University of Zurich[UZH]

Department of Business Administration

UZH Business Working Paper Series

Working Paper No. 384

Sonic Thunder vs. Brian the Snail

Are people affected by uninformative racehorse names?

Oliver Merz, Raphael Flepp and Egon Franck

November 2020

---

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

Sonic Thunder vs. Brian the Snail
Are people affected by uninformative racehorse names?

Oliver Merz, Raphael Flepp, and Egon Franck*

November 2020

*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 individuals' decision making is affected by fast-sounding horse names in a betting exchange market environment. In horse racing, the name of a horse does not depend on the horse's performance and is thus uninformative. If positive affect towards fast-sounding horse names is present, we expect less accurate prices (winning probabilities) and lower returns due to the increased demand for these bets. Using over 3 million horse bets, we find evidence that the winning probabilities of bets on horses with fast-sounding names are overstated, which impairs the prediction accuracy of such bets. This finding implies that prices in betting exchange markets are not efficient, as they become distorted by incorporating affective, misleading information from a horse's fast-sounding name. This bias translates into significantly lower betting returns for horses classified as fast-sounding compared to the returns for all other horses.

Keywords

Affect heuristic · Decision making · Market efficiency · Betting industry · Horse racing

JEL Classification

D40 · G40 · G41 · L83

1. Introduction

Imagine betting on a horse in a race without properly knowing the past performance or rankings of the horses involved in the race. You could choose to bet your money on a horse called *Sonic Thunder* or on a horse called *Brian the Snail*. On which horse would you bet?

We are faced with dozens of options and choices every day. Thus, we continuously make decisions, whether they are job-related or concern our private live. Often, it is too exhaustive, time-consuming or simply not feasible to consider all relevant factors when making a decision. Tversky and Kahneman (1974) explain that to simplify decision making, people often rely on mental shortcuts when faced with complicated choices. They elaborate three distinct heuristics “representativeness”, “availability” and “adjustment and anchoring” that individuals use to make judgments under uncertainty. Furthermore, Tversky (1972) shows how individuals employ a simplified strategy, “elimination-by-aspects”, to make choices. Subsequently, the literature on cognitive strategies underlying judgment and choice has been growing rapidly and extended by various models such as the “recognition heuristic” (Goldstein & Gigerenzer, 1999, 2002), the “advantage model” (Shafir, Osherson & Smith, 1989, 1993) and models based on the “construction of preference” (Payne, Bettman & Johnson, 1993; Slovic, 1995). These examples illustrate that the focus of descriptive decision research has long been cognitive rather than affective. Nevertheless, the affect heuristic has received increasing attention regarding human judgment and decision making (Finucane, Alhakami, Slovic & Johnson, 2000). An early advocate of its importance is Zajonc (1980), who argues that affective responses occur independently from cognitive operations. Moreover, affect can occur automatically and even before cognitive encoding, influencing information processing and judgment. Thus, all perceptions contain some level of affect. According to Zajonc (1980), we do not just see a house; we see a “handsome” house or an “ugly” house. Furthermore, Zajonc (1980) argues that we often buy the cars we like, choose jobs and houses we find attractive and justify these choices at a later stage for various reasons.

Thus, affect might even play an important role when making financial decisions. Hsee (1998) illustrates that people showed a higher willingness to pay for a smaller amount of ice cream presented in an overfilled cup than for a larger amount of ice cream in an underfilled cup. Statman, Fisher and Anginer (2008) provide the explanation that an overfilled cup is perceived as more generous, and its affect on the participants is positive, while an underfilled cup is perceived as stingy, and its affect on the participants is negative. Marketing companies seem to be well aware of the importance of the affect heuristic. For instance, background music is often played in shopping malls to lighten shoppers' mood and increase their willingness to spend money; models in advertisements smile to link positive affect to products; and cigarette advertising used to be designed in a way to overcome the perceptions of health risks by increasing the positive affect associated with smoking (Epstein, 1994). Both positive and negative affective feelings guide the judgment and decision making of people in a variety of circumstances (Slovic, Finucane, Peters & MacGregor, 2007). In a different setting, Hahari and McDavid (1973) find that teachers' evaluations of children's performance are associated with stereotyped perceptions of first names. The authorship of essays was randomly linked to children with common, popular and attractive names and with rare, unpopular and unattractive names. The assessed quality of the essays was higher when the attached author had a name associated with positive stereotypes. Erwin and Calev (1984) also find that individuals assigned with more desirable names achieved higher grades than individuals with less desirable names because of the implied expectations of the evaluators.

In contrast to previous research, in which most studies are conducted in a controlled environment, we consider empirical data in a real-life setting in which substantial amounts of money are in play. Similar to Krčál, Kvasnička and Staněk (2016), we use a betting market setting to analyze whether decision research insights from laboratory experiments can also be found in the real world. More specifically, we analyze a betting exchange market in which individuals bet on horses to win a race. Following the findings of Hahari and McDavid (1973) and Erwin and Calev (1984), we assume that people experience positive affect towards certain racehorse names. Because the goal in horse betting is to pick a fast-running horse that is likely to win the race, we expect to find positive affect towards horses with fast-sounding names that suggest good performance[1].

Importantly, the name of a horse such as *Sonic Thunder* or *Brian the Snail* is completely uninformative because the regulations concerning the naming of a horse do not consider past success. Rather, the name of a horse usually cannot be changed after the horse has participated in a race. Thus, the name of the horse has no signaling effect regarding its performance.[2] However, according to Zajonc (1980), affective reactions to stimuli are often people's very first reactions that subsequently guide their information processing and judgments. Thus, a casual bettor might simply make his or her decision based on the most obvious attribute of a horse, its name. If the bettor associates the horse's name with speed and thus expects that the horse will run fast, then the bettor might already be inclined to bet on the horse before making an extensive assessment of all publicly available factors, such as the offered odds, past placements or the physical appearance of the horse. As Slovic et al. (2007) state, affective responses occur rapidly and automatically, and people quickly associate feelings with stimulus images or words such as "treasure" or "hate". Therefore, the bettor might already have subconsciously experienced positive affect towards a fast-sounding horse well before evaluating the horse's chance of winning.

To test whether affect influences individuals' decision making, we analyze whether objectively irrelevant information, i.e., the names of racehorses, impact prices in the betting market. In betting (prediction) markets, individuals trade on the outcomes of future events (Brown, Reade & Vaughan Williams, 2019). Thus, as betting (prediction) markets are

[1] We are aware that other types of names might attract positive affect, as many horse names are catchy, funny or resemble celebrities. However, as the main goal of betting is to win, we limit our study to the names that imply superior performance.

[2] Indeed, our empirical results show that the correlation of horses with fast-sounding names and actual wins is basically zero (-0.002).

designed to allocate resources and aggregate information, market prices forecast future events (Berg, Nelson & Rietz, 2008). As the participants in these markets profit from accurate predictions, they have strong incentives to acquire useful information, and thus, market prices are expected to accurately forecast the underlying outcome (Brown et al., 2019). In the prediction market literature, there is a wide consensus regarding the high accuracy of prediction market prices (see, e.g., Wolfers & Zitzewitz, 2004; Berg et al., 2008; Spann & Skiera, 2009; Rothschild, 2015; Vaughan Williams & Reade, 2016). Similar to the efficient market hypothesis in financial markets (see Fama, 1970), betting (prediction) markets are efficient if the market prices reflect all historical information and are the best forecasts for the outcome probabilities of an event (Angelini & de Angelis, 2019). Methodologically speaking, a market price is efficient if no other variable has explanatory power regarding the outcome of an event after controlling for the price. While the overall evidence seems to point towards highly accurate and efficient prices in betting (prediction) markets, some studies have found deviations from efficient market prices due to various behavioral biases, such as favorite-longshot bias (see, e.g., Ottaviani & Sørensen, 2008; Forrest & Mchale, 2007; Buhagiar, Cortis & Newall, 2018), sentiment bias (e.g., Avery & Chevalier, 1999; Levitt, 2004; Feddersen, Humphreys & Soebbing, 2017) or home bias (e.g., Forrest & Simmons, 2008).

If fast-sounding names have a positive affect on bettors in horse racing, price accuracy impairments will result because of the disproportionate demand for such bets. Thus, we expect less accurate prices for bets on fast-sounding horses because the positive affect of bettors essentially increases the price for those bets. Consequently, we expect lower returns on bets on fast-sounding horses compared to bets on other horses.

We use data from over 400,000 horse races between 2008 and 2018 with more than 3 million horse bets. Data are obtained from the betting exchange Betfair, in which bettors trade bets against each other in a continuous double auction. Following previous research, e.g., Forrest and Simmons (2008) and Franck, Verbeek and Nüesch (2011), we use logit regressions with the outcome of a bet as the dependent variable (equaling 1 if the bet is won and 0 if lost), and as the explanatory variables, we use the probabilities implied in the odds and a binary variable indicating whether a horse is classified as fast-sounding. If the odds (prices) are efficient, all relevant information should be reflected in them, and no additional variables should have predictive power regarding the outcome of an event[3].

We find that a fast-sounding horse name has predictive power with regard to the race outcome beyond the probabilities implied in the odds. In particular, our results show that the winning probabilities of bets on horses with fast-sounding names are overstated, implying that the prices in betting exchange markets are not completely efficient, as prices become distorted due to the incorporation of misleading or false information. Furthermore, we find significantly lower returns for horses classified as fast-sounding compared to other horses. A simple trading strategy of betting against all horses classified as fast-sounding yields a return of approximately 2.9% before commission but a negative return of -1.6% after commission[4]. This finding could be bracketed under the "limits of arbitrage" argument of Gromb and Vayanos (2010) because the mispricing is not large enough to overcome the transaction costs; thus, potentially misleading or false information is not fully eliminated from prices. Nevertheless, this strategy generates significantly larger profits than a random betting strategy in which zero returns are achieved before commission and a negative return of -4.7% is achieved after the commission is considered. Despite wagering real money, a substantial share of the betting community seems to be systematically biased in preferring bets on fast-sounding horses to bets on other horses. This finding supports the view that objectively irrelevant factors affect people's judgment and even impair price accuracy and market efficiency to some degree.

[3] As in previous research, e.g., Forrest and Simmons (2008) and Franck et al. (2011), we calculate prices as the reciprocal of the odds. The price is the amount of money one has to 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 horse to win the race, i.e., the payoff would be 2.0 times the amount wagered, then the price would be 0.5, which also denotes the implied winning probability of the horse.

[4] We applied the default market base rate of 5% for UK and Ireland from Betfair.

Overall, this paper contributes to the literature on the role the affect heuristic plays in human decision making and extends the previous literature by examining the implications of positive affect in a real-life setting in which individuals' best interest is to be rational, as substantial amounts of money are in play. Interestingly, the aggregated impact of the biased betting decisions is large enough to yield significantly lower returns on such bets and falsely assign predictive power about the race outcome to uninformative outcome signals. Thus, this paper also contributes to the literature on (prediction) betting market efficiency and demonstrates that although markets are highly effective at allocating resources and aggregating information, their forecasting accuracy is impaired if large crowds with biased valuations are present.

The remainder of this paper is structured as follows. In section 2, we describe the data, the process of classifying fast-sounding horses and the empirical methodology used. In section 3, we present our results and a trading strategy. Section 4 concludes the paper.

2. Methods

2.1 Data

We collected betting data on horse races from Paddy Power Betfair, one of the largest online wagering operators in the world (Paddy Power Betfair PLC., 2018)[5]. Betfair operates markets across various sports, as well as in politics and economics (Paddy Power Betfair PLC., 2018). The betting exchange mechanism mirrors a standard limit order book in financial markets, where traders can submit limit and market orders in a continuous double auction (Flepp, Nüesch, & Franck, 2017). Thus, bettors can choose whether they want to place a bet that an event will occur (backing a bet) or bet against the event occurring (laying a bet). If two parties with opposing opinions agree on a price, their bets are matched, and a

[5] The data on Betfair starting prices (2020) are freely available on https://promo.betfair.com/betfairsp/prices

transaction takes place (Franck, Verbeek & Nüesch, 2010). Furthermore, Betfair offers the option to back or to lay a horse at an ex ante unknown starting price (SP). At the time the race begins, Betfair determines the market clearing SP from the aggregate volumes of back and lay bets.[6] For each race, we obtained data on the date, time and location of the race, the names of the participating horses, the winner of the race, the weighted average matched price (WAP) and the SP at the end of the prerace period. The WAP and the SP are both denoted in decimal odds.

We collected data on 443,850 horse races held in the UK, Ireland, the US, South Africa and Australia from March 2008 until May 2018. With an average number of approximately 9.2 horses per race, we observed a total of 4,066,445 horses. We exclude all races in which more than one horse won simultaneously and all races in which we cannot observe all odds for the participating horses because the values were either missing or erroneous[7]. Thus, the final samples used in our analysis consisted of 344,749 races and 3,193,458 horses using the WAP and 422,816 races and 3,903,604 horses using the SP. Summary statistics for the betting odds of WAP and SP are depicted in Table 1.

[Insert Table 1 about here]

2.2 Fast-sounding name categorization

The categorization of horse names into fast-sounding and not fast-sounding is of critical importance. To determine an objective list of terms that are associated with speed, we used four different sources. In particular, we used two commonly employed and very popular thesauri, the Cambridge and Oxford dictionaries, a word association API called *twinwords* and www.horses-names.com, an independent website that provides suggestions for horse

[6] The SP is calculated by Betfair to match the betting volumes from the back and lay sides. The idea of the SP is to generate a price that matches the largest share of the betting volume.

[7] In approximately 900 races, there were multiple winners. Our results are insensitive to the inclusion of those races. A value for the WAP or SP is deemed erroneous if the price is below 1.01 or above 1000, which corresponds to the odds range given by Betfair. Our results do not change if we exclude only the missing values for WAP or SP instead of the whole race.

names, to evaluate whether a horse's name is associated with speed[8]. Using the Cambridge and Oxford dictionaries, we looked up synonyms for the words "fast" and "speed" and then classified a horse as fast-sounding if one of those terms was contained in the horse's name[9]. Using the *twinwords* API, we captured horses whose name included a word associated with the terms "speed" or "fast". Additionally, we used the name suggestions for fast horses from www.horses-names.com to classify fast-sounding horses. While the dictionaries and word association API capture synonyms for "speed" and "fast", this website provides words that are commonly associated with speed, e.g., "rocket" or "comet." A comprehensive list of the terms determined to indicate a fast-sounding name can be found in the Appendix in Tables 8-11.

A horse is classified as fast-sounding if its name includes any of the terms on the list, e.g., "Speed Dragon" would be classified as fast-sounding because of the word "Speed"[10]. Table 2 shows 40 selected fast-sounding horses to further illustrate which names are classified as fast-sounding.

[Insert Table 2 about here]

The number of horses classified as fast-sounding from each source is shown in Table 3. Panel A shows the individual contributions of the four sources to the share of horses categorized as fast-sounding when using the WAP, and Panel B analogically shows the corresponding contributions to the share of horses categorized as fast-sounding when using

[8] The data for the different sources is freely accessible and was retrieved from the following URLs.
Cambridge Dictionary: https://dictionary.cambridge.org/topics/moving-quickly-and-slowly/fast-and-rapid/
Oxford Dictionary: https://en.oxforddictionaries.com/thesaurus/fast and https://en.oxforddictionaries.com/thesaurus/speed
Twinwords API: https://www.twinword.com/api/word-associations.php
Horse names website: http://www.horses-names.com/fast-horse-names.php

[9] Unfortunately, the Cambridge Dictionary does not provide a synonym search for speed. However, the term "speed" appears as a synonym for "fast". The Oxford Dictionary allows for a more thorough approach, as it distinguishes between adverbs and adjectives for "fast" as well as between nouns and verbs for "speed". Further, the Oxford Dictionary groups synonyms depending on the context. We only include terms that relate to speed of movement. We did not include words declared by Oxford to be vulgar, informal, rare, literary or archaic, as these are likely to be misinterpreted.

[10] For the classification of horse names, we consider only words starting with a capital letter, e.g., we search for "Speed" but not "speed". With this approach, we can capture almost all relevant names because our data provide the names with a capital letter between spaces. At the same time, this approach avoids cases in which preceding letters change the meaning of a word, e.g., the term "top" would also capture "stop".

the SP. The underlying reason for the categorization approach used in this paper is twofold. First, it is important to obtain a comprehensive list of terms indicating speed to capture a large share of horses with fast-sounding names and thus enable comparisons between the two groups. As shown in Table 3, the individual sources provide relatively small shares of the horses classified as fast-sounding. However, if we combine all sources additively, we can establish a more comprehensive subsample. As some classifications overlap, we have a final subsample of fast-sounding horse names that represents approximately 4.3% of observations for the WAP variable and approximately 4.6% for the SP variable.

[Insert Table 3 about here]

Second, the use of four independent and complete sources helps to mitigate the potential subjectivity of the name categorization process. Using the whole list of words associated with speed inevitably leads to some errors in categorization, as some words might not be distinctively related to movement speed and could be interpreted differently depending on the context. However, to ensure an objective classification procedure, we refrain from excluding individual words that are potentially subject to misinterpretation and always use the complete list of words associated with "fast" and "speed" according to the four sources[11].

2.3 Statistical methods

Following previous research, e.g., Forrest and McHale (2007), Forrest and Simmons (2008) and Flepp, Nüesch and Franck (2016), we use the reciprocal of the decimal odds to calculate the market's forecasting probability of a certain bet to win. We calculate the implied winning probabilities $impliedprob_{i,wap} = \frac{1}{WAP}$ and $impliedprob_{i,SP} = \frac{1}{SP}$ for each horse. A favorite horse that is more likely to win has a higher implied winning probability, as it trades

[11] In an alternative approach, we exclude the terms we consensually deemed nonsensical to validate our results. Despite having a smaller share of fast-sounding horses, we obtain marginally stronger results when excluding the terms that have little to no association to speed.

at lower odds. For example, a horse with an SP of 1.25 is expected to win in $\frac{1}{1.25} = 80\%$ of cases.

We follow Franck et al. (2011) and Forrest and Simmons (2008) by examining whether our indicator variable for fast-sounding horse names has explanatory power beyond the implied probabilities with regard to the actual outcomes. We test a binary model with the actual outcome of a bet (1 for a winning bet; 0 for a losing bet) as the dependent variable; the implied probability and our indicator variable (fast-sounding) are explanatory variables. Specifically, we estimate our multivariate logit model as follows:

$$Ln\left[\frac{P(Win_i = 1)}{P(Win_i = 0)}\right] = \beta_0 + \beta_1 impliedprob_i + \beta_2 fast\text{-}sounding_i + \varepsilon_i \qquad (1)$$

where *impliedprob* is the probability that is implied by the WAP or SP and *fast-sounding* is a binary variable indicating whether a horse's name is classified as fast-sounding or not.

Under the null hypothesis, we assume that betting exchange markets are efficient and thus that prices fully reflect all available information, including the names of the horses. In other words, we assume the prices to be the best outcome forecasts of the underlying events. Thus, the coefficient β_2 of the fast-sounding variable should be zero. If our indicator variable has explanatory power in addition to the implied probabilities, then on average, the odds are not efficient, and bets on horses with fast-sounding names are not equally profitable as bets on other horses. We expect a negative sign for the fast-sounding variable if a large enough share of bettors with a biased preference towards fast-sounding names demand bets on those horses. As we include multiple observations of the same race (bets on the participating horses), the independence assumption between those observations is violated. To account for this, we compute clustered heteroscedasticity-robust standard errors at the race level.

In an alternative approach, we use t-tests and nonparametric Wilcoxon rank-sum tests to check whether betting returns significantly differ among bets on horses that have a fast-sounding name and horses without a fast-sounding name. We calculate returns on one-unit bets using the formula $return_i = \frac{stake \times odds_i\text{-}stake}{stake}$, where *odds*$_i$ represent either the WAP or the SP. Because the returns are not normally distributed, i.e., for approximately 89% of the observations, the return is -1 whenever a bet is lost, we also conduct Wilcoxon rank-sum tests. We determine the returns by conducting one-unit bets on horses to win a race. If markets are rational and efficient, there should be no systematic difference in returns between any subgroups of bets.

3. Results

3.1 Price accuracy of fast-sounding horse bets

Table 4 depicts the summary statistics and correlation coefficients for the variables *impliedprob* and *fast-sounding*. The correlation coefficients indicate that there is almost no relationship between fast-sounding horses and well-performing horses. If anything, fast-sounding horses seem to be slightly less likely to win because the correlation coefficient between *Win*$_i$ and *Fast-sounding*$_i$ is negative. The correlations between fast-sounding horses and the probabilities implied in the odds are slightly positive for the WAP and slightly negative for the SP. Overall, the correlations are close to zero, indicating a weak relationship.

[Insert Table 4 about here]

The results of the logit regression are depicted in Table 5. The results are shown in terms of marginal effects measured at a point where the continuous *impliedprob* variable is set to its mean and the binary *fast-sounding* variable is set to zero. Panel A in Table 5 shows the results using the WAP to calculate the implied probabilities, and Panel B shows the results using the SP to do so. The results for the two estimations for the *impliedprob* variable stemming from WAP and SP are consistent. The sign for *impliedprob* is positive and significant at the 1% significance level for both estimations. More importantly, the sign of the variable *fast-sounding* is negative and significant at the 1% significance level for the WAP implied probabilities and negative and significant at the 5% significance level for the SP implied probabilities. Thus, the information about fast-sounding horses is not correctly reflected in the market odds (prices), and the null hypothesis of market efficiency is rejected. The variable *fast-sounding* has a significant impact on predicting a win and has a negative sign, implying that the implied probabilities for fast-sounding horses are too high. This result suggests that horses with fast-sounding names are overvalued by bettors.

[Insert Table 5 about here]

To test the robustness of our results, we conduct several variations of our main specification. First, as an alternative to the logit model that assumes a logistic distribution, we run the regressions using a probit model and standard OLS and obtain virtually the same results (not shown for brevity). Second, we randomly select a bet on one horse per race as an alternative to using clustered standard errors and obtain very similar results. Finally, following Forrest and Simmons (2008), we adjust the implied probabilities of the horses such that they sum to one for each particular race. Again, our results remain virtually the same.

3.2 Comparison of betting returns

Using two-sided t-tests and nonparametric Wilcoxon rank-sum tests, we examine how the returns differ on average for horses classified as fast-sounding compared to horses that are not classified as fast-sounding. Table 6 shows that on average, the returns for fast-sounding horses are significantly lower than the returns for the other horses. Panel A shows the results using the WAP to calculate returns, and Panel B shows the results using the SP to calculate returns. Independent of the odds used to calculate the returns, we observe significantly lower returns on bets on horses that are classified as fast-sounding. The returns on bets on fast-sounding horses are approximately 4.1 percentage points lower when using the WAP, and the difference is significant at the 1% significance level. Using the SP leads to similar results, a 3.6-percentage-point lower return on bets on fast-sounding horses, and the difference is significant at the 5% significance level. Nonparametric Wilcoxon rank-sum tests confirm the results, with the differences being significant at the 1% significance level for both the WAP and the SP. Overall, the results of the t-tests and Wilcoxon rank-sum tests suggest that the returns on bets on fast-sounding horses are systematically lower than the returns on bets on horses without fast-sounding names.

[Insert Table 6 about here]

Because we observe lower returns for fast-sounding horses, we can derive a simple trading strategy to exploit this finding. Due to the nature of betting exchange markets with no intermediary, the losses of the losing bettors equal the gains of the winning bettors. Thus, our findings imply positive returns on bets against all horses classified as fast-sounding[12]. Table 7 shows the returns from backing fast-sounding horses and laying fast-sounding horses before and after commission costs are considered. As lay bets work differently compared to back bets, the potential gains and losses differ. For a back bet, the maximum loss a bettor risks is the stake, e.g., 1 unit in our case. For lay bets, a bettor potentially risks much more than the stake. The potential loss of a lay bet is called the liability and can be calculated using the following formula: $liability = stake \times odds\text{-}stake$. Table 7 illustrates that for back bets, a bettor never loses more than 1 unit, while for lay bets, the largest loss equals 739 units. In our sample using the same stakes, returns on back bets are characterized by many small losses and few large gains, while the returns on lay bets are characterized by few large losses and many small gains.

[Insert Table 7 about here]

Theoretically, a return of approximately 2.9% could be achieved by simply betting against all fast-sounding horses. However, the returns become -1.6% after incorporating the Betfair commission rate of 5%[13]. Thus, the mispricing in the odds for fast-sounding horses is

[12] A bettor could simply choose to bet against all horses classified as fast-sounding at the Betfair SP. Although the WAP shows even stronger results, it could not be used to form a trading strategy, as the WAP is the volume-weighted average of the odds traded during the preplay period that is unknown to a bettor ex ante.
[13] At Betfair, 5% is the default market base rate for UK and Ireland.

not large enough to overcompensate for the commission costs imposed by Betfair. Although our proposed strategy of laying fast-sounding horses is not economically viable, the betting returns are significantly higher that the returns from a random betting strategy.[14]

4. Conclusion

The purpose of this paper is to examine whether positive affect influences decision making and, as a result, price accuracy in a prediction market environment in which prices are formed by individuals trading with each other. Specifically, we analyze whether a horse name that indicates speed affects bettor behavior and, consequently, the forecasting accuracy of prices on the betting exchange.

We find that the winning probabilities of bets on horses with fast-sounding names—for which positive affect is likely—are overstated, which impairs the prediction accuracy of such bets. Thus, the prices in the betting exchange market are not efficient because they incorporate misleading information from a horse's fast-sounding name. Furthermore, we find that returns on horses with fast-sounding names are systematically lower than the returns on all other horses. This result suggests that bettors should avoid jumping on the bandwagon when many other bettors are tempted to base their investment decisions on irrelevant factors and instead be aware of the potential mispricing of such bets.

This paper contributes to the role affect plays in human decision making. It extends the previous literature by using a real-life setting in which decisions have a substantial financial impact. In betting markets, people wager real money, and it is in their best interest to act rationally. Nevertheless, our findings suggest that people are biased due to positive affective feelings towards fast-sounding horses. Furthermore, this paper contributes to the general discussion of prediction market accuracy and extends it by analyzing the impact of emotional and affective betting decisions on price accuracy and market efficiency. We show

[14] For this strategy, we randomly select 177,422 bets from our entire sample. After commission, this strategy generates a lay return of approximately -4.7%.

that the presence of large crowds relying on affective preferences harms the price accuracy in prediction markets to some degree.

Data availability statement

The data that support the findings of this study are openly available in the Directory Listing of Betfair price files at https://promo.betfair.com/betfairsp/prices.

References

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

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

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.

Buhagiar, R., Cortis, D., & Newall, P. W. (2018). Why do some soccer bettors lose more money than others?. *Journal of Behavioral and Experimental Finance*, 18, 85-93.

[dataset] Betfair Starting Price (2020). *Directory Listing of Betfair price files*. Available from: https://promo.betfair.com/betfairsp/prices.

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

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, 49(8), 709.

Erwin, P. G., & Calev, A. (1984). The influence of Christian name stereotypes on the marking of children's essays. *British Journal of Educational Psychology*, 54(2), 223-227.

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. (2017). Sentiment bias and asset prices: Evidence from sports betting markets and social media. *Economic Inquiry*, 55(2), 1119-1129.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of behavioral decision making*, 13(1), 1-17.

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., & 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. (2010). Prediction accuracy of different market structures—bookmakers versus a betting exchange. *International Journal of Forecasting*, 26(3), 448-459.

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

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In *Simple heuristics that make us smart* (pp. 37-58). Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological review*, 109(1), 75.

Gromb, D., & Vayanos, D. (2010). Limits of arbitrage. *Annu. Rev. Financ. Econ*., 2(1), 251-275.

Harari, H., & McDavid, J. W. (1973). Name stereotypes and teachers' expectations. *Journal of educational psychology*, 65(2), 222.

Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, 11(2), 107-121.

Krčál, O., Kvasnička, M., & Staněk, R. (2016). External validity of prospect theory: The evidence from soccer betting. *Journal of Behavioral and Experimental Economics*, 65, 121-127.

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

Ottaviani, M., & Sørensen, P. N. (2008). The favorite-longshot bias: An overview of the main explanations. In Handbook of Sports and Lottery markets (pp. 83-101). Elsevier.

Paddy Power Betfair PLC. (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

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge university press.

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

Shafir, E. B., Osherson, D. N., & Smith, E. E. (1989). An advantage model of choice. *Journal of Behavioral Decision Making*, 2(1), 1-23.

Shafir, E. B., Osherson, D. N., & Smith, E. E. (1993). The advantage model: A comparative theory of evaluation and choice under risk. *Organizational behavior and human decision processes*, 55(3), 325-378.

Slovic, P. (1995). The construction of preference. *American psychologist*, 50(5), 364.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European journal of operational research*, 177(3), 1333-1352.

Statman, M., Fisher, K. L., & Anginer, D. (2008). Affect in a behavioral asset-pricing model. *Financial Analysts Journal*, 64(2), 20-29.

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.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.

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.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American psychologist*, 35(2), 151.

Tables (main text)

Table 1: Summary statistics for betting odds

	Observations	Number of Races	Mean	SD	Min	Max
WAP	3,193,458	344,749	38.41	86.20	1.01	1000
SP	3,903,604	422,816	47.32	106.87	1.02	1000

Notes: WAP is the weighted average of the traded prices prior to the race based on their traded volume. SP is the Betfair starting price calculated by Betfair based on the volumes from backing and laying customers.

Table 2: Examples of fast-sounding horse names

Sonic Power	Speed Dragon	Zippy Lad	Lightening Vault
Powerful Jet	Orbit Express	Swift Chap	Blazing Tempo
Brave Falcon	Rush Now	Top Magic	Dixie Flyer
Esprit De Bullet	Strike Fast	Hustle Hard	Diamond Rush
Crown Me Fast	Hot Seat	Top Gear	Bright Bullet
Quick Art	Rush Of Blood	Top Boy	Meteoric Moments
One Wild Guy	Sonic Thunder	Grand Gallop	Zippy Speed
Run for Roses	Saratoga Wildcat	Quick Beers	Sudden Rush
Flyingwithoutwings	Fast On	Dazzlem Quick	You Drive I Fly
Irish Rocket	Hot Sauce	Mighty Flying Thomas	Fullshot

Notes: To illustrate the types of names that are classified as fast-sounding, 40 names have been selected.

Table 3: Sample composition

Panel A: Share of horses classified as fast-sounding using WAP					
Classified as fast-sounding	Cambridge Dictionary	Oxford Dictionary	Twinwords	Horses-names.com	Combined
0	3,133,678 (98.13%)	3,120,071 (97.70%)	3,165,247 (99.12%)	3,150,886 (98.67%)	3,055,314 (95.67%)
1	59,780 (1.87%)	73,387 (2.30%)	28,211 (0.88%)	42,572 (1.33%)	138,144 (4.33%)

Panel B: Share of horses classified as fast-sounding using SP					
Classified as fast-sounding	Cambridge Dictionary	Oxford Dictionary	Twinwords	Horses-names.com	Combined
0	3,829,343 (98.10%)	3,810,480 (97.61%)	3,867,983 (99.09%)	3,846,086 (98.53%)	3,726,182 (95.45%)
1	74,261 (1.90%)	93,124 (2.39%)	35,621 (0.91%)	57,518 (1.47%)	177,422 (4.55%)

Notes: This table shows the share of horses classified as fast-sounding across different sources. Panel A shows the distributions of horses classified as fast-sounding for the WAP, and Panel B shows the distributions of horses classified as fast-sounding for the SP.

Table 4: Summary statistics and correlation coefficients

Panel A: Sample using WAP							
Variable	Mean	SD	Min	Max	1	2	3
1 Win_i *(0/1)*	0.1078	0.3102	0.0000	1.0000	1.0000		
2 $Impliedprob_{i,WAP}$	0.1185	0.1270	0.0010	0.9901	0.3681	1.0000	
3 $Fast\text{-}sounding_i$	0.0433	0.2034	0.0000	1.0000	-0.0018	0.0046	1.0000
Panel B: Sample using SP							
Variable	Mean	SD	Min	Max	1	2	3
1 Win_i *(0/1)*	0.1083	0.3110	0.0000	1.0000	1.0000		
2 $Impliedprob_{i,SP}$	0.1088	0.1205	0.0010	0.9804	0.3856	1.0000	
3 $Fast\text{-}sounding_i$	0.0455	0.2083	0.0000	1.0000	-0.0019	-0.0013	1.0000

Notes: Panel A shows the summary statistics and correlation coefficients using the WAP to calculate the implied probabilities, and Panel B shows those when using the SP. Win_i represents the actual outcome of bet i (0/1), *impliedprob* is the probability odds, and *fast-sounding* is an indicator variable for fast-sounding horse names.

Table 5: Results of logit regressions

Panel A: Using winning probabilities implied by WAP

		Win (0/1)
Impliedprob$_{WAP}$		0.509***
		(0.001)
Fast-sounding		-0.005***
		(0.001)
Number of observations	3,193,458	
Number of clusters	344,749	
Pseudo R^2	0.145	
Log pseudolikelihood	-933,625.89	

Panel B: Using winning probabilities implied by SP

		Win (0/1)
Impliedprob$_{SP}$		0.553***
		(0.001)
Fast-sounding		-0.002**
		(0.001)
Number of observations	3,903,604	
Number of clusters	422,816	
Pseudo R^2	0.158	
Log pseudolikelihood	-1,126,406.1	

Notes: The dependent variable *Win* is binary variable equaling 1 if a horse won the race or 0 if the horse did not win the race. Marginal effects of the variables *impliedprob* and *fast-sounding* are depicted. Heteroscedasticity-robust and clustered standard errors at the race level are reported in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively.

Table 6: Betting returns comparisons for bets on fast-sounding horses and bets on other horses

Panel A: Returns using the WAP

		t-test			Wilcoxon rank-sum test		
	N	Mean	SE	t	Rank sum	Expected	z
Fast-sounding	138,144	-0.1610	0.0139		$2.199 \cdot 10^{11}$	$2.206 \cdot 10^{11}$	
Other	3,055,314	-0.1206	0.0030		$4.879 \cdot 10^{12}$	$4.879 \cdot 10^{12}$	
Δ	3,199,790	-0.0405	0.0146	-2.780***			-3.536***

Panel B: Returns using the SP

		t-test			Wilcoxon rank-sum test		
	N	Mean	SE	t	Rank sum	Expected	Z
Fast-sounding	177,422	-0.0289	0.0155		$3.454 \cdot 10^{11}$	$3.463 \cdot 10^{11}$	
Other	3,726,182	0.0067	0.0038		$7.274 \cdot 10^{12}$	$7.273 \cdot 10^{12}$	
Δ	3,903,604	-0.0356	0.0179	-1.988**			-3.656***

Notes: The table displays the results of simple two-sided t-tests and Wilcoxon rank-sum tests based on the two groups of horses whose names are classified as fast-sounding and other horses. Panel A shows the returns using the WAP, and Panel B shows the returns using the SP.

Table 7: Returns on back and lay bets on fast-sounding horses

Panel A: Returns before commission

	N	Mean	SD	Min	Max
Back return	177,422	-0.0289	6.5348	-1	739
Lay return	177,422	0.0289	6.5348	-739	1

Panel B: Returns after commission

	N	Mean	SD	Min	Max
Back return	177,422	-0.0722	6.2147	-1	702.05
Lay return	177,422	-0.0158	6.5281	-739	0.95

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 5% has been deducted.

Tables (appendix)

Table 8: *Cambridge Dictionary* synonyms for “fast”

a mile a minute	helter-skelter	quick-fire
apace	high-speed	quickly
as fast as your legs would carry you	hot	quickness
as if it is going out of style	hotfoot	rapid
at a rate of knots	hustle	rapid-fire
at full pelt	in the twinkling of an eye	rate
at full speed	Jack Robinson	say
at full tilt	lick	shot
at full tilt	lickety-split	smartly
before you can say Jack Robinson	lightning	souped-up
blistering	like a shot	spanking
breakneck	like a streak of lightning	speed
brisk	like lightning	spread like wildfire
chop-chop	meteoric	streak
crash	mile	style
express	nimble	superfast
fast	nimbleness	supersonic
fleet	nimbly	swift
full	nippy	swiftly
full steam ahead	pdq	thick
gallop	pell-mell	thick and fast
galloping	poky	tilt
go like hot cakes	posthaste	top
have a heavy foot	precipitous	twinkling
headlong	precipitously	whoosh
heavy	prompt	wildfire
hell	promptly	zippy
hell for leather	quick	

Notes: List of synonyms for “fast” from the Cambridge Dictionary, retrieved from: https://dictionary.cambridge.org/topics/moving-quickly-and-slowly/fast-and-rapid/

Table 9: *Oxford Dictionary* synonyms

Panel A: Synonyms for "fast"		
accelerated	high-speed	pell-mell
at full speed	hurried	post-haste
at full tilt	hurriedly	quick
at speed	in a flash	quickly
at the speed of light	in a hurry	rapid
blistering	in a trice	rapidly
breakneck	in a wink	smart
brisk	in haste	speedily
briskly	in in time	speedy
energetically	in no time at all	sporty
expeditious	in the blink of an eye	sprightly
expeditiously	like a flash	swift
express	like a shot	swiftly
fast	like an arrow from a bow	turbo
fast-moving	lively	unhesitating
fleet-footed	meteoric	whirlwind
flying	nimble	with all haste
hastily	on the double	with dispatch
hasty	pell-mell	without delay

Panel B: Synonyms for "speed"		
acceleration	haste	scutter
alacrity	hasten	sharpness
blast	hurriedness	shoot
bolt	hurry	spank along
bowl along	hurry	speed
briskness	hurtle	speediness
career	immediacy	sprint
celerity	momentum	stampede
charge	pace	streak
dart	precipitateness	sweep
dash	promptness	swiftness
dispatch	quickness	swoop
expedition	race	tempo
expeditiousness	rapidity	uzz
fastness	rate	velocity
flash	rattle along	whirl
fly	run	whizz
gallop	rush	whoosh
go hell for leather	scramble	wing

go like lightning	scud	zoom
hare	scurry	

Notes: List of synonyms for “fast” (both adjectives and adverbs) from the Oxford Dictionary, retrieved from: https://en.oxforddictionaries.com/thesaurus/fast
We only include terms regarding movement speed, i.e., the first section of adjectives and adverbs. Terms belonging to informal, British informal, North American informal, literary or rare categories are not included. List of synonyms for “speed” (both nouns and verbs) from the Oxford Dictionary, retrieved from:
https://en.oxforddictionaries.com/thesaurus/speed
Terms belonging to informal, British informal, Scottish informal, North American informal, North American vulgar slang, literary, archaic or rare categories are not included.

Table 10: *Twinword API* words associated with “fast” and “speed”

Panel A: Words related to “fast”		
abrupt	impetuous	rushed
agility	outrun	scramble
dash	overhasty	speed
disconcerted	overrun	speedily
dodge	promptly	speedy
haste	quick	sudden
hastily	quickly	suddenly
hurried	rapid	swift
hurriedly	rapidly	swiftly
hurry	rush	zoom
Panel B: Words related to “speed”		
accelerate	haste	race
acceleration	hasten	rapidity
agility	hie	rush
airspeed	hurriedly	speedy
celerity	hurry	stronghold
dash	pace	swift
decelerate	quick	swiftness
expedite	quicken	tempo
fast	quickly	urgently
fastness	quickness	velocity

Notes: List of words associated with “fast” from the Twinword API, retrieved from: https://www.twinword.com/api/word-associations.php

Table 11: Fast horse name suggestions from the horse-names website

Apache	Bentley	Blustery
Bullet	Buzz	Comet
Cougar	Falcon	Faster
Flash	Ghost rider	Harley
Jet	Jump	Jumping
Miles	Mustang	Pony express
Quick	Quicky	Racer
Rapid	Rapide	Rocket
Sonic	Speedy	Taz
Tornado	Traveler	Wildfire
Voyager	Wild	Velocity

Notes: List of fast-sounding name suggestions from an independent website, retrieved from: http://www.horses-names.com/fast-horse-names.php