University of Zurich UZH

Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 390

Underestimating randomness: Outcome bias in betting exchange markets

Oliver Merz, Raphael Flepp and Egon Franck

August 2021

http://www.business.uzh.ch/forschung/wps.html

UZH Business Working Paper Series

Contact Details

Oliver Merz

oliver.merz@business.uzh.ch

Raphael Flepp

raphael.flepp@business.uzh.ch

Egon Franck

egon.franck@business.uzh.ch

University of Zurich

Department of Business Administration

Plattenstrasse 14, CH-8032 Zurich, Switzerland

Underestimating randomness: Outcome bias in betting exchange markets

Oliver Merz, Raphael Flepp, and Egon Franck*

August 2021

*Corresponding author: Oliver Merz, Plattenstrasse 14, CH-8032 Zurich, Switzerland. oliver.merz@business.uzh.ch. Coauthors: Dr. Raphael Flepp, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland, and Prof. Dr. Egon Franck, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland.

Abstract

This paper examines whether the outcome bias harms price efficiency in betting exchange markets. In soccer, the match outcome is an unreliable performance measure, as it underestimates the high level of randomness involved in the sport. If bettors overestimate the importance of past match outcomes and underestimate the influence of good or bad luck, we expect less accurate prices for lucky and unlucky teams. Analyzing over 8,900 soccer matches, we find evidence that the prices are overstated for previously lucky teams and understated for previously unlucky teams. Consistent with the outcome bias, the betting community overestimates the importance of past match outcomes. Consequently, this bias translates into significantly negative betting returns on lucky teams and positive betting returns on unlucky teams. Based on this finding, we propose a simple betting strategy that generates positive returns in an out-of-sample backtest.

Keywords

Behavioral biases · Market efficiency · Forecasting · Betting industry · Soccer

JEL Classification

D40 · G40 · L83

1. Introduction

Suppose a coworker tells you that he earned more money in the stock market than in his job during the last year. Despite having no profound knowledge of investing, you decide to put your savings into a few stocks that have recently caught your attention in the media. As it turns out, you triple the size of your investment within the next couple of months. You have made an excellent decision, haven't you?

People tend to associate good decisions with good outcomes and bad decisions with bad outcomes. However, in a world of uncertainty, a good decision can lead to a bad outcome, and a bad decision can lead to a good outcome (Hershey & Baron, 1992). People often ignore or underestimate the causal role of external, random or extraneous factors that influence outcomes (Allison, Mackie & Messick, 1996). The outcome bias is present whenever individuals tend to assign too much importance to the outcomes when evaluating past decisions (see, e.g., Baron & Hershey, 1988). Many laboratory studies have found an outcome bias in different settings, e.g., legal decisions (Alicke, Davis & Pezzo, 1994), medical decisions (Baron & Hershey, 1988), investment decisions (König-Kersting, Pollmann, Potters, & Trautmann, 2021; Ratner & Herbst, 2005) and ethical judgments (Gino, Moore & Bazerman, 2009; Gino, Shu & Bazerman, 2010). Consistent with the literature of psychology and economics, recent research in sports economics has also shown that decision-makers underestimate the role of randomness in match outcomes and assign too much weight to the observed outcomes when they evaluate performance (Flepp & Franck, 2021; Gauriot & Page, 2019; Kausel, Ventura & Rodríguez, 2019; Lefgren, Platt & Price, 2015).

While these studies have demonstrated that the outcome bias is present in various settings both in the laboratory and in the field, the main focus has been on individual rather than collective decision-making. However, a group of people is often better than an individual at solving cognitive problems and producing accurate predictions, even if the individual is an expert on the subject (e.g., Sjöberg, 2009; Surowiecki, 2004). This phenomenon is commonly referred to as the "wisdom of the crowds" (Frey & Van De Rijt, 2020; Galton, 1907; Surowiecki, 2004) and has been demonstrated in a variety of settings, including weight estimations of oxen (Galton, 1907), political forecasts (e.g., Murr, 2015; Sjöberg, 2009) and financial forecasts (e.g., Kelley & Tetlock, 2013; Nofer & Hinz, 2014). Most of previous researchers have found that the crowd becomes wiser as the number, ability, and diversity of its members increase (Keuschnigg & Ganser, 2017) and that the high accuracy of collective judgments can be explained because individual biases cancel each other out in an aggregated context (e.g., Hong & Page, 2004; Keuschnigg & Ganser, 2017; Larrick & Soll, 2006). However, the findings of Simmons, Nelson, Galak and Frederick (2011) indicated that crowds can also become unwise if their members are systematically biased. Thus, based on the extensive evidence on the outcome bias at an individual level, the question arises of whether the outcome bias is persistent enough to exist in large crowds or whether it is canceled out by the "wisdom of the crowds" mechanism.

We address this research gap by analyzing whether the outcome bias exists in a betting exchange market setting. In betting exchange (prediction) markets, the participants trade with each other on the outcome of future events, such as political elections or sports (Brown, Reade & Vaughan Williams, 2019). As the prices of these bets aggregate the dispersed information of numerous independent, well-informed and financially incentivized individuals, betting exchange markets are an ideal setting for the "wisdom of the crowds" mechanism. Consequently, we would observe distorted prices only if the betting community is systematically outcome biased.

Specifically, we test whether the outcome bias influences the betting market prices of soccer matches within the top five leagues in Europe. In soccer, the match outcome is not necessarily a reliable performance indicator because soccer is a low-scoring sport where randomness plays an important role and winning or losing is often determined by a single goal (Brechot & Flepp, 2020). We analyze whether the participants in betting exchange markets adequately consider the role of randomness in soccer matches or whether they might be prone to the outcome bias by over-relying on past match outcomes and simultaneously underestimating the role played by good or bad luck. Thus, to test whether bettors attribute too much weight to the actual outcome of a match instead of the true performance of the teams, we must consider a more reliable performance measure. Brechot and Flepp (2020) demonstrated that a model based on expected goals, i.e., the quantification of goal scoring opportunities, better measures performance and better predicts future match outcomes than considering actual past match outcomes. Based on this model, we can form variables that measure the good or bad luck of soccer teams and then determine whether good and bad luck are correctly reflected in the betting market prices.

We use data from over 8,900 soccer matches between 2013 and 2018. Soccer data are obtained from Gracenote, a subsidiary of Nielsen Holdings Plc., and odds data are obtained from the betting exchange Matchbook via www.oddsportal.com. Following Brechot and Flepp (2020), we estimate an expected goals model to derive the lagged table difference (*LTD*) between the official league table (OLT) rank and the rank in an alternative table based on expected goals (xGT). Based on the findings of Brechot and Flepp (2020), the xGT should cancel out good and bad luck by measuring performance more accurately. Thus, *LTD* denotes the amount of good or bad luck a team has had. A positive value for *LTD* implies that the rank in the OLT was lower than that in the xGT; thus, the team is considered "unlucky". Conversely, a negative value for *LTD* implies that the team is considered "lucky". Based on the *LTD* variable, we construct the binary variables *Goodluck*, *Badluck* and *Neutral*, denoting lucky, unlucky and neutral teams to better distinguish the impact of good and bad luck on betting market price efficiency.

Following previous research, e.g., Brown, Rambaccussing, Reade and Rossi (2018), Forrest and Simmons (2008), and Franck, Verbeek and Nüesch (2011), we use a binary probability model with the outcome of a bet as the dependent variable (equaling 1 if the bet is won and 0 if lost). As the explanatory variables, we use the winning probability implied in the betting odds, i.e., the reciprocal value of the odds, a dummy variable for home teams to exclude a potential home team bias (see, e.g., Forrest & Simmons, 2008), and our variable of interest, *LTD*, or *Goodluck* and *Badluck*. If the odds are efficient, all relevant information should be reflected in them, and no additional variables should have predictive power regarding the outcome of an event. In other words, if the bettors are not deceived by the outcome bias and instead correctly assess the randomness component in soccer, the *LTD* variable and the derived variables *Goodluck* and *Badluck* should not have explanatory power beyond the implied winning probabilities. However, if the bettors are prone to overweight the importance of past match outcomes, the information contained in *LTD, Goodluck* and *Badluck* might not be correctly reflected in the betting market prices.

We find that *LTD* has a positive and significant impact on the match outcome while controlling for implied winning probabilities. Thus, betting market prices are not entirely efficient. Furthermore, our results stemming from the regression using *Goodluck* and *Badluck* show that the prices of bets on previously lucky teams are overstated. Conversely, the prices of previously unlucky teams are understated. This finding is mirrored in consistently negative returns for bets on previously lucky teams and consistently positive returns for bets on previously unlucky teams. Thus, we form a simple betting strategy by betting on unlucky teams and betting against lucky teams. An out-of-sample backtest of this strategy yields a return of 11.8% before commission and transaction costs and 4.2% after these costs are deducted. This strongly indicates that participants in betting markets do not correctly regard the randomness component in soccer when making betting decisions. Rather, they overweight the importance of past match outcomes and exhibit behavior consistent with the outcome bias.

This paper contributes to the literature in various ways. First, we extend the previous literature of psychology, economics and sports economics by demonstrating that the outcome bias is not limited to individuals but also exists in large crowds. Interestingly, the "wisdom of the crowds" mechanism is unable to counteract outcome-biased individuals in a setting where people are highly incentivized to make optimal decisions. Second, we contribute to the literature on betting market efficiency by demonstrating that the outcome bias harms the forecasting power of betting prices. Finally, our findings imply that performance evaluations in soccer, and potentially also in other sports, might be fundamentally outcome biased. Consequently, the outcome bias might not be limited to betting markets but might also lead to inefficiencies in other large-scale areas where random factors influence the outcome.

The remainder of this paper is structured as follows. In section 2, we summarize the relevant literature and state our hypotheses. In section 3, we describe the data, variables and empirical methodology used. In section 4, we present our results and an out-of-sample betting strategy. Section 5 concludes the paper.

2. Related literature and hypotheses

2.1 Outcome bias

Research from psychology has shown that people often struggle to make decisions under uncertainty and exhibit various cognitive biases that distort not only decision-making but also the evaluation of past decisions (see, e.g., Earl, 1990; Rabin, 1998; Tversky & Kahneman, 1974). For our setting, the outcome bias is of great interest. The outcome bias refers to a phenomenon of people overweighting the importance of the outcome when evaluating a past decision (Baron & Hershey, 1988). When judging the quality of a past decision, an objective evaluator should consider all the information known to the decision-maker at the time of the decision to assess whether the decision was optimal. However, the evaluation of decision quality should not depend on randomly determined outcomes (Bazerman & Moore, 2012; Hastie & Dawes, 2009). Outcomes should be considered in the evaluation process only if they provide additional information about intentionality, culpability, or characteristics of the actor's personality (Hershey & Baron, 1992; Mazzocco, Alicke & Davis, 2004). However, when the outcome information lacks any additional information about the actor, the quality of the decision should not be judged differently depending on the outcome (Gino et al., 2010). Even if the outcome is clearly caused by external, random or extraneous factors, people tend to believe that good decisions lead to good outcomes and ignore or underestimate the causal role of such contextual factors (Allison et al., 1996). Baron and Hershey (1988) discovered that students are prone to be outcome biased when evaluating medical procedures. The students assessed decisions as more appropriate when the outcome was successful than when it was unsuccessful despite all other information being equal. After this study (Baron & Hershey, 1988), the existence of the outcome bias was demonstrated by many subsequent studies in laboratory settings (e.g., Brownback & Kuhn, 2019; Cushman, Dreber, Wang & Costa, 2009; Gurdal, Miller & Rustichini, 2013; König-Kersting et al., 2021; Marshall & Mowen, 1993; Mazzocco et al., 2004; Mowen & Stone, 1992; Rubin & Sheremeta, 2016). The outcome bias has also been found in various contexts. For instance, Gino et al. (2009, 2010) found that ethically questionable behavior is perceived as more unethical when it produces negative outcomes than when it produces positive outcomes. Consistent results regarding the outcome bias have also been found in legal contexts (e.g., Alicke et al., 1994; Mazzocco et al., 2004), salespeople's performance evaluations (Marshall & Mowen, 1993), and investment decisions (König-Kersting et al., 2021; Ratner & Herbst, 2005).

Consistent with insights from the laboratory, a few studies have provided the first evidence of the outcome bias in the field (Emerson et al., 2010; Tinsley, Dillon & Cronin, 2012). Recently, research in sports economics has shown that decision-makers underestimate the role of randomness in match outcomes and assign too much weight to the observed outcomes when they evaluate performance (Flepp

& Franck, 2021; Gauriot & Page, 2019; Kausel et al., 2019; Lefgren et al., 2015). Using a formal model based on Bayesian updating, Lefgren et al. (2015) found that coaches in the NBA change their strategy more often after losing a game than after winning a game, even when comparing narrow wins and losses. Furthermore, coaches react equally to expected and unexpected performance and revise their strategies independent of whether a loss is based on factors outside their control. These findings are consistent with the outcome bias literature. Gauriot and Page (2019) investigated whether the outcome bias is present in performance evaluations of soccer players. They found that players' shots on the post that resulted in a goal overly influenced the players' evaluations compared to shots that hit the post and missed. Players who scored in a match after hitting the post not only were rewarded by their manager with more play time in upcoming matches but also were rated higher by journalists and sports fans. Similarly, Kausel et al. (2019) showed that journalists rated players significantly better than their opponents when their team won the penalty shootout despite comparable in-game performance. These findings indicate that luck is overly rewarded in such performance evaluations and consequently might lead to inappropriate future strategy adjustments. For instance, Flepp and Franck (2021) showed that coach dismissals in soccer lead to a boost in subsequent team performance only when previous match outcomes were the result of actual bad performance rather than bad luck. This finding emphasizes the importance of unbiased performance and decision-making evaluations, as coach dismissals are very costly at the top level of soccer.

Overall, the literature shows that decision-makers have difficulty evaluating the quality of decisions or performance when random events also influence the final outcome. A favorable outcome often justifies a decision, strategy or performance, even if there is other evidence indicating otherwise. While it is unsurprising that coaches, journalists and fans focus on the outcomes of matches, which are of critical importance, this is not necessarily the best approach to make optimal decisions in the future. These studies have demonstrated that individual decision-makers are outcome biased in a variety of different settings, both in the laboratory and in the field. In this paper, we extend the previous literature by analyzing whether these individual biases are also present on a larger scale. To do so, we analyze price efficiency in a betting exchange market environment where prices reflect the aggregated beliefs of a crowd consisting of numerous individuals.

2.2 Research hypotheses

As the participants in betting exchange markets trade with each other on the outcomes of future events, i.e., soccer match outcomes in our setting, the prices reflect the aggregated beliefs of the participants (Brown et al., 2019). If prices reflect all relevant information and are the best forecasts of the match outcomes, they are considered efficient (e.g., Angelini & De Angelis, 2019). Prediction market research has demonstrated that these kinds of markets are indeed very efficient (see, e.g., Berg, Nelson & Rietz, 2008; Rothschild, 2015; Spann & Skiera, 2009; Vaughan Williams & Reade, 2016; Wolfers & Zitzewitz, 2004).

Thus, if the bettors in our setting are not outcome biased in the aggregate, we expect the betting markets to be efficient and the prices to correctly reflect all available information, including the randomness component involved in soccer match outcomes. Neither good luck nor bad luck, measured by our variables of interest, is expected to have additional predictive power toward the match outcome. The null hypothesis is stated as follows:

> *H0: If bettors are not outcome biased in the aggregate, the betting market prices will accurately incorporate the influence of good and bad luck on previous match outcomes.*

However, if the bettors are outcome biased in the aggregate and underestimate the role played by randomness in soccer, we expect that in addition to the betting prices, good and bad luck will have explanatory power. We assume that the unluckier a team has been, the more understated the price of a bet on this team will be. Thus, hypothesis H1 is stated as follows:

> *H1: If bettors are outcome biased in the aggregate, the unluckier a team previously was, the more underestimated the winning probability is by the betting prices.*

More specifically, we expect the prices for bets on previously lucky teams to be overstated, as bettors overly attribute prior match outcomes to good performance while underestimating the influence of good luck. Conversely, we expect the prices for bets on previously unlucky teams to be understated, as bettors overly attribute prior match outcomes to bad performance while underestimating the influence of bad luck. Thus, we hypothesize the following:

H2a: If bettors are outcome biased in the aggregate, the betting prices will overestimate the winning probabilities of previously lucky teams.

H2b: If bettors are outcome biased in the aggregate, the betting prices will underestimate the winning probabilities of previously unlucky teams.

3. Methods

3.1 Data

We obtained soccer data from Gracenote, a subsidiary of Nielsen Holdings Plc. Among other things, Gracenote provides sports metadata for a variety of different sports. Our main dataset contains shot information for 9,130 matches from the top five European soccer leagues for the 2013/2014 through the 2017/2018 seasons. More specifically, we have data on the goals and shots of 1,530 matches from the German Bundesliga and 1,900 matches each from the French Ligue, the Italian Serie A, the English Premier League and the Spanish La Liga. For each shot, we know the exact location, the rule setting and the part of the body that was used. Additionally, we collected data on 1,826 matches from the same leagues for the 2018/2019 season to test the robustness of the results and to test an out-of-sample betting strategy. The odds data stem from Matchbook and were collected from www.oddsportal.com. Unfortunately, some odds were not recorded on www.oddsportal.com due to technical problems. Thus, we have odds data for only 8,965 of the 9,130 matches in the main dataset.

3.2 Expected goals model

Brechot and Flepp (2020) show that the expected goals metric contains more relevant information about future team performance than match outcomes do, as this metric is less prone to randomness. We follow Brechot and Flepp (2020) in estimating the scoring probabilities of shots based on the distance, angle, rule setting of the shot (i.e., open play, free kick or penalty kick), and body part used. Additionally, we include team fixed effects in the logistic regression to account for unobserved team quality characteristics, such as defensive or goal scoring skills. In total, we estimate the scoring probability of 214,194 shots from all 9,130 matches in our sample. Finally, we aggregate the quantified scoring chances for each team within each match to derive the number of expected goals per match. For instance, if Manchester United had 5 shots against Leicester City with expected scoring probabilities of 0.80, 0.70, 0.50 0.25, and 0.15, the expected goals for Manchester United would be 2.40. For each match, we calculate the expected goals for both teams. Instead of awarding points based on the actual match outcome, i.e., three for a win, one for a draw and zero for a loss, we award the difference in the expected goals to each team. Therefore, in the abovementioned example, if Manchester United's expected goals were 2.40 and Leicester City's were 0.80, Manchester United would receive the expected goal difference (xGD) of 2.40–0.80=1.60, and Leicester City would receive the xGD of -1.60. Thus, the expected goal metric allows us to determine the better team on the pitch in terms of creating valuable scoring chances. Finally, we rank the teams according to their points based on expected goals and construct an expected goal league table (xGT). The rank in the xGT should reflect a team's playing quality on the pitch more accurately than the rank in the OLT because the OLT is based solely on actual match outcomes, in which bad luck fully translates into fewer points and a lower rank. For example, Manchester United could play well on the pitch and win a match in terms of expected goals because it created scoring chances of higher total value than its opponent. However, Manchester United might actually lose the match in terms of outcomes because the scoring chances did not translate into actual goals. In such situations, the rank in the xGT should be higher than the rank in the OLT. Conversely, in situations where the team played poorly on the pitch but won the match, the rank in the xGT should be worse than the rank in the OLT.

3.3 Variables of interest

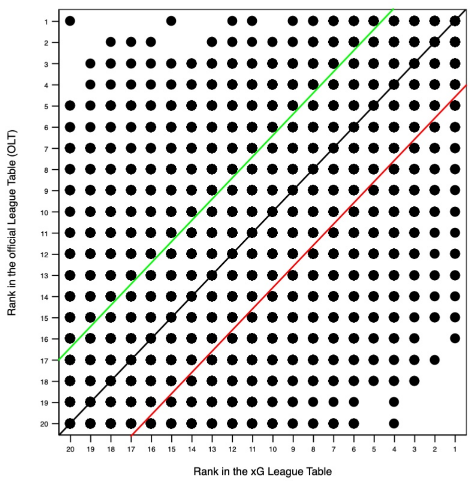
We use the difference between rank in the OLT and rank in the xGT to test whether betting market prices correctly include good and bad luck. Specifically, we define our variable of interest as follows:

$$LTD_{i,j,k} = Rank\ OLT_{i,j-1,k} - Rank\ xGT_{i,j-1,k}$$

where i denotes a team, j denotes the match week and k denotes the season. Thus, a positive value for *LTD* indicates that a team has been unlucky, as its rank is worse in the OLT than in the xGT. Conversely, a negative value for *LTD* indicates that the team has been lucky, as its rank is better in the OLT than in the xGT.

Following Brechot and Flepp (2020) and Flepp and Franck (2021), we plot the OLT rank against the xGT rank to show the distribution of the rank combinations for all teams in all but the first match weeks of the seasons (see Figure 1).[1] For all rank combinations above the black line in Figure 1, the *LTD* variable is negative, and for all rank combinations below the black line, the *LTD* variable is positive.

Based on the *LTD* variable, we form the binary subvariables *Goodluck*, *Badluck* and *Neutral* to further distinguish the impact of previously lucky and unlucky teams on betting market price efficiency. In our specification, we characterize teams as lucky if the value of the *LTD* variable is smaller than -3, i.e., all observations above the green line in Figure 1, and as unlucky if the value is larger than 3, i.e., all observations below the red line. The remaining teams are characterized as neutral, i.e., all observations between the green and red lines.



Using these thresholds, we roughly classify the first quintile of *LTD* values as lucky, the last quintile as unlucky and everything in between as neutral. More specifically, we classify approximately 16% of teams as lucky and 16% of teams as unlucky. In section 4.2, we use different thresholds, such as -1/1, -2/2, -4/4 and -5/5, to test the robustness of the proposed threshold values.

[1] We do not display the rank combinations of the first match weeks of the seasons, as they are not used for the calculation of *LTD*. The correlation between the two rankings stays approximately the same independent of whether the first match weeks are included (0.76 versus 0.77).

Formally, the variables are defined as follows:

$$Goodluck_{i,j,k} = \begin{cases} 1 & if\ LTD_{i,j,k} < -3 \\ 0 & otherwise \end{cases}$$

$$Neutral_{i,j,k} = \begin{cases} 1 & if -3 \leq LTD_{i,j,k} \leq 3 \\ 0 & otherwise \end{cases}$$

$$Badluck_{i,j,k} = \begin{cases} 1 & if\ LTD_{i,j,k} > 3 \\ 0 & otherwise \end{cases}$$

3.4 Statistical methods

As in previous research, e.g., Brown et al. (2018), Flepp, Nüesch and Franck (2016), and Forrest and Simmons (2008), we calculate prices as the reciprocal of the odds. The price is the amount of money one must bet in order to collect 1 unit if the bet wins. Thus, the price can also be seen as an implied winning probability. For instance, if the betting odds are 2.0 for a team to win the match, then the price would be $\frac{1}{2.00} = 0.5$, which also denotes the implied winning probability of the team. We calculate the implied winning probabilities $Impliedprob = \frac{1}{odds}$ for each team in all matches.

Betting exchange markets are efficient if the market prices reflect all historical information and the prices are the best forecasts of the outcome of a match (Angelini & De Angelis, 2019). Consequently, no other variable should have explanatory power regarding the match outcome after the implied winning probabilities are controlled for. Thus, following previous research, e.g., Brown et al. (2018), Forrest and Simmons (2008), and Franck et al. (2011), we estimate two logit models as follows:

$$Ln\left[\frac{P(Win_{i,j,k} = 1)}{P(Win_{i,j,k} = 0)}\right] = \beta_0 + \beta_1 Impliedprob_{i,j,k} + \beta_2 Home_{i,j,k} + \beta_3 LTD_{i,j,k} \tag{1}$$

$$Ln\left[\frac{P(Win_{i,j,k} = 1)}{P(Win_{i,j,k} = 0)}\right] = \beta_0 + \beta_1 Impliedprob_{i,j,k} + \beta_2 Home_{i,j,k} + \beta_3 Goodluck_{i,j,k} + \beta_4 Badluck_{i,j,k} \tag{2}$$

where the dependent variable is the actual outcome of a bet (1 for a winning bet and 0 for a losing bet), $Impliedprob_{i,j,k}$ is the probability for each team *i* in match week *j* and season *k* that is implied by the odds, $Home_{i,j,k}$ is a dummy variable controlling for a potential home team bias (see, e.g., Forrest & Simmons, 2008), and $LTD_{i,j,k}$, $Goodluck_{i,j,k}$ and $Badluck_{i,j,k}$ are the variables of interest, measuring whether a team has been lucky or unlucky in the past. Following previous research (see, e.g., Brown et al., 2018; Forrest & Simmons, 2008), we compute clustered heteroscedasticity-robust standard errors at the match level because the independence assumption between observations is violated, as we include multiple observations of the same match (bets on both teams).

Under the null hypothesis *H0*, we expect efficient betting prices and the information of good and bad luck to be fully incorporated into *Impliedprob*. In other words, we expect only the coefficient β_1 to be significant, while we expect the coefficients of *Home* and *LTD* in regression (1) and *Home*, *Goodluck* and *Badluck* in regression (2) to be zero. If bettors are outcome biased on aggregate, we expect betting prices to underestimate the winning probabilities of previously unlucky teams and overestimate the winning probabilities of previously lucky teams. Thus, under *H1,* we expect a positive and significant sign for *LTD*; under *H2a,* we expect a negative and significant sign for *Goodluck*; and under *H2b,* we expect a positive and significant sign for *Badluck* (see section 2.2).

4. Results

4.1 Main results

Table 1 displays summary statistics for the variables *Win*, *Impliedprob*, *Home*, *LTD*, *Goodluck* and *Badluck*. Table 1 shows that the implied winning probabilities very accurately mirror the actual winning probabilities, indicating efficient prices on average. The *LTD* variable has a mean of zero as for each team that is ranked better in the expected goals table at least one other team must be ranked worse.

Table 1: Summary statistics

Variable	N	Mean	SD	Min	Max
Win (0/1)	17,770	0.3771	0.4847	0.0000	1.0000
Impliedprob	17,484	0.3780	0.2023	0.0066	0.9901
Home	17,770	0.5000	0.5000	0.0000	1.000
LTD	17,770	0.0000	3.9554	-19.0000	16.0000
Goodluck	17,770	0.1573	0.3641	0.0000	1.0000
Badluck	17,770	0.1599	0.3666	0.0000	1.0000

Notes: We display summary statistics on a win indicator variable *Win*, the implied winning probability *Impliedprob* and on *LTD*, *Goodluck* and *Badluck* for all observations except the first match week of the season.

The results of the logit regressions are depicted in Table 2. The results are shown in the form of marginal effects measured at a point where the variables are set to their means. In column (1), we depict the results of regression (1) using the *LTD* variable to measure good and bad luck. As expected, the sign of the *Impliedprob* variable is positive and significant at the 1% significance level. The coefficient of the *Home* variable is nonsignificant, indicating that no home team bias is present. Interestingly, the sign of the *LTD* variable is positive and significant at the 1% significance level. This indicates that good and bad luck are not fully incorporated into market prices. More specifically, the coefficient of *LTD* states that a one-unit increase in *LTD* leads to a 0.3 percentage point increase in *Win* taking the value of 1. As positive values of *LTD* indicate that a team was unlucky, column (1) shows that teams that were unluckier in past matches are associated with a higher probability of winning than implied in the prices. In other words, the prices of bets on teams that were unlucky in the past are traded at a discount because they are undervalued by the bettors.

Table 2: Results of logit regressions

	Win (0/1)	
	(1)	(2)
Impliedprob	1.126***	1.127***
	(0.028)	(0.028)
LTD	0.003***	-
	(0.001)	
Goodluck	-	-0.020*
		(0.011)
Badluck	-	0.021*
		(0.011)
Home	0.007	0.007
	(0.011)	(0.011)
Number of observations	17,484	17,484
Number of clusters	8,742	8,742
Log pseudolikelyhood	-9904.2	-9904.7
Pseudo R^2	0.145	0.145

Notes: The table reports the logit estimates for *Win*; *Win* takes the value of 1 if the team won the game and 0 otherwise. Column (1) shows the results of the *LTD* variable, and column (2) shows the results of the variables *Goodluck* and *Badluck*. *Home* controls for teams that play at home. The heteroscedasticity-robust and for the first regression clustered standard errors at the match level are reported in parentheses. ***, **, and * denote significance at the 1-, 5-, and 10-percent levels, respectively.

In column (2) we depict the results of regression (2), where we analyze the impact of good and bad luck separately to better distinguish their individual effects on market prices. Consistent with the first model, *Impliedprob* is positive and significant at the 1% significance level, while the *Home* variable remains nonsignificant. More importantly, the coefficients of *Goodluck* and *Badluck* are both significant after the implied winning probabilities are controlled for. As expected, the sign of *Goodluck* is negative, indicating that bets on teams that were lucky in the past are less favorable, as the implied winning probabilities are too high and the odds are biased downwards. Conversely, the sign of *Badluck* is positive, indicating that bets on teams that were unlucky in the past are more favorable. The marginal effects show that teams that were previously lucky (unlucky) are associated with a 2.0 (2.1) percentage point decrease (increase) in the probability of *Win* taking the value of 1 at the mean values of the variables.

In conclusion, the null hypothesis of market efficiency is rejected, as information about good and bad luck is not fully incorporated into the market prices. Rather, our results support hypotheses *H1*, *H2a* and *H2b* and indicate that teams that were unlucky in the past are systematically undervalued, while teams that were lucky in the past are systematically overvalued by bettors.

4.2 Robustness tests

To test the robustness of the results, we conducted several variations of our main model. The results are displayed in Table 3. First, we conducted probit and standard OLS regressions instead of logit regressions. Columns (1) and (2) in Table 3 show that the results are virtually identical to the logit results. Second, following Forrest and Simmons (2008), we adjusted the implied probabilities for the match outcomes, i.e., home win, away win and draw, so that they sum to one for each match. Column (3) shows that our results are again insensitive to this altercation. Third, we randomly selected a bet on one team per match instead of using clustered standard errors and again obtained consistent results (see column (4) in Table 3). Fourth, we used a different approach to calculate *LTD* by using a variation of the alternative league table that was based on expected goals. Instead of calculating the differences in expected goals and then continuously adding these differences to create a table ranking, we formed thresholds for expected goal differences to determine whether the match outcome was considered a win, draw or loss. Following Flepp and Franck (2021), we considered an expected goal difference greater than 0.5 to be a win, an expected goal difference less than 0.5 to be a loss and everything in between a draw. We then allocated points as in the OLT, i.e., 3 points for a win, 1 point for a draw and 0 points for a loss. One potential upside of this approach is that it uses the same values for points as the OLT. Another advantage is that the maximal number of points for a match is capped by 3 points, whereas individual matches with the xGD approach could more strongly influence the ranking. Thus, the table position might be more comparable to the OLT position. A potential downside is the necessity of determining thresholds where small differences in xGD (around 0.5) can lead to a difference of up to 2 points. However, as displayed in column (5), the results are again very similar in sign, magnitude and significance to those of the main specification. Fifth, as it could be that the table positions are ossified toward the end of the season, we measured good and bad luck based on the last three matches instead of using the table rank differences. Specifically, using the same thresholds for xGD as in the specification of the *LTD* alternative, we classified the past three matches as wins, draws or losses, summed up the points and then subtracted the actual points won. Thus, a positive value for the *Last_three* variable indicates that a team was recently unlucky. Column (6) confirms that, consistent with the results of the main specification, the prices for unluckier teams are also more understated when recent matches are considered.[2] Finally, instead of using -3/3 as *LTD* thresholds to classify teams as lucky and unlucky, we tested a number of alternative threshold values, i.e., -1/1, -2/2, -4/4 and -5/5, to define the *Goodluck* and *Badluck* variables (results are depicted in Table A1 in the appendix). Compared to the results of the main specification, those shown in Table A1 are similar in sign and magnitude but less consistent in significance. For the threshold values of -1/1, -2/2 and -4/4, price inefficiencies seem to be driven by *Goodluck,* while they seem to be driven by *Badluck* for the threshold value of -5/5. Nevertheless, the null hypothesis of market efficiency is rejected for all alternative threshold values.

[2] Similar findings could be achieved if we considered only the last two matches to determine whether a team was lucky or unlucky in recent matches.

Table 3: Robustness tests of main model

	Win (0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Impliedprob	1.113***	1.012***	-	1.152***	1.127***	1.126***
	(0.027)	(0.020)		(0.034)	(0.028)	(0.020)
Impliedprob_adjusted	-	-	1.142***	-	-	-
			(0.028)			
LTD	0.003***	0.003***	0.003***	0.004**	-	-
	(0.001)	(0.001)	(0.001)	(0.001)		
LTD_alternative	-	-	-	-	0.003***	-
					(0.001)	
Last_three	-	-	-	-	-	0.003*
						(0.002)
Home	0.008	0.006	0.007	0.011	0.007	0.006
	(0.010)	(0.009)	(0.011)	(0.001)	(0.011)	(0.009)
Number of observations	17,484	17,484	17,484	8,742	17,484	16,372
Number of clusters	8,742	8,742	8,742	-	8,742	8,186
Log pseudolikelyhood	-9896.4	-	-9899.9	-5019.6	-9904.9	-9269.3
Pseudo R^2/R^2	0.146	0.182	0.145	0.140	0.145	0.146
Model	Probit	OLS	Logit	Logit	Logit	Logit

Notes: The table reports the logit estimates for *Win*; *Win* takes the value of 1 if the team won the game and 0 otherwise. Column (1) displays the results using a probit model, and column (2) displays the results of the standard OLS model. In column (3), we use adjusted implied probabilities, and column (4) shows the results of randomly choosing one team per match. Column (5) shows the results using an alternative approach to calculate our main variable, *LTD*, and column (6) shows the results using only the last three matches to determine good luck/bad luck. The *Home* variable controls for teams that play at home. The heteroscedasticity-robust and for the first regression clustered standard errors at the match level are reported in parentheses. ***, **, and * denote significance at the 1-, 5-, and 10-percent levels, respectively.

4.3 Betting returns comparisons

As the prices for previously unlucky teams are expected to be too low, betting on these teams should yield positive returns. Conversely, as prices for previously lucky teams are too high, betting on those teams should result in negative returns. We calculated betting returns on one-unit bets as follows:

$$return_i = \begin{cases} odds_i - 1 & if \;\; bet \; is \; successful \\ -1 & if \;\; bet \; is \; unsuccessful \end{cases}$$

Table 4 displays the betting returns on teams for various thresholds of the *LTD* variable. Table 4 shows that the returns on bets on previously lucky teams are strictly negative, while the returns on bets on previously unlucky teams are strictly positive. The returns on bets on previously lucky teams show a slight downward trend for stricter *LTD* threshold values, ranging from -6.4% for a threshold of -1 to -7.1% for a threshold of -5. In contrast, the returns on bets on previously unlucky teams increase for stricter *LTD* thresholds, ranging from virtually zero returns for a threshold of 1 to 8.8% for a threshold of 5. This finding indicates that the luckier a team was in the past, the more it is overvalued by bettors, and conversely, the unluckier a team was in the past, the more it is undervalued by bettors. Single-sample *t*-tests show that the returns on lucky teams are significantly smaller than zero for all thresholds, while only the return using a threshold value of 5 is significantly larger than zero for bets on unlucky teams. Nevertheless, consistent with the results of the logit regressions, the returns for previously unlucky teams are consistently higher than the returns for bets on previously lucky teams.

Table 4: Betting returns comparisons

Panel A: Returns for bets on lucky teams				
LTD threshold	N	Mean	SD	t-test (mean<0)
-5	1,379	-0.071	1.751	-1.500*
-4	1,945	-0.068	1.989	-1.509*
-3	2,786	-0.068	1.896	-1.903**
-2	3,973	-0.043	2.106	-1.300*
-1	5,403	-0.064	1.954	-2.401***
Panel B: Returns for bets on unlucky teams				
LTD threshold	N	Mean	SD	t-test (mean>0)
1	5,339	0.001	1.700	0.050
2	3,915	0.024	1.734	0.866
3	2,824	0.039	1.747	1.187
4	1,980	0.047	1.795	1.171
5	1,359	0.088	1.747	1.861**

Notes: The table displays the returns on bets with a stake equaling 1. Panel A shows the returns for bets on lucky teams, and panel B shows the returns for bets on unlucky teams.

4.4 Out-of-sample betting strategy

We used the 2018/2019 season to externally test a betting strategy based on the insights from our main analysis over the 2013/2014–2017/2018 seasons. Using the same dataset and model as in the main analysis, we estimated the expected goals for each team in every match week and specified the *LTD* variable as in the main analysis. Table 5 shows the out-of-sample betting returns for various *LTD* thresholds. As in the main analysis, the returns are negative for teams characterized as lucky and positive for teams characterized as unlucky. There even seems to be a trend that the absolute values of the returns increase depending on whether teams were luckier or unluckier in the past.

Table 5: Out-of-sample betting returns

Panel A: Returns for bets on lucky teams

LTD threshold	N.	Mean	SD
-5	342	-0.128	1.770
-4	470	-0.138	1.677
-3	654	-0.077	1.802
-2	864	-0.040	1.741
-1	1,184	0.010	1.888

Panel B: Returns for bets on unlucky teams

LTD threshold	N	Mean	SD.
1	1,219	0.040	2.095
2	873	0.078	2.158
3	653	0.041	2.091
4	469	0.055	2.039
5	327	0.094	2.208

Notes: The table displays the out-of-sample returns on bets with a stake equaling 1. Panel A shows the returns for bets on lucky teams, and panel B shows the returns for bets on unlucky teams.

On that basis, we can derive simple betting strategies to exploit this finding. The simplest way to profit from this finding is to bet on all teams characterized as unlucky to win (i.e., backing unlucky teams) and to bet against all teams characterized as lucky to win (i.e., laying lucky teams). Table 6 summarizes the returns on this simple strategy if we use the threshold of -3/3, as in our main specification of the variables *Goodluck* and *Badluck*. By strictly backing all teams that were unlucky in the past and laying all teams that were lucky in the past, a positive return of 11.8% could be achieved before commission and transaction costs, and a return of 4.1% could be achieved after the standard commission fee of 4% for winning bets and transaction costs of 2% for lay bets are deducted.[3]

The size of the achieved return is striking, especially considering the simplicity of the deployed betting strategy.

[3] In betting exchange markets, there is a spread between back and lay odds, commonly referred to as transaction costs. As we have no data on lay odds in this dataset, we estimate the spread based on Betfair data from Fracsoft, which do include back and lay odds for matches of the English Premier League in the 2013/2014 season. In that dataset, the spread relative to the back odds is approximately 1.8%. Given some margin of error, we estimate transaction costs of 2% for this sample.

Table 6: Betting strategy returns

Panel A: Returns before commission and transaction costs

	N	Mean	SD	Min	Max
Backing unlucky teams	653	0.041	2.091	-1	21
Laying lucky teams	654	0.077	1.802	-18	1
Combined	**1,405**	**0.118**			

Panel B: Returns after commission and transaction costs

	N	Mean	SD	Min	Max
Backing unlucky teams	653	0.013	2.021	-1	20.16
Laying lucky teams	654	0.029	1.787	-18	0.96
Combined	**1,405**	**0.042**			

Notes: The table displays the returns on back and lay bets with a stake equaling 1. Panel A shows the returns before commission, and panel B shows the returns after a commission rate of 4% on all winning bets and transaction costs of 2% for lay bets have been deducted.

5. Conclusion

Our paper contributes to the literature on the outcome bias by demonstrating that this bias is not limited to individuals but also exists in larger crowds. This finding is rather surprising, as a crowd usually solves cognitive problems better than individuals due to the "wisdom of the crowds" mechanism. We use a betting exchange (prediction) market setting where the prices reflect the collective beliefs of the crowd. We show that the outcome bias leads to less efficient prices, indicating that this bias is widespread in the betting community. More specifically, we demonstrate that bettors do not fully consider the influence of good and bad luck on soccer match outcomes. This is reflected in higher returns on bets on previously unlucky teams and lower returns on bets on previously lucky teams. A simple betting strategy based on this finding leads to a net return of 4.2% in an out-of-sample backtest.

The implications of this paper are twofold. First, our findings imply that prediction markets might not be as efficient in forecasting future events as is commonly assumed in the literature. Cognitive biases cannot dissipate through the "wisdom of the crowds" mechanism if people are not aware of them.

Second, our findings imply that the performance evaluation in soccer, and potentially also in other sports, might be fundamentally outcome biased. Considering the economic importance of soccer, it seems vital to increase awareness of this bias and foster the development of performance measures such as expected goals. Installing more accurate performance measures will not only help people make better-informed betting decisions but also help coaches choose the right players and develop the most promising talents.

As random factors influence the outcomes in various contexts, it seems likely that the outcome bias is not limited to prediction markets but might also exist in other large-scale environments, such as labor markets, politics and financial markets. In such areas, the socioeconomic impact of inefficiencies is certainly substantial. Thus, we encourage future research to further investigate the outcome bias in such areas to help limit this bias.

Appendix

Table A1
Alternative threshold values for good and bad luck.

	Win (0/1)			
	(1)	(2)	(3)	(4)
Impliedprob	1.124***	1.127***	1.126***	1.129***
	(0.028)	(0.028)	(0.028)	(0.028)
Goodluck	-0.022**	-0.018*	-0.028**	-0.022
	(0.010)	(0.010)	(0.013)	(0.015)
Badluck	0.005	0.014	0.018	0.038**
	(0.010)	(0.010)	(0.013)	(0.015)
Home	0.008	0.007	0.007	0.007
	(0.011)	(0.011)	(0.011)	(0.011)
Number of observations	17,484	17,484	17,484	17,484
Number of clusters	8,742	8,742	8,742	8,742
Log pseudolikelihood	-9904.8	-9905.2	-9905.0	-9904.3
Pseudo-R^2	0.145	0.145	0.145	0.145

Notes: The table reports the logit estimates for *Win*; *Win* takes the value of 1 if the team won the game and 0 otherwise. Column (1) shows the results for *Goodluck* and *Badluck* using threshold values of -1/1. In column (2), threshold values of -2/2 are used; in column (3), threshold values of -4/4 are used; and in column (4), threshold values of -5/5 are used. *Home* controls for teams that play at home. The heteroscedasticity-robust and for the first regression clustered standard errors at the match level are reported in parentheses. ***, **, and * denote significance at the 1-, 5-, and 10-percent levels, respectively.

References

Alicke, M. D., Davis, T. L., & Pezzo, M. V. (1994). A posteriori adjustment of a priori decision criteria. *Social Cognition*, *12*(4), 281-308.

Allison, S. T., Mackie, D. M., & Messick, D. M. (1996). Outcome biases in social perception: Implications for dispositional inference, attitude change, stereotyping, and social behavior. In Zanna, M. P. (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 53-93). San Diego, CA: Academic Press.

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

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*(4), 569-579.

Bazerman, M. H., & Moore, D. A. (2012). *Judgment in managerial decision making*. Hoboken, NJ: John Wiley & Sons.

Berg, J. E., Nelson, F. D., & Rietz, T. A. (2008). Prediction market accuracy in the long run. *International Journal of Forecasting*, *24*(2), 285-300.

Brechot, M., & Flepp, R. (2020). Dealing with randomness in match outcomes: how to rethink performance evaluation in european club football using expected goals. *Journal of Sports Economics*, *21*(4), 335-362.

Brown, A., Rambaccussing, D., Reade, J. J., & Rossi, G. (2018). Forecasting with social media: evidence from tweets on soccer matches. *Economic Inquiry*, *56*(3), 1748-1763.

Brown, A., Reade, J. J., & Vaughan Williams, L. (2019). When are prediction market prices most informative?. *International Journal of Forecasting*, *35*(1), 420-428.

Brownback, A., & Kuhn, M. A. (2019). Understanding outcome bias. *Games and Economic Behavior*, *117*, 342-360.

Cushman, F., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a "trembling hand" game. *PloS one*, *4*(8), e6699.

Earl, P. E. (1990). Economics and psychology: a survey. *The Economic Journal*, *100*(402), 718-755.

Emerson, G. B., Warme, W. J., Wolf, F. M., Heckman, J. D., Brand, R. A., & Leopold, S. S. (2010). Testing for the presence of positive-outcome bias in peer review: a randomized controlled trial. *Archives of Internal Medicine*, *170*(21), 1934-1939.

Flepp, R., & Franck, E. (2021). The performance effects of wise and unwise managerial dismissals. *Economic Inquiry*, *59*(1), 186-198.

Flepp, R., Nüesch, S., & Franck, E. (2016). Does bettor sentiment affect bookmaker pricing? *Journal of Sports Economics*, *17*(1), 3-11.

Forrest, D., & Simmons, R. (2008). Sentiment in the betting market on Spanish football. *Applied Economics*, *40*(1), 119-126.

Franck, E., Verbeek, E., & Nüesch, S. (2011). Sentimental preferences and the organizational regime of betting markets. *Southern Economic Journal*, *78*(2), 502-518.

Frey, V., & Van De Rijt, A. (2020). Social influence undermines the wisdom of the crowd in sequential decision making. *Management Science*, *67*(7), 4273-4286.

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature*, *75*(7), 450-451.

Gauriot, R., & Page, L. (2019). Fooled by performance randomness: overrewarding luck. *Review of Economics and Statistics*, *101*(4), 658-666.

Gino, F., Moore, D. A., & Bazerman, M. H. (2009). No harm, no foul: The outcome bias in ethical judgments. *Working paper, Harvard University.*

Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless+ harmless= blameless: When seemingly irrelevant factors influence judgment of (un) ethical behavior. *Organizational Behavior and Human Decision Processes*, *111*(2), 93-101.

Gurdal, M. Y., Miller, J. B., & Rustichini, A. (2013). Why blame?. *Journal of Political Economy*, *121*(6), 1205-1247.

Hastie, R., & Dawes, R. M. (2009). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Los Angeles, CA: Sage Publications.

Hershey, J. C., & Baron, J. (1992). Judgment by outcomes: When is it justified?. *Organizational Behavior and Human Decision Processes*, *53*(1), 89-93.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101*(46), 16385-16389.

Kausel, E. E., Ventura, S., & Rodríguez, A. (2019). Outcome bias in subjective ratings of performance: Evidence from the (football) field. *Journal of Economic Psychology*, *75*, Article 102132.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? Insights from retail orders and stock returns. *The Journal of Finance*, *68*(3), 1229-1265.

Keuschnigg, M., & Ganser, C. (2017). Crowd wisdom relies on agents' ability in small groups with a voting aggregation rule. *Management Science*, *63*(3), 818-828.

König-Kersting, C., Pollmann, M., Potters, J., & Trautmann, S. T. (2021). Good decision vs. good results: Outcome bias in the evaluation of financial agents. *Theory and Decision*, *90*(1), 31-61.

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

Lefgren, L., Platt, B., & Price, J. (2015). Sticking with what (barely) worked: A test of outcome bias. *Management Science*, *61*(5), 1121-1136.

Marshall, G. W., & Mowen, J. C. (1993). An experimental investigation of the outcome bias in salesperson performance evaluations. *Journal of Personal Selling & Sales Management*, *13*(3), 31-47.

Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, *26*(2-3), 131-146.

Mowen, J. C., & Stone, T. H. (1992). An empirical analysis of outcome biases in constituent evaluations of public policy decision makers. *Journal of Public Policy & Marketing*, *11*(1), 24-32.

Murr, A. E. (2015). The wisdom of crowds: Applying Condorcet's jury theorem to forecasting US presidential elections. *International Journal of Forecasting*, *31*(3), 916-929.

Nofer, M., & Hinz, O. (2014). Are crowds on the internet wiser than experts? The case of a stock prediction community. *Journal of Business Economics*, *84*(3), 303-338.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, *36*(1), 11-46.

Ratner, R. K., & Herbst, K. C. (2005). When good decisions have bad outcomes: The impact of affect on switching behavior. *Organizational Behavior and Human Decision Processes*, *96*(1), 23-37.

Rothschild, D. (2015). Combining forecasts for elections: Accurate, relevant, and timely. *International Journal of Forecasting*, *31*(3), 952-964.

Rubin, J., & Sheremeta, R. (2016). Principal–agent settings with random shocks. *Management Science*, *62*(4), 985-999.

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.

Sjöberg, L. (2009). Are all crowds equally wise? A comparison of political election forecasts by experts and the public. *Journal of Forecasting*, *28*(1), 1-18.

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.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations.* New York, NY: Doubleday.

Tinsley, C. H., Dillon, R. L., & Cronin, M. A. (2012). How near-miss events amplify or attenuate risky decision making. *Management Science*, *58*(9), 1596-1613.

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

Vaughan Williams, L., & Reade, J. J. (2016). Forecasting elections. *Journal of Forecasting*, *35*(4), 308-328.

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