC model are analyzed, after which this model's characteristics are compared to the outcomes of the four benchmark prediction methods. Finally, a binomial test will show the relative accuracy of the different models.

Table 2

League results Premier League, Bundesliga and Primera División 2013-2016

Season	Home wins	Away wins	Draws
Premier League			
2013/2014	47.11%	32.37%	20.53%
2014/2015	45.26%	30.26%	24.47%
2015/2016	41.32%	30.53%	28.16%
Total	**44.56%**	**31.05%**	**24.39%**
Bundesliga			
2013/2014	47.39%	31.70%	20.92%
2014/2015	47.39%	25.82%	26.80%
2015/2016	44.12%	32.68%	23.20%
Total	**46.30%**	**30.07%**	**23.64%**
Primera División			
2013/2014	47.11%	30.26%	22.63%
2014/2015	45.00%	31.05%	23.95%
2015/2016	48.16%	27.63%	24.21%
Total	**46.75%**	**29.65%**	**23.60%**

Source: http://www.football-data.co.uk/

In all competitions over all seasons, home wins are most common. It is also clear that a draw is the least common match outcome (which explains why bookmakers never expect a draw). The Bundesliga and Primera División show a similar distribution of match outcomes, with around the 46.5% home wins, 30% away wins and 23.5% draws. The Premier League shows relatively less home wins (44.56%) and more away wins (31.05%) and draws (24.39%).

4.1 WOC model

Table 3

Results WOC model

	Number of matches	Correctly predicted	Prediction accuracy
Home wins	2398	1271	53.00%
Away wins	800	412	51.50%
Primera División	1140	618	54.21%
Premier League	1140	590	51.75%
Bundesliga	918	475	51.74%
Total	3198	1683	52.63%

The WOC model analyzes a total of 3198 matches over a period of three seasons in England, Germany and Spain. The home team value multiplier of 2.15 results in a home win prediction of 74.98% of the games (2398 matches), whereas the other 25.02% of the games (800 matches) is predicted to result in an away win. In contrast with the results of the calibration model, the prediction accuracy of both home and away wins is much closer (53.00% and 51.50%). The WOC model consequently forecasts 1683 out of the 3198 matches correctly, which results in a total prediction accuracy of 52.63%. The prediction accuracy of the WOC model regarding the three different competitions is similar to the accuracy of the calibration model: the German Bundesliga and the British Premier League show almost the same forecasting accuracy, whilst the Spanish Primera División appears to be easier to predict.

4.2 Comparisons with benchmarks

The 'home team wins' model, is clearly a better forecasting method than the pure chance model is. During the three seasons, 45.84% (1466 out of 3198) of the league games in the three largest European competitions resulted in a home win. It follows that this 'home team wins' model likewise predicts 45.84% of the league games correctly. Besides the home wins, a total of 968 away wins (30.27%) and 764 draws (23.89%) have been observed in the analyzed dataset.

Table 4

Comparisons prediction accuracy WOC model and benchmark models

Model	Home pred. acc.	Away pred. acc.	Total pred. acc.
Aggregated betting odds	55.13%	51.27%	53.99%
Wisdom of crowds	53.00%	51.50%	52.63%
ELO ratings	60.54%	42.09%	51.31%
'Home team wins'	45.84%	-	45.84%
Pure chance	45.84%	30.27%	38.05%

Note. 'Home pred. acc.' stands for the prediction accuracy of the games that are forecasted to result in a home win. 'Away pred. acc.' stands for the prediction accuracy of the games that are forecasted to result in an away win. 'Total pred. acc.' stands for the prediction accuracy of all games.

Pure chance predicts that exactly half of the matches will result in a home win and the other half of the matches will result in an away win (1599 home wins & 1599 away wins). Half of the actual home wins is predicted correctly by this model, which results in a home prediction accuracy of 45.84% ((1466/2)/1599). Accordingly, this forecasting method yields an away prediction accuracy of 30.27%.

The forecasting method relying on ELO ratings predicts more than half of the league games correctly. With a prediction accuracy of 51.31% ELO ratings appear to be a slightly less accurate predictor than the Transfermarkt valuators are. A big contrast with the WOC model is the fact that ELO ratings exactly predict the same number of home wins as away wins, whereas the former model predicts three times as much home as away wins. This difference is also reflected by the prediction accuracy of both home and away wins: more than 60% of the predicted home games actually resulted in a home win, whilst only a mere 42% of the forecasted away games was predicted correctly. The distribution of the prediction accuracy over the three analyzed competitions however, is roughly comparable to the other models: the Primera División shows a slightly higher prediction accuracy than the Premier League and Bundesliga.

Finally, the last benchmark, the prediction method based on aggregated betting odds, shows a prediction accuracy of almost 54%. Existing literature already suggested the likelihood of the dominance of predictions based on the odds of bookmakers, compared to other forecasting methods. This method is often used as an upper benchmark when comparing several prediction models (Vlastakis, Dotsis, & Markellos, 2009; Hvattum & Arntzen, 2010); results in this thesis suggest that this is justified. Aggregated betting odds do not only appear to be better predictors than the naïve prediction methods, but also than the more sophisticated forecasting models using ELO ratings

and Transfermarkt valuations. According to this model, just over 70% of the matches is predicted to result in a home win, of which 55.14% is predicted correctly. The remaining 29.60% of the analyzed games is predicted to result in an away win, of which 51.24% is predicted correctly. Just like all other benchmarks, the prediction method based on aggregated betting odds shows similar prediction accuracies over the three analyzed national leagues.

Table 5

Comparisons prediction accuracy between the three competitions

Competition	ELO ratings	Aggregated betting odds	WOC
Primera División	52.72%	55.18%	54.21%
Premier League	50.96%	53.68%	51.75%
Bundesliga	50.00%	52.89%	51.74%

4.3 Binomial test results

Comparisons of the predictions of the WOC model, along with its benchmark models, to the league results of the Premier League, Bundesliga and Primera División over the seasons 2013/2014, 2014/2015 and 2015/2016, result in a distinctive sequence, regarding prediction accuracy. The WOC model is only topped by the prediction method based on aggregated betting odds. ELO ratings and the remaining two naïve predictors appear to be no match for the Transfermarkt valuators. A binomial test is conducted to find out whether these overall prediction accuracies are valid and generalizable. The WOC model will, in succession, be compared to all other prediction methods.

Under the null hypothesis I assume that the WOC model is equally good in predicting league games as the benchmark models; this results in a two-tailed binomial test.

$H_0: \pi = .5$
$H_1: \pi \neq .5$

Every method predicts three seasons in three different national leagues, which creates nine observations. Every observation (e.g. Premier League, season 2014/2015) has its own prediction accuracy, depending on which model is used. All observations (i.e. prediction accuracies) of the benchmark models are compared to the observations of the WOC model.

Table 6 displays the binomial test results. Every time the WOC model's prediction accuracy beats the benchmark models, it is called a 'success'. Subsequently, the two-sided binomial test will show the corresponding p-values.

Table 6

Binomial test results

Model	Successes	P-value (two-tailed)
Pure chance	9/9	0.00390626**
'Home team wins'	9/9	0.00390626**
ELO ratings	7/9	0.1796875
Aggregated betting odds	1/9	0.9609375*

Note: Statistical significance levels at 5% and 1% are indicated by * and **.

The WOC model beats the two naïve models nine out of nine times, which gives a two-tailed P-value of 0.004. Because this P-value is smaller than either a 5%- or 1%-significance level, the null hypothesis is rejected. This means that the alternative hypothesis has to be accepted, which consequently means that the model based on Transfermarkt valuations is a significantly better prediction method than the models based on pure chance and 'home team wins'.

A comparison with ELO ratings, shows that the WOC model beats the former method seven out of nine times. The corresponding two-tailed P-value is 0.180, which is not below the 5%-significance level and indicates that the null hypothesis cannot be rejected. The WOC model is therefore no significantly better prediction method, than the model based on ELO ratings.

Finally, the only method that was able to display a higher prediction accuracy than the WOC model is analyzed. The model based on aggregated betting odds beats the predictions of the WOC model eight out of nine times (WOC model is only one out of nine times the better predictor). The accompanying P-value is 0.961; because the null hypothesis assumes that both models are equally good in predicting ($H_0: \pi = .5$), the binomial distribution is symmetric (Lane, n.d.). This means that the two-tailed P-value of 0.961 can be interpreted as a P-value of 0.039 (= 1 - 0.996), which is smaller than the 5%-significance level. Consequently the null hypothesis is rejected and the alternative hypothesis is accepted. This indicates that aggregated betting odds are a significantly better basis for predicting football matches, than Transfermarkt valuations are.

4.4 Biases of Transfermarkt users and Surowiecki's four criteria

Even though Surowiecki provides evidence for the superiority of the wisdom of crowds, the observed results in this thesis are contradicting. This section therefore investigates why the Transfermarkt valuations are not able to predict as well as they *in theory* should be.

The first of Surowiecki's criteria is diversity of opinion: a diverse crowd should be able to cancel out each other's errors. Due to the fact that Transfermarkt is originally a German website, their registered users are mainly German; especially Bundesliga teams are supported (Peeters, 2016). This means that wishful thinking is given free rein to have a distorting impact on the wisdom of crowds. Football fans are more likely to give disproportionate high values to players of their favorite team, compared to players of teams they don't support. This bias can therefore, boosted by the relatively homogenous group of valuators, result in inaccurate player valuations on Transfermarkt.

However, Peeters (2016) tested for the effect of wishful thinking on the valuations of Transfermarkt users, and found no evidence for its existence. Even though Peeters did not find any evidence for the (disturbing) influence of wishful thinking, diversity of opinion could be harmed in other ways. In order to acquire complete team values, many players need to be evaluated. To cancel errors out, player values need to be created by a sufficiently large and diverse crowd. However, in practice the 'bigger' names in football are attracting more people to the value discussion, than the less known players. For example, Bundesliga players Sascha Burchert (at the time player of Hertha BSC) has been valued by just two people, whereas Pierre-Emerick Aubameyang of Borussia Dortmund has been analyzed by 365 Transfermarkt users (Transfermarkt, 2016). Less known players are therefore at greater risk of receiving 'biased' valuations, which could hamper the prediction accuracy of the team values of Transfermarkt.

Surowiecki's independence criterion is harmed by several cognitive biases. Especially the three cognitive biases anchoring, conformity and the bandwagon effect, are likely to reduce the prediction accuracy of the Transfermarkt valuations by violating this second criterion. Because initial player market values on Transfermarkt can be perceived as the first bit of information an individual gets, it is not unlikely that this will act as a reference point registered users employ to estimate. This anchor however, is already set by other Transfermarkt users, which makes anchoring a possible threat to independence. Besides anchoring, the registered Transfermarkt users could also fall victim to the conforming bias. When the market value of a player is regarded either much too high or low, and this value is constantly being confirmed (i.e. on the forums) by other Transfermarkt users (market value

stays that high or low), individuals could become afraid to disagree with the majority. However, it is possible that this individual has some specific (local) information to support his or her valuation; because of conforming, this will not be taken into account in the updated Transfermarkt player valuations. This bias therefore harms independence and subsequently the probability that Transfermarkt valuations are as accurate as possible. Finally, the bandwagon effect could distort the valuations of the registered users of Transfermarkt. An increase in market value on Transfermarkt, indicates that the player of interest on average is gaining appreciation by the registered users. Subsequently, this could lead to individuals 'following the hype', which in its turn causes systematic errors and a violation of Surowiecki's independence criterion.

The third criterion, decentralization, requires local knowledge of the crowd members. This can (partly) be acquired by demographically widely dispersed Transfermarkt users. Therefore, by preference, the valuators need to come from different regions and nations around the world. Because the majority of the Transfermarkt users are from Germany, it is however the question if this criterion will hold.

Finally, the last of Surowiecki's criteria requires an aggregating mechanism that collects all individual valuations. Transfermarkt processes the input of their users, and composes player and team values based on the 'wisdom of the crowd'; therefore the aggregation criterion is met.

5 Discussion

5.1 Conclusion

Player valuations of registered users of the football statistics website 'www.transfermarkt.de', provide the basis for this thesis' 'wisdom of crowds' prediction model. This model is compared to four benchmark prediction methods, to find out if larger groups are able to relatively accurately predict the outcome of sports matches. The conducted binomial test shows that the WOC model proves to be a significantly better prediction method than the two less sophisticated benchmark methods, based on pure chance and 'home team wins'. Besides that, it is able to compete with the much more sophisticated prediction method based on ELO ratings. However, in comparison with the model based on aggregated betting odds, the WOC model is a clearly less accurate prediction method. Various studies show similar results: the crowd's wisdom is able to beat the less sophisticated prediction methods, but appears to be less accurate than models based on betting odds. This is mainly due to the fact that betting odds can be updated up till the very last moment before a match; a prediction model based on betting odds is therefore better able to take the most recent available information into account. On the other hand, the WOC model is presumably not as effective in predicting as it could be, which makes the difference between the two models even larger. Cognitive biases like anchoring, conformity and the bandwagon effect are working through in the predictions of the Transfermarkt users and harm the required independence. Besides that, the nationality of the Transfermarkt users is skewed because of Transfermarkt's German roots; this detracts diversity of opinion and decentralization. According to Surowiecki, the Transfermarkt valuators need to meet these three criteria, in order to extract optimal crowd predictions. Even though this is not the case, the crowd is still able to outperform simple and compete with sophisticated prediction methods. Therefore the crowd could indeed be considered 'wise', but in the end it is wiser to rely on aggregated betting odds when predicting sports matches.

All used figures are verifiable. From Transfermarkt team values, match statistics, ELO ratings to betting odds: all can be retrieved by visiting the indicated sources. The ideology behind the wisdom of crowds, is that a group's collective knowledge potentially could be the foundation of relatively good judgments, decisions and thus predictions. In this thesis the player valuations of the registered users of Transfermarkt are used to create a prediction model; they can therefore be seen as the crowd's predictions, hence the wisdom of the crowd.

Three major European football competitions are used in this thesis´ analysis: the Premier League, the Bundesliga and the Primera División. By including three different national leagues, the influence of

demographical and cultural factors is mitigated. Besides, because of the fact that these competitions are among the largest of the world, they are - for Transfermarkt users - the most interesting to analyze. As a result, many player valuations are provided in those three analyzed competitions, which strengthens the validity of the aggregated Transfermarkt player- and team values in these competitions. All of this should contribute to the generalizability of the results in this thesis.

In accordance with results in existing literature, I initially expected the WOC model to outperform at least all naïve benchmark models (pure chance and 'home team wins'). Because of the potential danger of cognitive biases on the independence of the player valuations on Transfermarkt, I however assumed that the models based on ELO ratings and aggregated betting odds, models that are better able to include the most recent (updated) information, would be superior.

Eventually this appears to be the case for the aggregated betting odds model, but not for the model based on ELO ratings. The low amount of team value updates and the - mainly driven by harmful cognitive biases - violation of Surowiecki's four criteria, cannot undo the strong prediction accuracy of the WOC model, compared to the prediction method based on ELO ratings. On the other hand, aggregated betting odds clearly form the best basis for a prediction model, as it proves to be significantly better than all other forecasting methods.

The WOC model in this thesis is deliberately kept much more 'simple' than the majority of the models in the existing literature. Despite this simplicity, it is still able to outperform other (and compete with even more sophisticated) prediction models. The time- and therefore cost-saving nature makes the model interesting to employ when predicting.

5.2 Limitations

Violation of the four criteria of Surowiecki causes the eventual 'wise' effect of the crowd to be reduced. This does not directly pose a threat to the conclusions in this thesis, but due to the fact that examining the exact impact of cognitive biases is beyond the scope of this study, the true WOC effect (i.e. no violation of Surowiecki's criteria) can only be approached. Another more technical limitation, is the fact that Transfermarkt in the past has been susceptible to manipulation. Real transfer fees are increasingly established by looking at suggestive crowd's valuations; that's why players and their agents benefit from higher valuations. Transfermarkt is free accessible; this offers manipulation opportunities. Especially relatively unknown players from 'corrupt' countries are often the subject matter. Even though Transfermarkt watches sudden player value changes closely, it is not unimaginable that player valuations can still be affected by this phenomenon (Zeit Online, 2009).

5.3 Recommendations

To find out the true prediction accuracy of the Transfermarkt valuations, it has to be made sure that Surowiecki's four criteria are met. Because it is highly likely that this is not the case in this thesis, further research should try and focus on examining the impact of cognitive biases on the valuations of Transfermarkt users. Detecting the magnitude of these confounders should enable us to demonstrate the true WOC effect; find out whether the crowd's wisdom - under perfect conditions - is able to compete with upper benchmarks like aggregated betting odds when predicting.

References

Armstrong, J. S. (2005). The forecasting canon: nine generalizations to improve forecast accuracy. *International Journal of Applied Forecasting, 1*, 29-35.

Asch, S. (1951). Effects of group pressure on the modification and distortion of judgments. *Groups, leadership and men*, 222-236.

Australian Institute of Family Studies. (2014, November). *''The growth of sports betting in Australia''*. Retrieved June 8, 2016, from Australian Institute of Family Studies: https://aifs.gov.au/agrc/publications/sports-betting-and-advertising/growth-sports-betting-australia

Babad, E. (1987). Wishful thinking and objectivity among sports fans. *Social Behaviour, Vol 2(4)*, 231-240.

Babad, E., Hills, M., & O'Driscoll, M. (1992). Factors influencing wishful thinking and predictions of election outcomes. *Basic and applied social psychology, 13(4)*, 461-476.

Bar-Hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking & Reasoning, 1(1)*, 71-103.

Bastardi, A. U., & Ross, L. (2011). Wishful Thinking Belief, Desire, and the Motivated Evaluation of Scientific Evidence. *Psychological Science, 22(6)*, 731-732.

Baur, D. G., & Lehmann, S. (2007, April). *"Does the mobility of football players influence the success of the national team?."*. Retrieved June 17, 2016, from Social Science Research Network: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=980936

Biddle, J. (1991). A bandwagon effect in personalized license plates? *Economic Inquiry, 29*, 375–388.

Bless, H., Fiedler, K., & Strack, F. (2004). *Social cognition: How individuals construct social reality.* New York: Psychology Press.

Bryson, A., Frick, B., & Simmons, R. (2013). The returns to scarce talent footedness and player remuneration in European soccer. *Journal of Sports Economics, 14(6)*, 606-628.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes, 79(2)*, 115-153.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annu. Rev. Psychol., 55*, 591-621.

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: a literature review. *Journal of Sport & Exercise Psychology, 14(1)*, 13-27.

Das, T., & Teng, B.-S. (1999). ''Cognitive Biases and Strategic Decision Processes: An Integrative Perspective.''. *Journal of Management Studies, 36*, 757–778.

Davis-Stober, C. P., Budescu, D. V., Dana, J., & Broomell, S. B. (2015). ''The Composition of Optimally Wise Crowds''. *Decision Analysis 12(3)*, 130-143.

Deloitte. (2016). *Top of the table: Football Money League.* Manchester: Sports Business Group.

Eysenck, H. J. (1939). The validity of judgments as a function of the number of judges. *Journal of Experimental Psychology, Vol 25(6)*, 650-654.

Forrest, D., Goddard, J., & Simmons, R. (2005). Odds-setters as forecasters: The case of English football. *International journal of forecasting, 21(3)*, 551-564.

Foutz, N. Z., & Jank, W. (2007, May 7). *"The wisdom of crowds: pre-release forecasting via functional shape analysis of the online virtual stock market.".* Retrieved June 11, 2016, from SSRN: http://ssrn.com/abstract=1432444

Franch, F. (2013). "(Wisdom of the Crowds) 2: 2010 UK election prediction with social media.". *Journal of Information Technology & Politics 10.1*, 57-71.

Frick, B. (2006). "The football players' labour market: emperical evidence from the major European leagues.''. *Scottish Journal of Political Economy 54.3*, 422-446.

Gaissmaier, W., & Marewski, J. N. (2011). Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. *Judgment and Decision Making 6.1*, 73-88.

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism under social pressure. *Review of Economics and Statistics, 87(2)*, 208-216.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, Vol 109(1)*, 75-90.

Goldstein, D. G., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting, 25(4)*, 760-772.

Harville, D. A., & Smith, M. H. (1994). The Home-Court Advantage: How Large is it, and does it vary from Team to Team? *The American Statistician, 48(1)*, 22-28.

Haselton, M. G., Nettle, D., & Murray, D. R. (2005). *The evolution of cognitive bias.* New York: John Wiley & Sons.

Herm, S., Callsen-Bracker, H. M., & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review, 17(4)*, 484-492.

Herzog, S. M., & Hertwig, R. (2011). The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition. *udgment and Decision Making 6(1)*, 58.

Hong, C. S., & Konrad, K. (1998). Bandwagon effects and two-party majority voting. *Journal of Risk and Uncertainty, 16(2)*, 165-172.

Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of forecasting*, 460-470.

Investopedia. (2016). *''Anchoring''*. Retrieved June 22, 2016, from Investopedia: http://www.investopedia.com/terms/a/anchoring.asp

Jefferys, W. H. (1992). Ockham's razor and Bayesian analysis. *American Scientist, 80(1)*, 64-72.

Jeffrey Young. (2014, May 23). *''Worldwide Sports Betting Draws Big Bucks''*. Retrieved June 13, 2016, from Voice of America: http://www.voanews.com/content/worldwide-sports-betting-draws-big-bucks/1921071.html

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives, 5(1)*, 193-206.

Krizan, Z., Miller, J. C., & Johar, O. (2010). Wishful thinking in the 2008 US presidential election. *Psychological Science, 21(1)*, 140-146.

Krueger, J. I., & Massey, A. L. (2009). "A rational reconstruction of misbehavior.". *Social Cognition 27.5*, 786.

Lane, D. M. (n.d.). *One- and Two-Tailed Tests*. Retrieved August 7, 2016, from Online Statistics Education: An Interactive Multimedia Course of Study: http://onlinestatbook.com/2/logic_of_hypothesis_testing/tails.html

Larrick, R. P., & Soll, J. B. (2006). Intuitions about combining opinions: Misappreciation of the averaging principle. *Management science, 52(1)*, 111-127.

Larrick, R. P., Mannes, A. E., Soll, J. B., & Krueger, J. I. (2011). *The social psychology of the wisdom of crowds.* New York: Psychology Press.

Long, D. S., Fox, N. R., & York, T. (2007). *Calculated futures: theology, ethics, and economics.* Waco: Baylor University Press.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). "How social influence can undermine the wisdom of crowd effect.". *Proceedings of the National Academy of Sciences 108.22*, 9020-9025.

Marsh, C. (1985). Back on the Bandwagon: The Effect of Opinion Polls on Public Opinion. *British Journal of Political Science*, 51-74.

Massey, C., & Simmons, J. P. (2011). Hope Over Experience Desirability and the Persistence of Optimism. *Psychological Science, 22(2)*, 274-281.

Mayraz, G. (2011, October 25). *Wishful thinking.* Retrieved June 25, 2016, from Social Science Research Network: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1955644

McAllister, I., & Studlar, D. T. (1991). Bandwagon, underdog, or projection? Opinion polls and electoral choice in Britain, 1979-1987. *Journal of Politics, 53(3)*, 720-741.

Mehrabian, L. (1998). Effects of poll reports on voter preferences. *Journal of Applied Social Psychology, 28(23)*, 2119-2130.

Murr, A. E. (2011). ""Wisdom of crowds"? A decentralised election forecasting model that uses citizens' local expectations.". *Electoral Studies 30.4*, 771-783.

Mussweiler, T., & Strack, F. (2000). Numeric judgments under uncertainty: The role of knowledge in anchoring. *Journal of experimental social psychology, 36(5)*, 495-518.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review, 4(1)*, 79-95.

Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport. *Sports Medicine, 28(4)*, 221-236.

Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica, 125(1)*, 99-116.

Pavlović, V., Milačić, S., & Ljumović, I. (2014). Controversies about the Accounting Treatment of Transfer Fee in the Football Industry. *Management Journal for Theory and Practice Management*, 70.

Peeters, T. (2016, May). *Erasmus University Rotterdam.* Retrieved July 17, 2016, from Erasmus University Rotterdam: http://papers.sin-online.nl/190627x6e38a985117af6be6aff213b8d6195c0/Wisdom%20of%20crowds%200520 16.pdf

Price, P. C. (2000). Wishful thinking in the prediction of competitive outcomes. *Thinking & Reasoning, 6(2)*, 161-172.

Ray, R. (2006). "Prediction Markets and the Financial" Wisdom of Crowds".". *The Journal of Behavioral Finance 7.1*, 2-4.

Schiefler, L. (2015, August 8). *The system*. Retrieved May 27, 2016, from Football Club ELO Ratings: http://clubelo.com/System

Schiefler, L. (2016). *Football Club Elo Ratings*. Retrieved May 27, 2016, from Football Club Elo Ratings: http://clubelo.com/

Schwenk, C. R. (1988). ''The cognitive perspective on strategic decision making.''. *Journal of Management Studies, 25*, 41–55.

Seybert, N., & Bloomfield, R. (2009). Contagion of wishful thinking in markets. *Management Science, 55(5)*, 738-751.

Simmons, J. P., Nelson, L. D., Galak, J., & Frederick, S. (2011). ''Intuitive biases in choice versus estimation: Implications for the wisdom of crowds.''. *Journal of Consumer Research, 38(1)*, 1-15.

Sinha, S., Dyer, C., Gimpel, K., & Smith, N. A. (2013, October 25). *Predicting the NFL Using Twitter.* Retrieved June 15, 2016, from Cornell University: http://arxiv.org/pdf/1310.6998v1.pdf

Smithson, M. (2010). When less is more in the recognition heuristic. *Judgment and Decision Making, Vol. 5, No. 4*, 230–243.

Spann, M., & Skiera, B. (2009). Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting, 28(1)*, 55-72.

Stat Trek. (2016). *Statistics and Probability Dictionary*. Retrieved July 3, 2016, from Stat Trek: http://stattrek.com/statistics/dictionary.aspx?definition=coefficient_of_determination

Strack, F., & Mussweiler, T. (1997). "Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility.". *Journal of Personality and Social Psychology 73 (3)*, 437–446.

Straffin Jr, P. D. (1977). "The bandwagon curve.". *American Journal of Political Science*, 695-709.

Štrumbelj, E. (2014). On determining probability forecasts from betting odds. *International journal of forecasting*, *30*(4), 934-943.

Surowiecki, J. (. (2005). *The wisdom of crowds.* Anchor.

The Daily Mail. (2015, April 15). *''Global sports gambling worth 'up to $3 trillion'''*. Retrieved June 13, 2016, from The Daily Mail: http://www.dailymail.co.uk/wires/afp/article-3040540/Global-sports-gambling-worth-3-trillion.html

Torgler, B., & Schmidt, S. L. (2007). What shapes player performance in soccer? Empirical findings from a panel analysis. *Applied Economics, 39(18)*, 2355-2369.

Transfermarkt. (2016). *Football players and team values*. Retrieved June 21, 2016, from Transfermarkt: http://www.transfermarkt.de/

Transfermarkt. (2016). *Market Value Analysis*. Retrieved July 15, 2016, from Transfermarkt: http://www.transfermarkt.de/marktwertanalyse/detail/forum/67

Tunaru, R. S., & Viney, H. P. (2010). "Valuations of Soccer Players from Statistical Performance Data.". *Journal of Quantitative Analysis in Sports 6.2* , Article 10.

Tversky, A., & Kahneman, D. (1974). ''Judgment under uncertainty: Heuristics and biases.''. *science, 185(4157)*, 1124-1131.

UEFA. (2016, June 1). *UEFA rankings for club competitions*. Retrieved June 28, 2016, from UEFA: http://www.uefa.com/memberassociations/uefarankings/country/index.html

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

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). "A new look at anchoring effects: Basic anchoring and its antecedents.". *Journal of Experimental Psychology: General 125 (4)*, 387–402.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *The Journal of Economic Perspectives, 18(2)*, 107-126.

Zeit Online. (2009, December 10). *Manipulierter Marktwert*. Retrieved July 17, 2016, from Zeit
    Online: http://www.zeit.de/sport/fussball/2009-12/transfermarkt-spieler-
    marktwert/komplettansicht

Appendix

Calculations home team value multiplier calibration model

To find out which value the home team multiplier in the WOC model has to take, 31 values are plotted in a graph (2.00, 2.01, 2.02 … 2.28, 2.29, 2.30). These 31 values are taken into account when plotting a trend line; the top of this trend line will reflect the home team value multiplier in the model I compose. The trend line is set up by Microsoft Excel function 'trend line', which gave the following equation:

$y = -0,1568x^2 + 0,6754x – 0,1987$

To calculate the x-value (home team multiplier value) of the top of the trend line, the following formula is used:

-b / 2a   (whereas a = -0.1568 and b = 0.6754)

Which results in the following home team multiplier value:

= -0,6754 / (2 * -0,1568) = <u>2,154</u> (2.15)