University of Zurich[UZH]

Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 381

Does sentiment harm market efficiency?

An empirical analysis using a betting exchange setting

Oliver Merz, Raphael Flepp and Egon Franck

August 2019

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

Affolternstrasse 56, CH-8050 Zurich, Switzerland

Does sentiment harm market efficiency?
An empirical analysis using a betting exchange setting

Oliver Merz, Raphael Flepp, and Egon Franck*

August 2019

*Corresponding author: Oliver Merz, Affolternstrasse 56, CH-8050 Zurich, Switzerland. oliver.merz@business.uzh.ch. Coauthors: Dr. Raphael Flepp, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland, and Prof. Dr. Egon Franck, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland.

Abstract

This paper investigates whether the sentimental preferences of investors influence market efficiency. We use a betting exchange market environment to analyze the influence of sentimental bettors on market efficiency in 2,333 soccer matches played between 2006-2014 during the last three hours of the pre-play period. Contrary to bookmaker markets, there is no intermediary in a betting exchange and, thus, the market prices solely reflect the beliefs of person to person betting. We use three different proxy variables to measure the bettor sentiment and find that price changes are more likely to be inefficient for betting events that are more prone to sentiment. Based on that finding, we propose a trading strategy that generates positive returns before considering the transaction costs and commission fees. Although the returns turn negative after considering the transaction costs and commission fees, the proposed trading strategy still outperforms a random betting strategy.

Keywords

Sentiment bias · Market efficiency · Forecasting · Betting markets · Soccer

JEL Classification

D40 · G40 · L83

1. Introduction

Investors are constantly confronted with new information from various news outlets on the internet, television and radio. To optimize portfolio returns, investors must pay close attention to the most promising investment opportunities and evaluate and compare their potential. However, attention is a scarce cognitive resource and investors have limited information processing capabilities (Kahneman, 1973; Barber & Odean, 2007). Thus, investors tend to focus on stocks that have recently caught their attention instead of investigating all possible alternatives (Odean, 1999). For example, stocks in the news, stocks experiencing high abnormal trading volumes and stocks with extreme one-day returns are more likely to be bought by individual investors than stocks that are out of the limelight (Barber & Odean, 2007). Similarly, Chen, Hu and Hwang (2014) find that the views expressed on a popular social media website could strongly predict future stock returns and earnings surprises. Furthermore, initial public offerings from well-known firms in the social media industry and other high-profile industries draw a great deal of attention from investors. For example, the IPO of Facebook in 2012 was the third largest IPO in U.S. history (Geron, 2012). A more recent example is Snap, the parent company of the social media application Snapchat, which raised over $3.4 billion in its IPO in 2017 (Hirsch, 2017). The IPO of Snap seems controversial, as Snap does not yet generate profits, does not intend to pay dividends in the upcoming future and does not permit voting rights to shareholders (Levine, 2018; United States Securities and Exchange Commission, 2017). Thus, investors have no voting or cash flow rights and their only benefit stems from potential capital gains if the stock appreciates.

If an unknown share of investors exhibit sentimental preferences towards more popular or glamorous stocks, the question arises if sentiment-motivated investments influence stock prices due to the increased demand for such stocks. On the one hand, if too many investors act on sentimental preferences rather than rational decisions, stock prices may deviate from their true fundamental valuations and thus make markets less efficient (e.g., De Long, Shleifer, Summers & Waldmann, 1990). On the other hand, rational institutional or professional investors could exploit this behavior and counter potential inefficient stock price movements (e.g., Fama, 1965; Ke & Ramalingegowda, 2005). To assess if sentimental traders influence market efficiency, it is necessary to know the asset's true fundamentals. The problem in analyzing stock markets is that fundamental values are difficult to determine because stocks are essentially infinitely lived (Brown & Yang, 2017). As the true valuation of a stock remains unknown, it is difficult to infer whether movements in stock prices correctly reflect insightful new information or whether they are driven by noise factors such as sentiment. Thus, many researchers analyze prices in betting markets to explain theories about anomalies and biases in financial markets. In betting markets, the true fundamental value of each bet is unambiguously revealed once the underlying event (e.g., a football match or an election) is finished, as we can observe the result (Brown & Yang, 2017). Furthermore, the time frame of bets is relatively short so that there is no need to discount. Similar to more high-profile stocks in financial markets, there are, for example, more popular teams in sports with larger fan bases that are more prone to attract sentiment from bettors. The potential reasons for bettor sentiment in the sports environment include a perception bias, where bettors overestimate the winning probabilities of their supported team (e.g., Babad & Katz, 1991; Braun & Kvasnicka, 2013), or a loyalty bias, where bettors are not willing to bet against their favorite team (e.g., Forrest & Simmons, 2008).

The previous research used bookmaker settings to analyze bettor sentiment (e.g., Avery & Chevalier, 1999; Forrest & Simmons, 2008; Franck, Verbeek & Nüesch, 2011). A bookmaker has a similar role to a dealer in a quote-driven financial market, as the bookmaker offers prices against which the bettors can place their bets (Franck et al., 2011). The empirical evidence regarding whether sentiment influences the pricing decisions of bookmakers is mixed. Avery and Chevalier (1999), Strumpf (2003), Levitt (2004) and Feddersen, Humphreys and Soebbing (2017) show that bookmakers increase the prices (lower the odds) for more sentiment-prone teams, while Forrest and Simmons (2008), Franck et al. (2011) and Feddersen, Humphreys and Soebbing (2016) find that bookmakers decreased the prices (increase the odds). Braun and Kvasnicka (2013) find that some bookmakers increase the prices while others decrease the prices in the presence of sentiment. Page (2009) and Flepp, Nüesch and Franck (2016) find no sentiment bias in bookmaker pricing decisions. One explanation for these heterogeneous results is that bookmakers are profit maximizing and have to consider various factors when setting the prices. First, bookmakers must decide whether their books should be balanced or whether they allow for unbalanced books. With a balanced book, a bookmaker makes a guaranteed profit of its commission in the long run without exposing themselves to risk (e.g., Woodland & Woodland, 1994; Avery &

Chevalier, 1999). If the bookmaker decided to allow for an unbalanced book, the bookmaker could exploit bettor sentiment by adjusting the prices (e.g., Levitt, 2004; Paul & Weinbach, 2008). Second, Shin (1991, 1992, 1993) shows that bookmakers may deviate from the true prices to hedge themselves from bettors with superior information. Third, Franck et al. (2011) argue that the decision of the bookmaker relies on the price elasticity of the sentimental bettors. If sentimental bettors are price-insensitive, a bookmaker can increase the prices of sentiment-prone teams to increase its profits. If sentimental bettors are price-sensitive, a bookmaker can decrease the prices of sentiment-prone teams to draw more betting volume. Overall, whether bettor sentiment influences prices, and thus market efficiency, is ultimately dependent on the bookmaker's decision.

In this paper, we propose an alternative setting using a betting exchange market where individual bettors bet against each other. In contrast to bookmaker markets, betting exchanges do not use an intermediary setting the prices. Prices are the sole consequence of trading between individual bettors. We assume that there are two types of bettors. The first type are professional bettors, who act like financial investors and gain utility only through financial profits. The second type are sentiment-prone bettors who either gain additional utility by betting on their team due to a loyalty bias (e.g., Forrest & Simmons, 2008) or are unable to calculate the true probabilities due to a perception bias (e.g., Babad & Katz, 1991; Braun & Kvasnicka, 2013). Therefore, a bias in the prices might occur if the share of sentiment-prone bettors with biased beliefs is large enough. As inefficient price movements are assumed to originate from biased bettor beliefs in a betting exchange setting, the betting exchange should offer a more reliable setting to analyze bettor sentiment than the use of bookmakers. Betting exchange markets work similarly to order-driven financial markets, as they provide a public limit order book where bets can be placed throughout the day. As in financial markets, bettors can submit to the market and limit orders and bet that a certain team wins or loses (Brown & Yang, 2017).

For our analysis, we use the betting exchange Betfair, a leading worldwide sports betting exchange that is based in the UK. We conduct our analysis using 2,333 soccer matches played in the English Premier League throughout the seasons 2006/2007-2013/2014. Soccer is the largest spectator sport and strong emotional attachment for certain teams is very common. We analyze the last three hours of the pre-play period immediately before a match starts. This setting offers some advantageous characteristics. First, compared to in-play, only very limited new information arises during the pre-play period and changes in prices are expected to originate from the trading activity rather than new information. Furthermore, more than 70% of the total pre-play betting volume is generated during the last three hours before the match starts, on average. If sentimental bettors want to increase the excitement of watching a game or showing support towards their team, they are likely to enter a bet during this time frame. Thus, we examine whether bettor sentiment influences the implied predictive power of the odds over this pre-play period.

One major difficulty when analyzing the impact of bettor sentiment on price efficiency is how to measure sentiment. Drawing on the previous literature, we propose three different but logical measures to proxy bettor sentiment. Our first sentiment proxy is based on the winning probabilities implied in the betting odds. Levitt (2004) shows that bettors exhibit a systematic bias towards favorites in football point spread betting, where the spread is set so that the winning probability of the favorite equals the winning probability of the underdog[1]. Thus, we determine the favorite within each match by using the winning probabilities implied in the betting odds. The favorites tend to be the more popular teams with a larger fan base that will attract more attention from the bettors than the underdogs receive. We thus hypothesize that the favorites are more prone to sentiment.

Our second sentiment proxy variable is based on the betting volume. In betting exchanges, bettors are not only able to bet that a certain event occurs (i.e., backing), but they can also bet that a certain event does not occur (i.e., laying). As Forrest and Simmons (2008) point out, fan bettors might gain utility by betting on their team because they increase their stake in the game to demonstrate support and further align the interests between them and their team winning the game. Thus, for every event, we compare the cumulative betting volume that was initiated by back bets to the cumulative betting volume that was initiated by lay bets over the three-hour pre-play period[2]. We expect that sentimental bettors prefer to back their team, and we thus assume more sentiment for events where relatively more volume has been initiated by bettors backing an event rather than laying an event.

[1] Point spread betting in the U.S. implies that the spread is set to offer an approximately 50% winning bet on either the favorite or the longshot. Betting on the favorite is, thus, not attributable to risk preferences.
[2] By comparing the offered back and lay odds with the odds that were last matched, we can determine whether a bet was initiated by a bettor who wants to back a bet or by a bettor who wants to lay a bet.

Our third sentiment proxy follows Avery and Chevalier (1999) and uses the past performance of teams to measure bettor sentiment. The underlying idea is that the teams that performed well in the previous season are more in the news and memories of bettors. Thus, better placed teams are expected to draw more sentiment than worse placed ones. Our third sentiment proxy variable uses the difference in the league table positions of the previous season between two opposing teams.

As sentimental bettors bet on certain teams, the increasing demand for such bets will result in mispricing when there are not enough bettors who are willing to enter the opposite betting position, i.e., sentimental bettors drive prices away from the true fundamentals. We compare the prices at kick-off with the prices three hours before for all events of all matches. According to the efficient market hypothesis of Fama (1970), the prices should contain all the information of historical prices and update accordingly. Thus, as new information arises, the predictive power of the prices should increase or at least stay the same. We show that the price movements during the pre-play period are efficient on average, i.e., over all matches, the implied predictive power in the odds increases. However, we find that in over 40% of the cases, the implied predictive power decreases. This stands clearly in contrast to the proposition of Fama (1970). We find evidence that higher sentiment is associated with inefficient price movements. This paper extends the previous research, as it demonstrates that bettor sentiment impacts prices in a betting exchange setting where prices are the consequence of person to person betting without the influence of an intermediary. We developed simple trading strategies based on our findings. Although the trading strategies are not economically profitable after considering the transaction costs and commission fees, the proposed trading strategies consistently outperform a random betting approach. Thus, our findings imply that the price movements of sentiment-prone events are not completely efficient. Nevertheless, the mispricing evoked by this bias is not large enough to encourage professional bettors to fully counter sentimental bettors, as transaction costs and commission fees prevent profitable trading strategies.

The remainder of this paper is structured as follows. In Section 2, we review the relevant theoretical and empirical literature regarding bettor sentiment. In Section 3, we describe our data from the betting exchange Betfair and specify the dependent and independent variables and the protocol of the methodology used. Section 4 showcases the main results as well as a simple trading strategy, and Section 5 concludes.

2. Bettor Sentiment

In the literature, sentiment is often associated with the nonprofit-maximizing behavior of bettors or investors (e.g., Avery & Chevalier, 1999) and deviations from standard rational choice models (Feddersen et al., 2017). Forrest and Simmons (2008) exemplify two views regarding the existence of sentimental bettors. In the first scenario, there is a share of informed bettors that know the true objective winning probabilities and a share of uninformed, sentimental bettors with biased opinions about the winning probabilities of their preferred team. Sentimental bettors suffer from a perception bias and, thus, overestimate their teams winning probability (Braun & Kvasnicka, 2013). As Babad and Katz (1991) show, bettors are likely to overestimate the winning probabilities of their preferred club due to "wishful thinking," which is not in their best financial interest. In the second scenario, the sentimental bettors can objectively assess the winning probabilities but are willing to wager despite unfavorable odds because it gives them utility when demonstrating their support of their favorite team. It increases the extent to which they are stakeholders in the club. If bettors show a loyalty bias, they either bet on their team despite unfavorable odds or not at all (Braun & Kvasnicka, 2013).

Avery and Chevalier (1999) find that sentiment serves as a predictor of point-spread movements in football games over a 12-year period by comparing the point spreads one weak apart from each other. They look at three different sentiment variables, i.e., expert advice, hot hand and prestige, to investigate the influence of bettor sentiment on the betting line. Avery and Chevalier (1999) assume that the bookmaker, i.e., Las Vegas sports book, adjusts the spreads to balance their books. They find that bookmakers lower the odds (increase the prices) of teams that are sentiment prone. Levitt (2004) analyzed the price-setting mechanism of bookmakers using data of NFL games. By analyzing the betting volumes, Levitt (2004) shows that bettors have preferences towards the favorite within a game. In consensus with Avery and Chevalier (1999), bookmakers exploit these sentimental preferences by lowering the odds (increasing the prices) of the favorites to gain higher profits, as with the market clearing price; however, according to Levitt (2004), bookmakers do not necessarily balance their books. Strumpf (2003) analyzed baseball, football, ice hockey and basketball data from six illegal bookmakers in New York. A substantial share of bettors show sentimental preferences, i.e., high loyalty, towards local home-town teams based in New York, e.g., the New York Yankees in baseball. Thus, the finding of Strumpf (2003) indicates that bookmakers offer unfavorable prices on sentimental bets on local teams to maximize their profits. This is consistent with Kuypers (2000), who shows that bookmakers might set inefficient odds if the bookmaker is profit maximizing and bettors have biased expectations, and with Humphreys (2010), who shows that bookmakers operate with an unbalanced book to exploit the large share of bettors who bet on the stronger team in the NBA. Feddersen et al. (2017) analyze five European soccer leagues, the NBA and the NFL and use Facebook "likes" to proxy bettor sentiment, as more popular teams draw more attention from sentimental bettors. Feddersen et al. (2017) show evidence that bettor sentiment influences the pricing decisions of bookmakers, such that bookmakers offer less favorable odds for sentiment-prone wagers.

While these studies show a consensus in terms of bookmakers decreasing odds (increasing prices) for sentiment-prone bets, another string of studies find the opposite effect or no effect at all. Forrest, Goddard and Simmons (2005) demonstrate that bookmakers make very sophisticated calculations with many variables that can generate very precise forecasts. Thus, Forrest and Simmons (2008) assume that bookmakers do not misprice bets due to inability, but rather deliberately make a commercial decision that exploits the preferences or misperceptions of bettors. Forrest and Simmons (2008) use the difference in stadium attendance in a match as a proxy for bettor sentiment in the Spanish and Scottish soccer leagues and find that better odds (lower prices) are offered for sentiment-prone bets. Similarly, Franck et al. (2011) use the difference in stadium attendance and find evidence in the upper four English soccer leagues that favorable prices are extended to more popular teams winning and that this effect is amplified on weekends where the fraction of sentiment-driven bettors is expected to be higher. Feddersen et al. (2016) used the following two measures of team popularity to proxy bettor sentiment in NBA games: the difference in the arena utilization of two opposing teams and the difference in the share of All Star Game votes received by each of the two opposing teams. Feddersen et al. (2016) find that bookmakers offered more favorable prices to more popular teams with the intention to draw more volume to those bets. Finally, Braun and Kvasnicka (2013) analyze qualification soccer games for the UEFA Euro 2008 to investigate whether national sentiment affects the pricing decisions by bookmakers. Braun and Kvasnicka (2013) find that both positive and negative national sentiment biases towards their own national team exist. Some domestic bookmakers decrease the odds for corresponding national teams, while others increase the odds. Page (2009) analyzes whether the large share of English bettors introduce a bias by betting on English teams in international soccer games or European cup games. However, Page (2009) does not find biased odds due to an optimistic bias introduced by sentimental bettors in soccer matches among various betting firms. Flepp et al. (2016) find no evidence that bookmakers adjust prices due to the presence of bettor sentiment. Flepp et al. (2016) choose the over/ under 2.5 goals per match betting market to analyze the effect of bettor sentiment on bookmaker pricing decisions. Sentimental bettors are expected to bet on high scoring outcomes to increase the excitement in a game. In fact, more than 80% of the volume wagered is placed on the over bet. However, this imbalance in the betting volume distribution has no impact on bookmaker pricing.

Overall, the empirical evidence regarding the influence of bettor sentiment on bookmaker pricing decisions is mixed. The underlying issue is that we simply do not know which factors bookmakers consider when setting their prices. The profit functions may differ from bookmaker to bookmaker and could, thus, explain the heterogeneous findings. Some bookmakers might prefer a guaranteed profit by operating with a balanced book without exposing themselves to risk (e.g., Woodland & Woodland, 1994; Avery & Chevalier, 1999), while others may allow for an unbalanced book to actively exploit bettor sentiment by adjusting the prices accordingly (e.g., Levitt, 2004; Paul & Weinbach, 2008). Some bookmakers may even adjust the prices to hedge themselves from bettors with inside information (Shin, 1991, 1992, 1993). Furthermore, Franck et al. (2011) point out that the pricing decision of a bookmaker depends on the price sensitivity of the bettors. A bookmaker might increase the prices of sentiment-prone teams if sentimental bettors are price-insensitive to generate higher profits or decrease the prices of sentiment-prone teams if sentimental bettors are price-sensitive to an increase in the overall betting volume. Taken together, whether we observe the influence of sentiment on prices is largely dependent on the profit functions and pricing decisions of the bookmakers at hand.

This paper extends the research on bettor sentiment by using a betting exchange market setting where the prices are not set by an intermediary but are the mere consequence of person to person betting. Betting exchange markets are basically prediction markets that only rely on the market mechanism

(Page, 2009). Thus, biased prices are more likely to originate from biased bettor opinions rather than a third party's profit function.

As the summary of previous research illustrates, there is no distinct approach regarding how to measure sentiment. It is important to note that all sentiment proxy variables probably have weaknesses. For example, Feddersen et al. (2017) point out that attendance-based sentiment proxies have certain limitations. First, proxies based on stadium attendance might only capture the local popularity of a team and are, thus, dependent on the market potential of the surrounding region. Bettors with biased beliefs exist worldwide, though and are not only located in a specific region. Further, teams can even influence the attendance themselves by reducing ticket prices or increasing promotional activities (Feddersen et al., 2017). Thus, Feddersen et al. (2016) name two criteria that proxy variables should fulfill. First, they should not be only local measures, and second, they should not be under the direct control of the teams. In this paper, we address these requirements, as we use three different measures to proxy bettor sentiment that fulfill the criteria indicated by Feddersen et al. (2016).

3. Methods

The methods section is separated into three subsections. First, we depict the data used for our analysis. Second, we explain the specifications of the dependent variable and three independent sentiment proxy variables. Third, we show the methodology used.

3.1 Data

We collected data on 2,333 soccer matches in the English Premier League for the seasons 2006/2007-2013/2014. The data originates from Betfair, the largest betting exchange in the world that merged in 2016 with Paddy Power to become Paddy Power Betfair. Paddy Power Betfair is a diversified international sports betting and gaming operator with a total annual revenue of £1,745 million, an operating profit of £250 million and 7,640 employees as of December 2017 (Paddy Power Betfair, 2018). The data from Betfair were purchased through an authorized third-party data provider called Fragsoft[3]. While the data for most matches are only available for the 3-5 hours prior to the start of the match, there are a few exceptions where the provided pre-play data goes further back. To obtain more

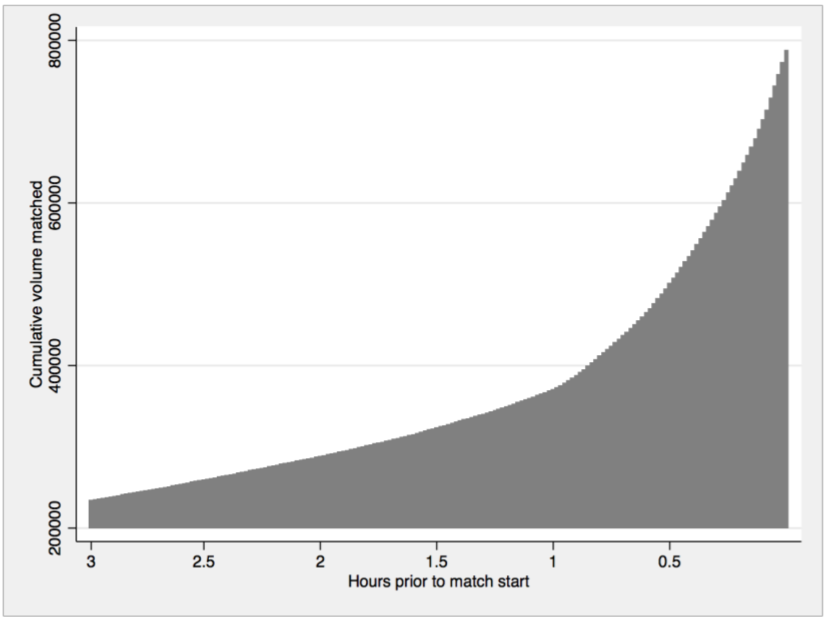
[3] Roughly 15% of the matches are missing. Choi and Hui (2014) point out that Fracsoft attributes the missing data to technical issues while recording the data from Betfair.

similar and consistent pre-play periods, we limit the maximum pre-play period to three hours. We do not include matches with pre-play periods of less than three hours to ensure consistent and comparable time periods across matches, leaving us with 2,333 matches[4].

For each match event (home team win, draw and away team win), the data contains detailed second-by-second data on the first three best back and lay decimal odds, the last price matched, as well as the corresponding volumes traded and the total matched volumes[5]. To conduct our empirical analysis, we use the last price matched values to calculate the implied prices. By using the last price matched variable, we consider the prices of trades that actually took place.

The three hours prior to the match start are particularly relevant because more than 70% of the total matched volume prior to the match start is generated during this time (see figure 1). If bettors with sentimental preferences influence price movements before a match starts, the mispricing will likely occur during this period. Further, the three hour pre-play period represents a good setting, as prices are generally far less volatile due to the absence of major news, such as goals, cards or substitutions, that occur during a match, and price movements are more likely to originate due to the trading behavior of the bettors.

Figure 1: Cumulative volume matched over the pre-play period.



Notes: Betting volume averaged per match.

[4] Instead of 2,562 matches if all matches were included.
[5] As in the previous research (e.g., Forrest and Simmons, 2008), we only consider home and away wins for our analysis and neglect the event "draw" because a draw is unlikely to be prone to sentimental preferences.

3.2 Measurement and Variables

3.2.1 Odds and implied winning probabilities

The decimal odds denote the payoff of a successful bet. For instance, if the odds for a win for Manchester United are 1.50, a one dollar bet pays \$1.50 if Manchester United wins, i.e., a profit of \$0.50 has been achieved. The reciprocal of the odds can be seen as the market's forecasting probability of a certain bet to win (see, e.g., Forrest & McHale, 2007; Forrest & Simmons, 2008 and Flepp, Nüesch & Franck, 2017). For each match *i*, event $e \in \{home\ win,\ away\ win\}$ and time *t* before the match start, the implied winning probability is defined as follows:

$$p_{iet} = \frac{1}{lpm_{iet}} \tag{1}$$

where *lpm* refers to the decimal odds that were last matched between two bettors in the betting exchange. The value of p tells us the implied winning probability, e.g., the probability of a Manchester United win from above is $p = \frac{1}{1.50} \approx 0.67$. In this case, the underlying assumption of the odds suggests that Manchester United wins the game with a 67% chance. The implied winning probabilities can also be interpreted as prices, as they indicate the stake a bettor must wager to earn \$1 in the event of a successful bet (Forrest & Simmons, 2008).

3.2.2 Dependent variable

Using the implied winning probabilities, we can form predictions about the outcome of a match and analyze whether the predictive power of the prices changes over the pre-play period. We follow Wolfers and Zitzewitz (2004) in calculating the absolute prediction errors. The use of prediction errors is a common tool in the forecasting of elections or sports match outcomes (Mincer & Zarnowitz, 1969; Williams & Reade, 2016; Brown & Yang, 2019; Brown, Reade & Williams, 2019). The absolute prediction error is calculated as follows:

$$APE_{iet} = |result_{ieT} - p_{iet}| \tag{2}$$

where *result* is a binary variable indicating the true *ex post* outcome of the match, i.e., *result* equals 1 for a win and 0 for a loss, and p refers to the implied winning probability as depicted in equation (1). We calculate the absolute prediction errors three hours before the match and at kick-off and compare their sizes to assess the change in the predictive power. Formally, this can be expressed as follows:

$$\Delta APE_{ie} = APE_{ie\ match\ start} - APE_{ie\ pre-play\ start} \tag{3}$$

The prices should fully reflect all available information in efficient markets (Fama, 1970). As the most recent price should contain at least the same amount of information as the price three hours before, the predictive power of this price should increase or stay the same. To measure if sentiment influences prices and, thus, market efficiency, we specify our dependent variable *inefficiency* as follows:

$$inefficiency = \begin{cases} 1 & if\ \Delta APE_{ie} > 0 \\ 0 & if\ \Delta APE_{ie} \leq 0 \end{cases} \tag{4}$$

The use of the difference in absolute prediction errors is comparable to using match fixed-effects, as it partials out match specific characteristics, e.g., the magnitude of the absolute prediction error[6].

3.2.3 Independent variables

We hypothesize that inefficient updating of prices occurs at least partially due to sentimental bettors who bet on their favorite teams regardless of the accuracy of the odds offered. We present three different approaches to proxy bettor sentiment.

3.2.3.1 Odds-based sentiment proxy

First, we specify a variable *favorite* based on the winning probabilities implied in the betting odds. Levitt (2004) shows in football spread betting that even if the probability of a favorite win is approximately equal to the probability of a win for a non-favorite, bettors still prefer to bet on the favorite[7]. This indicates the importance of the popularity of a team over other potential factors, e.g., the risk-aversion of bettors. We suggest that favorites tend to be more successful teams, attract more supporters and, thus, are more prone to sentiment.

Further, we observe that the betting volumes increase much more drastically during the pre-play period for favorites than for underdogs. For favorites, the average betting volume per match increases during the pre-play period by approximately 970,000, while the average betting volume per match for underdogs only increases by approximately 135,000. It illustrates how the favorite in a match receives

[6] A specification error might occur if only the magnitudes of the absolute prediction errors were compared across matches, as close matches with no distinct favorite would yield a relatively high prediction error regardless of who wins.

[7] In spread betting, the bookmaker can adjust the spread to even out the winning probabilities of two opposing teams.

much more attention by bettors, which is similar to the finding of Barber and Odean (2007) that more glamorous stocks receive more attention from financial investors.

Three hours before the match starts, we assess which team is considered the favorite to win a particular game according to the winning probabilities implied by the betting odds. We specify *favorite* as an indicator variable equaling 1 if the corresponding team is the favorite within a match, and 0 otherwise. We hypothesize that sentimental bettors prefer favorites and that this drives the inefficient updating of prices.

3.2.3.2 Volume-based sentiment proxy

Our second variable to proxy bettor sentiment is based on the betting volume. In bookmaker markets, bettors traditionally have three options to bet on, i.e., whether the home team wins, the away team wins or the match ends in a draw. In betting exchanges bettors can both bet for an event to occur (back a bet) and bet against an event to occur (lay a bet). Figure 2 illustrates this using an example of a match between Manchester United and Leicester. For illustration purposes, imagine a bettor chooses to either bet on Manchester United (back Manchester United at the best offered odds of 1.51) or to bet

Figure 2: Backing and laying odds on the Betfair betting exchange

3 selections			Back all	Lay all		
Man Utd	**1.48** €1090	**1.49** €505	**1.51** €234	**1.52** €110	**1.54** €495	**1.56** €244
Leicester	**7.6** €45	**7.8** €112	**8** €16	**8.4** €234	**8.6** €100	**8.8** €40
The Draw	**4.3** €337	**4.4** €77	**4.5** €168	**4.7** €160	**4.8** €410	**5** €336

Notes: The odds are depicted in bold and the corresponding betting volumes are shown below. The blue-colored fields represent the best offered back odds and the red-colored fields represent the best offered lay odds.

against a win of Manchester United (lay Manchester United at the best offered odds of 1.52). Backing Manchester United yields a profit if Manchester United wins the game against Leicester, and laying Manchester United yields a profit if Manchester United does not win the game against Leicester, i.e., if Manchester United loses or draws against Leicester.

Forrest and Simmons (2008) explain that fans of a certain team might gain utility by betting on their team because they increase the extent to which they are stakeholders in the team. By backing their team, they demonstrate support and align their financial interests with those of the team. Braun and Kvasnicka (2013) and Forrest and Simmons (2008) explain that sentimental bettors exhibit a loyalty bias towards their club and are unlikely to bet against their club even if the prices increase and are less favorable. This can be compared to fans that would not start buying jerseys from opposing teams because they offer the jerseys at a lower price.

For all events, we compare the cumulative betting volume that was initiated by the bettors backing the event to the cumulative betting volume that was initiated by the bettors laying the event over the course of the three hours' pre-play period. Thus, we form our sentiment proxy variable *back lay ratio* as the ratio of the cumulative back-initiated volume to the cumulative lay-initiated volume for all events. As we expect the sentimental bettors to back their preferred team, we assume more sentiment for events where relatively more cumulative volume has been generated by the initiated backing bets rather than by the initiated laying bets. To determine whether a bet is initiated by back or lay odds, we compare the last price matched with the best odds offered on the back and lay side. Therefore, in our example above, if a bet on Manchester United were matched at 1.51, then this bet would be characterized as a back initiated bet, as a bettor chose to back Manchester United at the offered odds of 1.51. As an illustration, imagine that over the observed pre-play period, $2,000 of the betting volume for Manchester United was characterized as back-initiated, as the last price matched was closer to the last back odds and $500 was characterized as lay-initiated, as the last price matched was closer to the last lay odds. For our sentiment proxy variable *back lay ratio,* we would obtain a ratio of 4 for Manchester United because the volume that has been matched coming from the back side is four times higher than the volume coming from the lay side over the course of the three-hour pre-play period. We hypothesize that the higher the amount of back-initiated volume is relative to the lay-initiated volume, the higher the share of sentimental bettors should be and the more likely we should observe inefficient price movements.

3.2.3.3 Past performance based sentiment proxy

Our third variable to proxy bettor sentiment follows an approach of Avery and Chevalier (1999) and is based on the past performance to determine whether a team is considered to be prestigious and, thus, sentiment prone. It is debatable what the appropriate time horizon is to determine whether to classify a team as prestigious or not. For simplicity, we choose the last year's table position to measure past performance. For each team and year, we assign the corresponding league table position at the end of the previous season[8]. The data for past league table positions originate from the official webpage of the English Premier League. For each team, we calculate the difference in the lagged league position using the following formula:

$$\Delta\, table\ position\ lag_i \;=\; opposition\ rank_i - team\ rank_i \tag{5}$$

For illustrative purposes, suppose Chelsea placed second and Wigan 16th in the season 2010/11. Our sentiment proxy variable for the games between Chelsea and Wigan in the season 2011/12 would therefore show the value of 14 for Chelsea and -14 for Wigan. A team with a relatively higher placement in the previous season obtains a higher value for the past performance sentiment proxy. Thus, the more positive the value of this variable, the more sentiment is expected. We hypothesize that the more positive the value of the past performance sentiment proxy variable is, the more likely we observe inefficient price movements due to sentiment.

3.3. Statistical method

By estimating simple logit regressions, we want to examine the influence of sentimental bettors on market efficiency. The logit regression is specified as follows:

$$inefficieny_{ie} = \theta_0 + \theta_1 \cdot sentiment_{ie} + \theta_2 \cdot home_{ie} + \vartheta_{ie} \tag{6}$$

In this logit regression, $inefficieny_{ie}$ is the dependent variable and *sentiment* represents one of the three aforementioned sentiment proxy variables (*favorite, ratio back lay, Δ table position lag*) for each match *i* and event *e* ∈ $\{home\ win,\ away\ win\}$. To partial out a potential home team bias (e.g., Forrest and Simmons, 2008), we include a control variable *home* that is a dummy variable equaling 1 for home teams and 0 for away teams. Following the approach of Forrest and Simmons (2008), we conduct our regression for both home and away wins, resulting in a total of 4,666 observations for 2,333 matches. Analogous to Forrest and Simmons (2008), we adjust for correlated error terms, as a win for a team implies a loss of its opponent, and use clusters for each match.

[8] For the three teams that were in the lower division a year before, we follow van Ours and Tuijl (2016) and assigned the last position, i.e., 20. We assume this to be a good approximation, as the teams in the lower division most likely received less attention from bettors. The results do not change if we use the average position of the teams that were relegated, i.e., $\frac{18+19+20}{3} = 19$, or, alternatively, if we excluded all matches that involved those teams.

4. Results

4.1 Main results

Using a t-test we examine how the absolute prediction error differs on average at the beginning and at the end of the pre-play period. Table 1 shows that, on average, the absolute prediction errors decrease over the pre-play period, indicating that the price movements are overall efficient. This is not surprising, as new information arises to better predict the match outcome.

Table 1: t-test of *absolute prediction error (APE)*

	At pre-play start			At pre-play end			
	N	Mean	SE	N	Mean	SE	t
APE	4,666	0.3839	.0031	4,666	0.3828	.0031	4.43***

Notes: ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

However, taking a deeper look at the share of increasing versus decreasing absolute prediction errors, we find that in over 41% of the events, the absolute prediction error increases, which stands in contrast to the efficient market hypothesis that suggests that prices should include all past information (Fama, 1970). According to our hypothesis, the share of inefficient price updating is explained by sentimental bettors that deteriorate prices by betting on their favorite team without taking into account how precisely the odds reflect the true winning probabilities. Table 2 shows the results of logit regressions using different variables to proxy bettor sentiment. Column (1) reports the estimate for the odds-based sentiment proxy. The indicator variable *favorite* has a positive impact on the likelihood of an inefficient price movement during the pre-play period and is statistically significant at the 5% level. This indicates that market inefficiency is more likely to be observed for the favorites in a match. Column (2) reports the estimate for the volume-based sentiment proxy *ratio back lay*. The positive sign of the coefficient indicates that the higher the ratio of back to lay initiated volume for a match is, the higher the probability that we observe an inefficient price movement. The coefficient is significant at the 1% level. Finally, column (3) reports the estimate for the past performance sentiment proxy *Δ table position lag,* which is significant at the 5% level. As expected, the sign of the coefficient for *Δ table position lag* is also positive. This suggests that the higher a team was placed than its opponent in the previous season, the higher the likelihood of inefficient price movements.

Table 2: Results of logit regressions for *inefficiency*

	inefficiency		
	(1)	(2)	(3)
favorite	0.1225** (0.0546)		
ratio back lay		0.1836*** (0.0549)	
Δ table position lag			0.0062** (0.0029)
home dummy	0.0348 (0.0545)	0.0844* (0.0496)	0.0852* (0.0498)
N	4,666	4,666	4,666
Log pseudolikelyhood	-3156.13	-3152.72	-3156.38
Pseudo R^2	0.0009	0.0020	0.0008

Notes: The table reports the logit estimates for *inefficiency*; *inefficiency* takes the value of 1 if the APE increased over the pre-play period, and 0 otherwise. Column (1) represents an odds-based sentiment proxy, column (2) represents a volume-based sentiment proxy and column (3) uses a sentiment proxy based on past performance. *Home dummy* controls for teams that play at home. The heteroscedasticity-robust and clustered standard errors at the match level are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Overall, the depicted sentiment proxies consistently indicate a positive impact of bettor sentiment on market inefficiency. The presence of bettor sentiment increases the probability that we observe an inefficient price movement over the last three hours of the pre-play period.

As a robustness check, we use the difference in implied winning probabilities of two opposing teams as an alternative to the indicator variable *favorite* to check if the results hold if the magnitude of the difference in the winning probabilities is considered. It yields similar results as the *favorite* variable and is significant at the 10% level, also indicating a positive relationship between the likelihood of observing inefficient price movements and the extent to which a team is relatively more favored than its opponent. Additionally, as an alternative to the one year lagged difference in table positions, we used the difference in one year lagged points won between two opposing teams[9]. The coefficient is significant

[9] For this regression, we exclude the teams that were in a lower league in the previous year, as the points won in lower leagues are difficult to compare to the points won in a higher league. For the ease of interpretation, we use

at the 5% level as well, and the sign is positive, which is consistent with the other sentiment proxies, as the higher the difference in points won, the more sentiment-prone a team is expected to be.

4.2 Trading strategy

In this section, we want to explore whether it is possible to exploit the observed inefficient price movements and obtain profits by constructing simple trading strategies.

As we expect a share of sentimental bettors to wager during the three-hour pre-play period, we expect the odds for sentiment-prone teams to decrease over this time. Thus, a simple trading strategy could be formed as follows: entering a one unit bet to back a team classified as sentiment-prone at the start of the observed pre-play period and then laying a bet on the same event at the start of the match. The underlying idea of the trading strategy is to be independent of the match outcome and profit from the expected decreasing odds of sentiment-prone bets. To ensure that the profit or loss is the same independent of the match outcome, the amount wagered on the lay bet is adjusted accordingly[10]. While we can form trading strategies for the *favorite* and the *Δ table position lag* sentiment proxy, it is not possible to develop a strategy for the *ratio back lay* sentiment proxy, as the ratio of the back- and lay-initiated volume is calculated over the course of the analyzed time and, thus, is not known *ex ante*. The results of the simple trading strategies for the other two sentiment proxies are depicted in table 3. As depicted in panel A, the average return for the trading strategy using *favorite* to classify sentiment-prone bets is 0.0023 before the transaction costs and commission fee. However, as shown in panel B, after the transaction costs and commission fees are incorporated, the return turns negative, i.e., to -0.0062[11]. Similarly, the depicted thresholds for the *Δ table position lag* sentiment proxy show positive returns without considering the transactions costs and commission fees but negative returns if these are included. The same trading strategies for teams classified as non-favorites or for any negative threshold for the difference in the lagged table position strictly yield negative returns (not depicted for brevity).

---

the formula $\Delta\ league\ points\ lag = team\ points - opposition\ points$ so that a higher value for this variable indicates more sentiment.

[10] To ensure that the same profit or loss for each match outcome is achieved, the following condition must hold: $stake_{lay} = stake_{back} \times \frac{odds_{back}}{odds_{lay}}$

[11] Transaction costs denote the spread between the back odds and lay odds. As a commission fee, we use the default market base rate of 5% for the UK and Ireland from Betfair. Therefore, all winning bets were reduced by 5%.

As a comparison, we employed a random betting strategy for all different constellations of matches chosen to be bet on by our sentiment proxies.

Table 3: Rates of return for strategies based on sentiment proxy *variables*

Panel A	Return before transaction costs and commission		
	Number of matches	Rate of return	Random betting rate of return
Favorite	2,330	0.0023	-0.0051
Δ table position lag > 0	2,302	0.0018	-0.0052
Δ table position lag > 6	1,159	0.0020	-0.0066

Panel B	Return after transaction costs and commission		
	Number of matches	Rate of return	Random betting rate of return
Favorite	2,330	-0.0062	-0.0214
Δ table position lag > 0	2,302	-0.0086	-0.0214
Δ table position lag > 6	1,159	-0.0069	-0.0243

Notes: The true number of bets is twice the depicted values of the number of matches as we back and lay the same event. The rate of return shows the mean return for the respective sentiment proxy. The thresholds for the *Δ table position lag* variable correspond to the 50% and 75% percentiles. Panel A depicts the returns before considering the transaction costs and commission fees (5%), while Panel B depicts the returns after correcting for transactions and commission costs. The random betting returns were calculated for the same sample of matches used for the different sentiment proxy variables.

As table 3 shows, the returns of random betting are consistently lower than the returns of the corresponding bets based on the sentiment proxy variables. The returns of random betting for the matches involved in the strategy using the *favorite* sentiment proxy are -0.0051 before the transaction costs and commission fees are included and -0.0214 after the transaction costs and commission fees are included. The corresponding returns for the different thresholds of *Δ table position lag* are very similar in size. While our proposed trading strategies theoretically work and are superior to a random betting strategy, the practical implementation is not economically viable.

As we used the whole dataset to conduct the analysis, the depicted trading strategies are tested *in-sample* only. Alternatively, we split the sample and conduct the main analysis without the most recent season as well as without the last 2, 3 or 4 seasons. The excluded seasons were then used to test the trading strategies. The impacts of the sentiment proxy variables on our measurement for inefficiency are virtually identical in sign, magnitude and significance across all different sample constellations. Similarly, the returns are positive without considering the transactions costs and commission fees and are negative when we incorporate the costs across all the abovementioned scenarios.

5. Conclusion

This paper is the first to use a betting exchange market setting instead of the bookmaker settings used in the previous research to analyze the impact of bettor sentiment on prices. An advantage of the betting exchange market setting is that there is no intermediary setting the prices. Thus, the prices effectively reflect the bettors' beliefs and trading decisions rather than a third party's profit function.

We examine 2,333 soccer matches in the English Premier League to analyze whether bettor sentiment influences market efficiency during the last three hours of the pre-play period. On average, over 70% of the total matched volume prior to the match start is generated during the last three hours of the pre-play period. We hypothesize that the potential presence of sentimental bettors can introduce a bias during this period and, thus, harm market efficiency. By comparing the absolute prediction errors at the start and at the end of the pre-play period, we find that the market is overall efficient. However, in over 41% of the events, we observe inefficient price movements. To examine whether bettor sentiment can explain the observed inefficient price updating, we form three different sentiment proxy variables, i.e., an odds-based sentiment proxy, a volume-based sentiment proxy, as well as past performance as an indicator for sentiment. We find evidence that bettor sentiment attributes to inefficiencies. Our sentiment proxies show a consistent picture of a positive relationship between bettor sentiment and inefficient price movements. This indicates that bettor sentiment can at least partially explain some of the inefficiencies observed during the pre-play period.

Based on our findings, we formed simple trading strategies to try to profit from these inefficiencies. We find positive returns for different specifications without incorporating the transaction costs and commission fees. However, after considering the transaction costs and commission fees, the trading strategies are no longer profitable. While our proposed trading strategies generate superior returns to randomly betting on matches, it is evident that the inefficiencies in the betting exchange are not large enough to be exploited economically.

This paper contributes to the wide discussion regarding whether sentiment influences prices. We extend the previous literature by observing the impact of bettor sentiment on prices in person to person markets without an intermediary, i.e., where prices are the sole consequence of trading. Thus, we can conclude that the found inefficiencies stem from bettors' beliefs and their betting behavior rather than from a bookmaker's individual profit function.

A potential implication for the found inefficiencies could either be that the share of purely profit-oriented professional bettors is not large enough to fully counter the inefficient price movements caused by sentiment prone bettors or more likely that the mispricing caused by the sentiment bias is too small to be exploited by professional bettors, as the transaction cost and commission fees cannot be overcompensated. Perhaps the mispricing introduced by bettors with sentimental preferences is exploited by professional bettors to the point where it is no longer profitable to counter it.

References

Avery, C., & Chevalier, J. (1999). Identifying investor sentiment from price paths: the case of football betting. *The Journal of Business*, 72(4), 493-521.

Babad, E., & Katz, Y. (1991). Wishful thinking—against all odds. *Journal of Applied Social Psychology*, 21(23), 1921-1938.

Barber, B. M., & Odean, T. (2007). All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21(2), 785-818.

Braun, S., & Kvasnicka, M. (2013). National sentiment and economic behavior: evidence from online betting on European football. *Journal of Sports Economics*, 14(1), 45-64.

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

Brown, A., & Yang, F. (2017). The role of speculative trade in market efficiency: Evidence from a betting exchange. *Review of Finance*, 21(2), 583-603.

Brown, A., & Yang, F. (2019). The wisdom of large and small crowds: evidence from repeated natural experiments in sports betting. *International Journal of Forecasting*, 35(1), 288-296.

Chen, H., De, P., Hu, Y., & Hwang, B. H. (2014). Wisdom of crowds: the value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5), 1367-1403.

Choi, D., & Hui, S. K. (2014). The role of surprise: understanding overreaction and underreaction to unanticipated events using in-play soccer betting market. *Journal of Economic Behavior & Organization*, 107, 614-629.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 98(4), 703-738.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.

Feddersen, A., Humphreys, B. R., & Soebbing, B. P. (2016). Sentiment bias in national basketball association betting. *Journal of Sports Economics*, 19(4), 455-472.

Feddersen, A., Humphreys, B. R., & Soebbing, B. P. (2017). Sentiment bias and asset prices: evidence from sports betting markets and social media. *Economic Inquiry*, 55(2), 1119-1129.

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

Flepp, R., Nüesch, S., & Franck, E. (2017). The liquidity advantage of the quote-driven market: Evidence from the betting industry. *The Quarterly Review of Economics and Finance*, 64, 306-317.

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

Forrest, D., & McHale, I. (2007). Anyone for tennis (betting)?. *The European Journal of Finance*, 13(8), 751-768.

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.

Geron, T. (2012). *Facebook prices third-largest IPO ever, valued at $104 billion.* https://www.forbes.com/sites/tomiogeron/2012/05/17/facebook-prices-ipo-at-38-per-share/#532a24cc728a. Accessed on 8 July 2019.

Hirsch, L. (2017). *Snap's shares pop after year's biggest IPO*. https://www.reuters.com/article/us-snap-ipo-idUSKBN1690I7. Accessed on 8 July 2019.

Humphreys, B. R. (2010). Point spread shading and behavioral biases in NBA betting markets. *Rivista di Diritto ed Economia dello Sport*, 6(1), 13-26.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Ke, B., & Ramalingegowda, S. (2005). Do institutional investors exploit the post-earnings announcement drift?. *Journal of Accounting and Economics*, 39(1), 25-53.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, 32(11), 1353-1363.

Levine, M. (2018). *Investors gave Snap a gift.* https://www.bloomberg.com/view/articles/2018-07-11/investors-gave-snap-a-gift. Accessed on 8 July 2019.

Levitt, S. D. (2004). Why are gambling markets organised so differently from financial markets?. *The Economic Journal*, 114(495), 223-246.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3-46). NBER.

Odean, T. (1999). Do investors trade too much?. *American Economic Review*, 89(5), 1279-1298.

Paddy Power Betfair (2018). *Annual report & accounts 2017*. Retrivied from https://www.paddypowerbetfair.com/~/media/Files/P/Paddy-Power-Betfair/documents/ppb-r-n-a-2017-180327.pdf

Page, L. (2009). Is there an optimistic bias on betting markets?. *Economics Letters*, 102(2), 70-72.

Paul, R. J., & Weinbach, A. P. (2008). Price setting in the NBA gambling market: Tests of the Levitt model of sportsbook behavior. *International Journal of Sport Finance*, 3(3), 137.

Shin, H. S. (1991). Optimal betting odds against insider traders. *The Economic Journal*, 101(408), 1179-1185.

Shin, H. S. (1992). Prices of State Contingent Claims with Insider Traders, and the Favourite-Longshot Bias. *The Economic Journal*, 102(411), 426-435.

Shin, H. S. (1993). Measuring the incidence of insider trading in a market for state-contingent claims. *The Economic Journal*, 103(420), 1141-1153.

Strumpf, K. (2003). Illegal sports bookmakers. Unpublished manuscript. Available at http://rgco. org/articles/illegal_sports_bookmakers. pdf.

United States Securities and Exchange Commission (2017). *Snap Inc*. https://www.sec.gov/Archives/edgar/data/1564408/000119312517068848/d270216d424b4.htm#rom270216_6. Accessed on 8 July 2019.

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

Williams, L. V., & 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.

Woodland, L. M., & Woodland, B. M. (1994). Market efficiency and the favorite-longshot bias: the baseball betting market. *The Journal of Finance*, 49(1), 269-279.