International Journal of Forecasting 39 (2023) 1413–1423

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Betting on a buzz: Mispricing and inefficiency in online sportsbooks☆

Philip Ramirez, J. James Reade, Carl Singleton *

Department of Economics, University of Reading, Whiteknights Campus, RG6 6EL, UK

ARTICLE INFO

Keywords:
Wisdom of crowds
Betting markets
Efficient Market Hypothesis
Forecast efficiency
Professional tennis

ABSTRACT

Bookmakers sell claims to bettors that depend on the outcomes of professional sports events. Like other financial assets, the wisdom of crowds could help sellers to price these claims more efficiently. We use the Wikipedia profile page views of professional tennis players involved in over 10,000 singles matches to construct a buzz factor. This measures the difference between players in their pre-match page views relative to the usual number of views they received over the previous year. The buzz factor significantly predicts mispricing by bookmakers. Using this fact to forecast match outcomes, we demonstrate that a strategy of betting on players who received more pre-match buzz than their opponents can generate substantial profits. These results imply that sportsbooks could price outcomes more efficiently by listening to the buzz.

1. Introduction

The size and ubiquity of online sports betting markets continues to increase. Most notably in recent years, the world's most successful online sportsbooks entered the U.S. after a 2018 Supreme Court ruling allowed states to legalise gambling at their own discretion.[1] As online sports betting markets have grown and replaced more traditional forms of gambling, lower transaction costs have increased competition and driven down bookmaker profit margins (i.e., the 'overround' or 'vig') (Forrest, 2008). Over the same period, the amount of online information that bettors can use to form expectations about sports outcomes has increased. This includes detailed historical data about the participants and the setting of an event, the commentary and predictions of sports pundits and tipsters, and the so-called 'wisdom of crowds'. This latter term is used widely to describe instances where information aggregated from the decisions of many individuals improves forecasting and decision-making processes, compared with relying on a small number of expert positions (Galton, 1907; Surowiecki, 2004). Given the small profit margins and competition with the crowd-based betting exchanges (prediction markets), odds-setters may need to forecast outcomes and price the claims they sell to bettors more efficiently than ever before. It is natural to ask whether bookmakers are doing this successfully. In this paper, we use a specific practical example to demonstrate how online sportsbooks are vulnerable to information that could represent the wisdom of crowds.

☆ We would like to thank Giovanni Angelini, Luca De Angelis, Sarah Jewell and Tho Pham for helpful comments, as well as participants at the 15th International Conference on Computational and Financial Econometrics (CFE 2021; London).

Replication files, code and instructions can be found on Philip Ramirez's GitHub page: https://github.com/philiprami/betting_on_a_buzz.

* Corresponding author.
E-mail addresses: p.ramirez@pgr.reading.ac.uk (P. Ramirez), j.j.reade@reading.ac.uk (J.J. Reade), c.a.singleton@reading.ac.uk (C. Singleton).

[1] See Murphy v. National Collegiate Athletic Association, No. 16-476, 584 U.S. (2018), which ruled that the Professional and Amateur Sports Protection Act of 1992 was unconstitutional. As of April 1, 2021, 11 states have legalised online sports betting: California, Delaware, Illinois, Indiana, Michigan, Nevada, New Hampshire, New Jersey, Pennsylvania, Rhode Island, and West Virginia.

https://doi.org/10.1016/j.ijforecast.2022.07.011

Wikipedia, the free online encyclopaedia, is an example of crowd wisdom. It has become the go-to online place for information about almost anything, including the characteristics and form of sports people.[2] We use this fact to construct what we call the Wiki Relative Buzz Factor, for over 10,000 Women's Tennis Association (WTA) singles matches since the beginning of the 2015 season.[3] These matches were all at the elite level of the sport and include the four annual Grand Slam tournaments. The buzz factor uses the number of page views on the Wikipedia profiles of players before their matches began. We call it 'relative' because it compares the players within a match. We call it 'buzz' because it uses the profile page views on the day before a match in proportion to the typical numbers over the past 12 months. We then adapt the Mincer and Zarnowitz (1969) forecast evaluation framework, showing that the Wiki Relative Buzz Factor can significantly predict the systematic mispricing of bookmaker odds, with the higher-buzz player being underpriced. There is no significant evidence of a favourite or longshot bias in these markets, but bookmakers tended to significantly underprice a player who was substantially lower-ranked than her opponent. Taking these results together, we can reject a sufficient condition for weak-form market efficiency. To prove that these markets are inefficient, we generate probability forecasts of tennis match results by using the same model that detected the mispricing. Combining these forecasts with the Kelly criterion, which can be motivated from expected utility theory, we demonstrate substantial and sustained profits from exploiting the information contained in the Wiki Relative Buzz Factor. Specifically, we found a potential return on investment of 17%–29% from applying the forecasting model at Bet365, the world's highest revenue online sportsbook, to over 5000 potential bets on WTA matches between the beginning of the 2019 season and March of 2020. In contrast, using probability forecasts from the widely used (Elo, 1978) rating systems and the Kelly criterion would have generated substantial losses over the same samples of matches.

These results contribute to the growing literature attempting to elicit the value of crowd wisdom from the field and using this to test the efficient market hypothesis (Fama, 1965, 1970). Relevant to our study of betting markets, research has demonstrated how information from social media can predict what happens in financial markets, including cross-sectional stock returns (e.g., Avery et al., 2016; Chen et al., 2014; Sprenger et al., 2014) and the price movements of cryptocurrencies (Kraaijeveld & De Smedt, 2020). Specifically using Wikipedia, Moat et al. (2013) found that activity on relevant financial pages could provide some early signs of stock market movements. Behrendt et al. (2020) also found that activity on Wikipedia pages could be used to infer collective investor behaviour and design a trading strategy for individual stocks. In a study closely related to our own, Brown et al. (2018) discovered that the aggregate tone extracted from a large number of Twitter posts contained significant information not present in live betting-exchange prices during football matches, especially in the aftermath of major events such as goals or red cards. Using a crowd explicitly making predictions, Brown and Reade (2019) found that the aggregated content from a community of online sports tipsters also contained information not present in betting prices. Betting when the majority of the community predicted a particular outcome generated a small average positive return. Peeters (2018) also found that a crowd of sports fans could improve forecasting accuracy and generate profitable opportunities on betting markets. Specifically, forecasts based on the football player transfer market values on *transfermarkt.de* and the implied strengths of international teams proved more accurate than other standard predictors of match results, such as official team rankings or form-based rating systems.

This paper contributes more generally to the literature on the efficiency of betting and prediction markets, specifically for sports, much of which has focused on the favourite–longshot bias (for reviews, see Williams, 1999, Ottaviani & Sørensen, 2008, or Newall & Cortis, 2021). There is some literature focused on the efficiency of tennis-match betting markets (Abinzano et al., 2016, 2019; Forrest & Mchale, 2007; J., 2014; Lyócsa & Výrost, 2018). This literature has tended to find evidence of a longshot bias that is not large enough to overcome the bookmaker profit margin and prove inefficiency. The present paper also contributes to the use of professional sports to learn about the practice of forecasting, in particular to some studies that have focused on professional tennis (e.g., Angelini et al., 2021a; Barnett & Clarke, 2005; Candila & Scognamillo, 2018; del Corral & Prieto-Rodríguez, 2010; Easton & Uylangco, 2010; Knottenbelt et al., 2012; Kovalchik, 2020; Kovalchik & Reid, 2019; McHale & Morton, 2011; Scheibehenne & Broder, 2007; Spanias & Knottenbelt, 2013). The forecasting models introduced by these studies cannot normally outperform bookmakers without shopping around to find the best available odds (Angelini et al., 2021a; Kovalchik, 2016).

The rest of the paper proceeds as follows: Section 2 describes our dataset, a model to detect mispricing, and a simple betting strategy to test market efficiency using the model; Section 3 presents the results; and Section 4 concludes.

2. Data & method

We collected information from tennis-data.co.uk for all WTA match results from the main draws of all tournaments, including the Grand Slams, between January 1, 2015 and February 16, 2020.[4] This information includes the identity of players and tournaments, as well as when

[2] Wikipedia is the seventh most visited website worldwide; see https://www.similarweb.com/top-websites/, retrieved May 11, 2022.

[3] We have no particular rationale for focusing on this sport and the women's game only. However, it is convenient that odds on all these events were offered by a large number of online sportsbooks. Further, we had built a dataset containing information about these events for other research projects before using it to explore the questions in this paper.

[4] These tennis match data are readily available before 2015, but our analysis period is restricted by the availability of historical Wikipedia page-view data.

(local date) and where matches took place.[5] The dataset represents 10,522 matches, 443 players, and 271 tournaments. It includes the WTA world rankings of the players immediately before each match, which are based on performances over the preceding year and are updated after a tournament is completed. We used the Python packages *geopy* and *timezonefinder* to locate the coordinates of each city in the dataset and the time zones for each match location.

The main draw for a WTA tournament normally takes place a few days before the first round begins, after any qualification matches. All tournaments are in a knock-out format and the draw is seeded, except for the end-of-season WTA Tour finals, which have a round-robin stage. The seedings are generally based on world rankings going into a tournament. The average length of a WTA tennis match in 2020 was 97 min.[6] A player can normally expect one to three days of rest between matches in a tournament. The lineup for a match is usually known at least the day before it starts, after either the first round draw or the completion of players' previous matches in the tournament, at which point betting odds will become available.

We collected betting odds from oddsportal.com for the winner and loser of a match at the time it began. In what follows, we generally use the average odds from the 40 to 60 online bookmakers (sportsbooks) that were posted for any given match on oddsportal.com. We also use the highest (or best) available odds from the bookmaker sample for each match, as well as the specific odds from Bet365, the largest single online bookmaker (sportsbook) in the world by revenue, number of customers, and visitors, which offered odds on almost every match in the dataset.[7]

2.1. The Wikipedia relative buzz factor

To construct a measure of the pre-match buzz about the players, we collected daily (Coordinated Universal Time, UTC) Wikipedia page views of their English-language profiles using the Pageview Application Programming Interface (API), a tool used to query the Wikipedia Foundation page-view data. A small number of observations in the WTA match dataset use maiden names, nicknames, or variations of abbreviations. Therefore, we were careful to ensure that every player in the WTA dataset was matched to her Wikipedia profile page views using manual checking. The mean number of page views for players on the day before a match took place was 1079, with a median of 139, a standard deviation of 6823, and a maximum of 429,245 (for Naomi Osaka on September 7, 2018, the day before she won the US Open final when her opponent, Serena Williams, accused the umpire of being a 'thief'). Panel (a) of Fig. 1 shows kernel density plots of the log profile page views of players the day before a match took place. The distribution for match winners is generally to the right of that for match losers, suggesting that players with higher levels of interest in their profiles before a match were more likely to win. Panel (b) of Fig. 1 shows the tighter distributions of the log daily median page views in the past year before a match, though with greater differences between the winner and loser distributions than in panel (a), suggesting that the typical past number of profile page views could be a better predictor of subsequent success in a match.

To generate our Wiki Relative Buzz Factor for each player–match observation in the dataset, we combine the information contained in panels (a) and (b) of Fig. 1. First, we subtract the log median daily page views of a player over the year before a match from the log page views the day before that match for the same player. Second, we subtract from this value the equivalent value for the player's opponent. As such, our Wiki Relative Buzz Factor measures whether the interest in a player's Wikipedia profile page was atypical the day before a match, and how much it was atypical relative to the player's opponent in the match. Precisely, for player i appearing in match j we calculate:

$$\text{WikiBuzz}_{ij} = \ln(w_{ij}/\widetilde{w}_{ij}) - \ln(w_{-ij}/\widetilde{w}_{-ij}) \,, \tag{1}$$

where w_{ij} is the previous day's page views for the player, $\widetilde{w}_{ij}$ is the median daily page views over the past year before the match, and $-i$ denotes the player's opponent in the match. This measure is plotted in panel (c) of Fig. 1 only for the winning player observations in the dataset. For the match winners, WikiBuzz is on average negative. Thus, when a player receives a greater log increase in daily pre-match page views relative to the typical number received over the previous year, than the player's opponent, it tends on average to predict the player's own defeat in the match (p-value < 0.001). By construction, the Wiki Relative Buzz Factor has zero mean over all winners and losers in the dataset, but we can reject normality with standard tests, due to excess kurtosis of 0.9.

We use the Wikipedia profile page views from the day before the match to construct the buzz factor, instead of from the day of the match, because the daily views are in UTC. If we instead used page views from the day of the match, then we could not be confident that the buzz factor was not caused by the outcome of the match (given that our data only record when each match began in local time), and we could not then use it to form a realistic betting strategy to test market efficiency. Therefore, by converting all times to UTC format and isolating Wikipedia article views from the day prior to the match, we ensure separation between whenever matches started on a particular day and the Wikipedia data. This rules out the potential for leakage of information about the progress or outcome of a match into the period where we observe and use the Wikipedia profile page views of the players involved.

2.2. Detecting mispricing

Let y_{ij} equal one if player $i = 1, 2$ won match $j = 1, \ldots, J$ and zero otherwise, where i distinguishes between the two players in a match and J gives the total

[5] The local date gives the match start, which is important since matches can be played over multiple days due to stoppages, for example, due to the weather.

[6] See http://www.tennisabstract.com/blog/category/match-length/.

[7] See for example https://bestonlinebookmakers.com/largest-bookmakers.html; retrieved June 9, 2022.

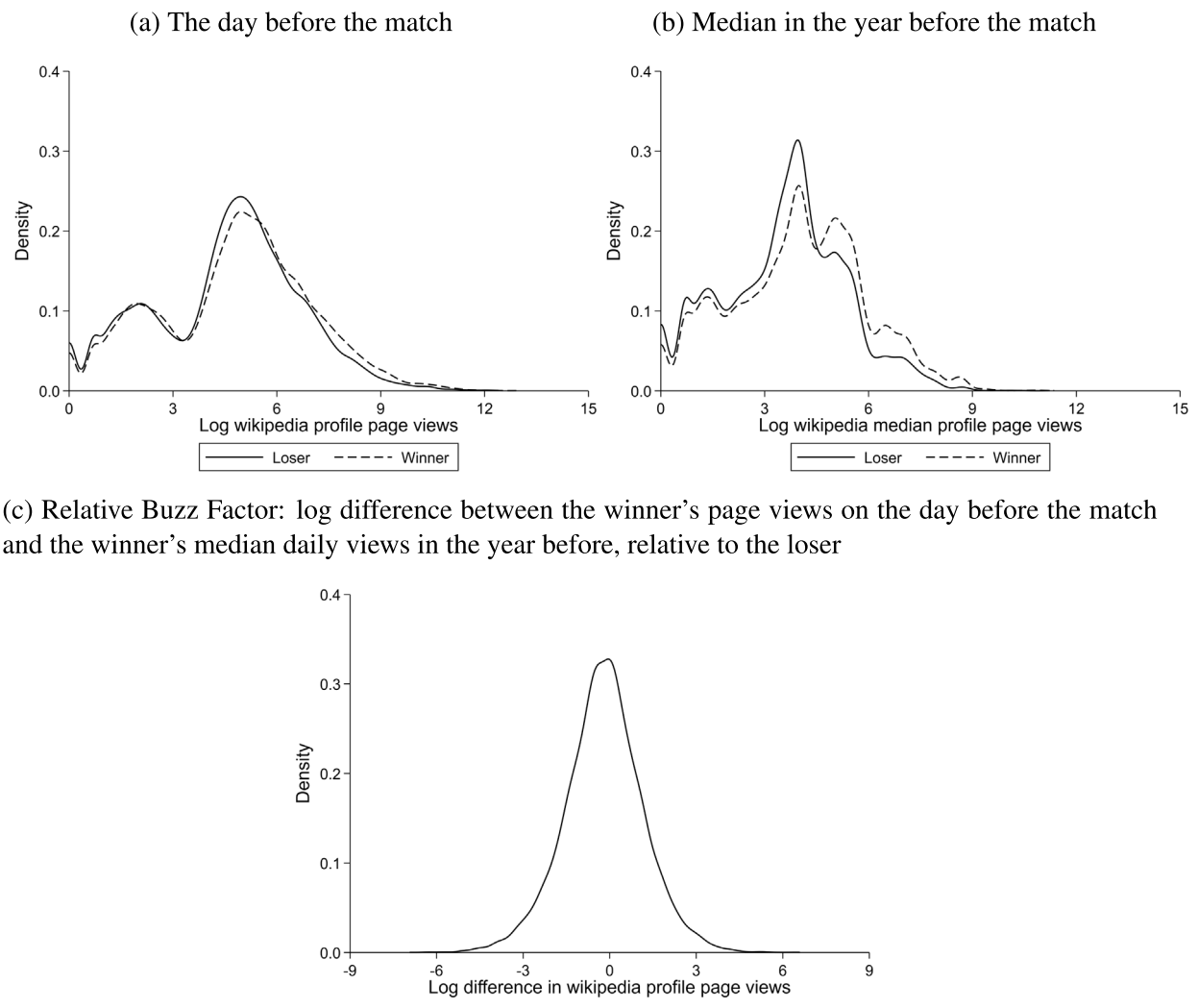

Fig. 1. Wikipedia daily page views of tennis players before WTA matches from 2015–2020. Notes. Author calculations using Wikipedia Foundation page-view data for the English-language profiles of WTAtennis players, collected daily (Coordinated Universal Time (UTC)) using the Pageview Application ProgrammingInterface (API). The densities are estimated with a Gaussian kernel and bandwidth of 0.2.

number of matches in a sample, such that the overall sample contains $2J$ player–match observations. Let p_j be the unobserved beliefs of the bookmaker about the probability of $y_{1j} = 1$ happening beforehand, i.e., player 1 winning match j. The bookmaker offers decimal odds o_{ij} on the two potential outcomes, meaning that, on taking a £1 bet, they return o_{ij} to the bettor if the outcome happens, and they gain £1 if it does not. Let $z_{ij} = 1/o_{ij}$ be the inverse odds or implied odds-based probability forecast of the bookmaker. For any match, $z_{1j} + z_{2j} = 1 + \kappa_j > 1$, where κ_j has often in the literature been termed as the bookmaker's expected rate of commission or profit margin on a match, also known among sports bettors as the 'overround' or 'vig'. This implies $z_{1j} = p_j + \alpha\kappa_j$ and $z_{2j} = (1 - p_j) + (1 - \alpha)\kappa_j$. If we denote $e_{ij} = y_{ij} - z_{ij}$, then an efficient bookmaker market requires that forecast errors on average are equal to the negative value of some sample average 'overround', $E_{ij}\left[e_{ij}\right] = -\bar{\kappa}$. In other words, the bookmaker is efficient if it makes some average level of commission across matches and outcomes, and no other information can predict e_{ij}, since it would already be priced into the odds.

We consider three potential sources of mispricing and departures from the efficient market hypothesis in WTA betting markets.

(1) Favourite–longshot bias: There is an empirical irregularity in some prediction and betting markets known as the favourite–longshot bias. When it has been used in the academic literature, this term most typically equates to a longshot bias, whereby the odds offered by bookmakers suggest an underestimation by the market about the chances of the most expected outcomes happening over the least expected outcomes, making bets on favourites generally more profitable than bets on longshots (see the summaries by Ottaviani & Sørensen, 2008 and Newall & Cortis, 2021). Many studies of professional sports betting markets have identified such a longshot bias, including the seminal study on horseracing by Ali (1977).

Several theoretical contributions, which could be broadly classified as coming from neoclassical economic theory, have demonstrated the sufficient conditions such that this longshot bias can arise in equilibrium, in terms of the preferences, budget constraints, and distribution of beliefs among the market participants (e.g., He & Treich, 2017; Manski, 2006; Ottaviani & Sørensen, 2015). The same theoretical frameworks also suggest that high risk aversion among bettors can lead to the bias reversing toward the favourite outcome in the market, which is often termed reverse favourite–longshot bias or just favourite bias. Besides the predictions from neoclassical economic theory, a competing set of behavioural explanations has been proposed to explain the favourite–longshot bias, which emphasises the misperception of probabilities by bettors (e.g., Snowberg & Wolfers, 2010; Vaughan Williams et al., 2018). Newall and Cortis (2021) suggest from their review of the empirical literature that sports markets with fewer potential outcomes tend to produce a favourite bias (e.g., team sports or tennis), whereas a longshot bias appears in markets with many outcomes (e.g., horseracing or golf). Nevertheless, previous studies of professional tennis have found a longshot bias (e.g., Abinzano et al., 2016, 2019; Forrest & Mchale, 2007; J., 2014), though not sufficient to suggest market inefficiency through positive mean returns from consistently betting on match favourites (e.g., Forrest & Mchale, 2007; Lyócsa & Výrost, 2018).

(2) Player ranking bias: We consider whether tennis betting markets systematically misprice the outcome of a match according to player rankings. Several studies have demonstrated how the recent performances of tennis players can provide relatively accurate forecasts compared with those implied by bookmaker odds as a benchmark, typically through enhanced (Elo, 1978) ratings (e.g., Angelini et al., 2021a; Kovalchik, 2020; Kovalchik & Reid, 2019). We use standard and more advanced Elo ratings below to provide benchmark probability forecasts of match results.[8] There is some suggestive evidence that bookmakers are more risk averse in tennis matches involving lower-ranked players, and the longshot bias thus increases in these cases (Abinzano et al., 2016; J., 2014). The WTA world rankings are ordered from one, for the best player cumulatively over the past year, to having no rank, for a player who has not earned enough points at WTA events over the past year to get one. We consider two measures based on these rankings. First, we consider the raw rank difference between the players in a match, $\text{RankDiff}_{ij} = rank_{ij} - rank_{-ij}$. Second, we assume that the performance difference between two consecutive players in the rankings is decreasing more so as one goes down the ranking list from the top. The difference in ability between the first- and second-ranked players is likely to be more than between the 100th- and 101st-ranked players, which can be evidenced by how much less often player rankings move at the top compared with the bottom. We construct a ranking distance measure for player i in match j as:

$$\text{RankDist}_{ij} = -\left(\frac{1}{rank_{ij}} - \frac{1}{rank_{-ij}}\right), \tag{2}$$

where we impute $1/rank_{ij} = 0$ if a player was unranked at the time of a match. RankDist_{ij} is bounded by –1, when the player considered is ranked first in the world and is playing somebody unranked, and 1, when it is the other way around, thus having the same sign interpretation as RankDiff_{ij}.

(3) Wikipedia Relative Buzz Factor bias: To the best of our knowledge, this sort of information has not been used to predict the outcome of tennis matches and the efficiency of their betting markets, or at least this has not been documented before. However, there are parallels with studies using information from social media and player evaluations to predict football match outcomes and betting inefficiencies (e.g., Brown et al., 2018; Peeters, 2018).

To detect mispricing and estimate the conditional mean effects on bookmakers' odds-implied probability forecast errors, we apply the general (Mincer & Zarnowitz, 1969) forecast evaluation framework (see Angelini & L., 2019, Angelini et al., 2021b, and Elaad et al., 2020, who tested for home bias, the favourite–longshot bias, and the weak-form efficiency of European football betting markets in much the same way). We estimate the following using least squares:

$$e_{ij} = \alpha + \beta_1 z_{ij} + \beta_2 \text{RankDist}_{ij} + \beta_3 \text{WikiBuzz}_{ij} + \psi_{S(j)} + \phi_{T(j)} + \varepsilon_{ij}, \tag{3}$$

where $\{\alpha, \beta_1, \beta_2, \beta_3, \psi_{S(j)}, \phi_{T(j)}\}$ are parameters. We expect a significantly negative estimate of α to capture the bookmaker's profit margin (overround). Positive values of β_1, β_2, or β_3 would respectively suggest a longshot bias, a high-rank bias, or a low-buzz bias in the markets, such that betting on a win by the favourite, the lower-ranked player, or the one with greater pre-match relative buzz could be profitable strategies, and vice versa if these parameters are negative. We also consider fixed effects in Eq. (3) for the season (year), $\psi_{S(j)}$, and tournament of the match, $\phi_{T(j)}$, where $S(j)$ and $T(j)$ are indicator functions, to address the potential heterogeneity over these dimensions in bookmaker overrounds or expected profit margins. The remaining heterogeneity is left in the residual term ε_{ij}. We construct standard errors for the estimates of Eq. (3) that are robust to clusters at the match and tournament levels. This addresses the heteroskedasticity from including both players in a match in the estimation sample, as well as the possibility that some tournaments may be less predictable than others.[9]

The mean of e_{ij} will be significantly negative for any reasonably sized sample of matches. Therefore, a sufficient condition for the betting market to be weak-form

[8] The Elo ratings are computed using all WTA tennis matches between the beginning of the 2007 season and March 2020.

[9] As a robustness check, we also considered estimates of Eq. (3) using weighted least squares, with elements of the diagonal weighting matrix approximated by $z_{1j} \times z_{2j}$, as suggested by Angelini and L. (2019). Although this estimator reduces the influence of more competitive matches, the results that follow are robust to using this instead of ordinary least squares.

efficient, according to Eq. (3), is given by the null hypothesis: $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$. If we find that estimates of $\{\beta_1, \beta_2, \beta_3\}$ are significantly positive or negative, then the associated variables provide information that is not fully incorporated in the pre-event prices. In this case, the markets may be inefficient if bettors can use the same information to make sustained positive returns.

2.3. Market inefficiency and a simple betting strategy

To test whether the bookmaker markets are inefficient, we use estimation results of the mispricing model in Eq. (3), an out-of-sample dataset of tennis matches, bookmaker odds and Wikipedia page-view data, and the Kelly (1956) criterion. This criterion is the solution to a bettor's maximisation problem on how much of the bettor's wealth should be invested in the claim offered by the bookmaker, assuming logarithmic utility and given the bettor's beliefs about the outcome of the claim and the odds posted by the bookmaker. Along with simpler strategies, such as 'bet one unit when the expected return is positive', the Kelly criterion has been widely used in the literature to evaluate betting market efficiency (e.g., Hvattum & Arntzen, 2010; Peeters, 2018; Ziemba, 2020). We assume that our bettor in this case forms expectations from estimating Eq. (3) using ordinary least squares (OLS), though without including the season or tournament fixed effects in the model, as these are impractical for forecasting. The other variables in Eq. (3) are all available to the bettor before a tennis match begins, allowing the bettor to use the estimated model to form probability forecasts of match outcomes. The bettor's out-of-sample expected probability of winning a bet on event i, a specific player to win match j, denoted by $\widetilde{y}_{ij}$, is thus given by:

$$\widetilde{y}_{ij} = \widehat{\alpha} + (1 + \widehat{\beta}_1) z_{ij} + \widehat{\beta}_2 \text{RankDist}_{ij} + \widehat{\beta}_3 \text{WikiBuzz}_{ij} \ , \quad (4)$$

where $\{\widehat{\alpha}, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3\}$ are in-sample OLS estimators. The Kelly criterion gives the share of a fixed amount of wealth or budget to invest in each bet, remembering that $o_{ij} = 1/z_{ij}$ are the decimal odds offered:

$$x_{ij} = \max\{\widehat{y}_{ij} - \frac{1 - \widehat{y}_{ij}}{o_{ij} - 1} \ , \ 0\} \ . \quad (5)$$

The bettor's return on investment (ROI) over $N = 2J$ potential bets, expressed as a percentage of the total amount invested over the sample period (some multiple of the per-bet budget), is given by:

$$\text{ROI} = \frac{\sum_{ij}^{2J} \left(x_{ij} o_{ij} \mathbb{1}\{y_{ij} = 1\} - x_{ij} \mathbb{1}\{y_{ij} = 0\} \right)}{\sum_{i}^{2J} x_{ij}} \ . \quad (6)$$

A substantially positive ROI, over a large out-of-sample number of matches, would provide evidence that tennis match betting markets are weak-form inefficient, due to some combination of the biases captured by the model. This would suggest that the relatively straightforward model and betting strategy could be applied profitably in real time. To provide benchmark ROIs, we construct alternative estimates of $\widetilde{y}_{ij}$ using the standard player form-based (Elo, 1978) ratings, with an updating factor (K-factor) of 20, and using all WTA match results since the beginning of the 2007 season. We also use the more sophisticated W-Elo forecasting model from Angelini et al. (2021a).

3. Results

3.1. Mispricing

Table 1 shows the results of estimating Eq. (3) for an in-sample period of the 2015–2018 WTA seasons, using as the dependent variable the value of the prediction error according to the mean pre-match odds offered by the K_j (normally 40–60) individual bookmakers ($k = 1, \ldots, K_j$) listed by oddsportal.com for any given match: $\bar{e}_{ij} = y_{ij} - \sum_k^{K_j} \left(z_{ijk}/K_j \right)$. Column (I) only tests for a favourite–longshot bias. We find on average a marginal favourite bias, but this is not statistically significant. Column (II) adds the difference in the pre-match WTA rankings of the players, RankDiff_{ij}, as a regressor, which is also not statistically significant. When taken together with the favourite–longshot bias, the null $H_0 : \beta_1 = \beta_2 = 0$ cannot be rejected, and there is no evidence that bookmaker betting markets for WTA tennis matches are mispriced according to the raw difference in ranks and the balance of the odds between players.

In column (III) of Table 1, we replace RankDiff_{ij} with our alternative measure of the rank distance between players, RankDist_{ij}. This measure significantly predicts the average bookmaker odds-implied forecast errors (p-value $= 0.035$), and the null $H_0 : \beta_1 = \beta_2 = 0$ can be rejected at the 10% level. The model estimates suggest that the probability of an unranked player winning against the number one ranked player in the world is 0.061 greater than what bookmaker odds tend to imply. In column (IV), we add the third potential source of mispricing to the model in the form of the Wiki Relative Buzz Factor. This measure positively and significantly predicts the average bookmaker odds-implied forecast errors (p-value $= 0.030$). As mentioned before, on average the player with a relatively larger pre-match increase in Wikipedia profile page views tends to lose a tennis match. However, the model estimates show that bookmaker odds generally imply a further under-prediction of that player's chances, making her more of a longshot or less of a favourite than she ought to be according to the Wiki Relative Buzz Factor and conditional on the other variables in the model. After including this source of mispricing in the model, the estimated rank distance mispricing remains positive and significant at the 10% level. In this specification, there is a small conditional longshot bias, consistent with the previous literature (Abinzano et al., 2016, 2019; Forrest & Mchale, 2007; J., 2014), though here it is not statistically significant. We can also reject the sufficient condition for weak-form market efficiency, $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ at the 5% level. In column (V), we add tournament fixed effects to the regression model: the estimates and test results are practically the same.[10] Table 2 shows results comparable to Table 1 after adding to the estimation samples matches from the 2019 and 2020 (before March) WTA seasons, which we use below for the out-of-sample

[10] We checked for misspecification of Eq. (3) using Ramsey RESET tests and did not reject the null hypothesis; the data generating process was not better approximated by including squared terms for any of the regressors.

Table 1
Model estimates and tests of betting market mispricing for WTA match results, 2015–2018: in-sample period only.

	(I)	(II)	(III)	(IV)	(V)
Odds-implied probability	−0.022	−0.061	0.002	0.025	0.025
	(0.024)	(0.040)	(0.027)	(0.029)	(0.029)
WTA rank diff. (player–opponent)		−0.013			
		(0.009)			
WTA rank distance to opponent			0.061**	0.055*	0.055**
			(0.029)	(0.029)	(0.029)
Wiki Relative Buzz Factor				0.009**	0.009**
				(0.004)	(0.004)
Constant	−0.018	0.002	−0.031**	−0.043***	−0.042***
	(0.013)	(0.022)	(0.014)	(0.015)	(0.015)
Year/season fixed effects	Yes	Yes	Yes	Yes	No
Tournament fixed effects	No	No	No	No	Yes
F-test: $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$		0.319	0.070	0.022	0.022
N of player-matches	15,854	15,826	15,854	15,854	15,854

Notes. ***, **, and * indicate significance from zero at 1%, 5%, and 10% levels, respectively, two-sided tests. Standard errors in parentheses were estimated robust to both match- and tournament-level clusters.
Column (I): linear regression estimates of Eq. (3), where the dependent variable is the forecast error implied by average bookmaker odds (oddsportal.com)—test of favourite–longshot bias.
Column (II): adds the pre-match raw WTA rank difference to the model in (I).
Column (III): uses the alternative differences in ranks measure described in the text—the coefficient effect should be interpreted as an unranked player against the number one ranked in the world, relative to two hypothetically equally ranked players.
Column (IV): adds the Wiki Relative Buzz Factor—preferred results.
Column (V): adds tournament fixed effects to the model in (IV).

Table 2
Model estimates and tests of betting market mispricing for WTA match results, 2015–2020: full sample period.

	(I)	(II)	(III)	(IV)	(V)
Odds-implied probability	−0.009	−0.040	0.013	0.036	0.036
	(0.020)	(0.034)	(0.022)	(0.024)	(0.024)
WTA rank diff. (player–opponent)		−0.010			
		(0.008)			
WTA rank distance to opponent			0.054**	0.049**	0.049**
			(0.024)	(0.024)	(0.024)
Wiki Relative Buzz Factor				0.009**	0.009**
				(0.003)	(0.003)
Constant	−0.025**	−0.009	−0.037***	−0.049***	−0.047***
	(0.011)	(0.018)	(0.012)	(0.013)	(0.013)
Year/season fixed effects	Yes	Yes	Yes	Yes	No
Tournament fixed effects	No	No	No	No	Yes
F-test: $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$		0.436	0.076	0.016	0.016
N of player-matches	21,044	20,992	21,044	21,044	21,044

Notes. ***, **, and * indicate significance from zero at 1%, 5%, and 10% levels, respectively, two-sided tests. Standard errors in parentheses were estimated robust to both match- and tournament-level clusters.
See Table 1. Each column's model estimates are equivalent to the respective columns in Table 1, but here matches from the 2019 and 2020 WTA seasons are included in the estimation samples.

forecasting and market efficiency analysis. All of the mispricing test results are robust to extending the sample period in this way.

Heterogeneity in match location and time differences could perhaps be relevant to the impact of the Wikipedia Relative Buzz Factor. To address this, column (I) of Table 3 repeats the model estimates from column (IV) of Table 1, and then columns (II)–(IV) show results after cumulatively dropping from the estimation sample matches in time zones from the east, starting with UTC+11&12 (Sydney/Auckland), then UTC+11&12 (Seoul/Tokyo), and finally UTC+7&8 (Singapore/Hong Kong). The influence of the Wiki Relative Buzz Factor and the rejection of the sufficient condition of weak-form efficiency are robust to dropping these matches from the estimation sample. After dropping matches from all six of the most eastern time zones in the dataset, the mispricing in odds predicted by the buzz factor is greater. This suggests that the Wikipedia profile page views less than 24 h before the start of a match may be less useful in predicting odds mispricing. This would be consistent with the buzz factor being a proxy for crowd judgements on the relative strengths of players' most recent performances within a tournament. To test whether this could alone explain why the buzz factor can predict bookmaker mispricing, in column (V) of Table 3 we re-estimate the model only for matches in the first round of tournaments. The coefficient on the Wiki Relative Buzz Factor remains marginally significant (p-value = 0.069) and is larger than when it is estimated over all matches in tournaments. This suggests that the

Table 3
Model estimates and tests of betting market mispricing for WTA match results, 2015–2018: preferred model and dropping time zones, and first-round matches only.

	(I)	(II)	(III)	(IV)	(V)	(VI)
Odds-implied probability	0.025	0.036	0.037	0.043	0.053	0.077
	(0.029)	(0.030)	(0.032)	(0.035)	(0.036)	(0.057)
WTA rank distance to opponent	0.055*	0.061**	0.058*	0.023	0.108*	0.123
	(0.029)	(0.030)	(0.030)	(0.031)	(0.064)	(0.089)
Wiki Relative Buzz Factor	0.009**	0.010**	0.010**	0.014***	0.012*	0.017
	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.010)
Constant	−0.043***	−0.049***	−0.050***	−0.053***	−0.059***	−0.071**
	(0.015)	(0.016)	(0.018)	(0.018)	(0.019)	(0.030)
Drop UTC+11&12	No	Yes	Yes	Yes	No	No
Drop UTC+9&10	No	No	Yes	Yes	No	No
Drop UTC+7&8	No	No	No	Yes	No	No
Year/season fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F-test: $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$	0.022	0.017	0.019	0.027	0.069	0.165
N of player-matches	15,854	14,448	13,620	11,358	7208	3914

Notes. ***, **, and * indicate significance from zero at 1%, 5%, and 10% levels, respectively, two-sided tests. Standard errors in parentheses were estimated robust to both match- and tournament-level clusters.
Column (I): repeats the preferred model estimates from column (IV) of Table 1.
Columns (II)–(IV): each column drops matches from an additional two time zones, starting with UTC+11&12 (Sydney/Auckland) and finally in column (IV) dropping UTC+7&8 (Singapore/Hong Kong).
Column (V): estimates the preferred model from column (I) here but only using matches from the first round of tournaments.
Column (VI): also drops from the estimation sample of column (V) any first-round matches which involved a player with a world ranking greater than 100 at the time (i.e., players who were very likely to have come through qualifying rounds in the previous week).

mispricing is not only driven by whatever happened in the previous round of a tournament, which may have generated interest in a player's Wikipedia profile page. As a further robustness check in this regard, in column (VI) we estimate the model using only first-round matches involving players who had a ranking no greater than 100 and, therefore, were less likely to have come through qualifying rounds in the previous week before entering the main draws of tournaments.[11] In this smaller sample of matches, the coefficient estimate for the Wiki Relative Buzz Factor is even larger than in the previous specifications, but it is also less precisely estimated and thus statistically insignificant at standard levels.

In summary, the results from estimating Eq. (3), and the tests of mispricing by bookmakers, suggest that there might be inefficiencies in the final-result markets of tennis matches. These inefficiencies could be proven by betting on players who are substantially lower-ranked than their opponents or whose Wikipedia profiles show unusually high interest before matches.

3.2. Market inefficiency and the betting strategy

Table 4 shows the results of applying the simple betting strategy described in Section 2.3, by using match outcome probability predictions according to Eq. (4) and applying the Kelly criterion. We estimated the model up to the end of the 2018 season, used this to forecast match outcomes in the 2019 and 2020 seasons, and then applied the Kelly criterion with these forecasts. Column (I) of Table 4 shows the results of the betting strategy for a hypothetical bettor who could place bets at the average pre-match odds offered by the 40–60 bookmakers sampled for each match. The average overround in these markets in 2019 and 2020 (before March) was 5.3%. The out-of-sample probability forecasts and Kelly criterion results suggest betting on 221 of the 5190 considered odds (2595 WTA matches in the period), with a total amount invested equal to 4.3 times the per-bet budget and a return on investment (ROI) of −6.4%, which is no better than the average bookmaker overround.

For curiosity, column (II) of Table 4 presents results whereby the model was estimated and predictions were made using average odds but the best available out-of-sample odds listed on oddsportal.com were used in the Kelly criterion. In this case, a much greater proportion of matches are bet on, the total amount invested over the sample period is 76.6 times the per-bet budget, and the ROI is 3.1%. However, despite the existence of 'odd-schecker' websites being available to the bettor, using the best available odds just before a match begins is not normally realistic, due to the transaction costs and time involved with managing a large number of online accounts. Further, there are restrictions that can prevent a bettor from obtaining the best available odds listed on oddsportal.com, such as the location of a bettor affecting which online sportsbooks can be used. This is evidenced by the average overround according to the best available odds being negative in the 2019 and 2020 WTA seasons, suggesting that theoretical arbitrage opportunities were common if not entirely practical.

As a more realistic test of bookmaker inefficiency, column (III) of Table 4 presents results from using the Kelly criterion and the odds from only one online sportsbook. We selected Bet365 because it is the highest-revenue sportsbook in the world and had odds listed on oddsportal.com for almost every WTA match since 2015. From using the model's predictions and the Bet365 odds, we find an out-of-sample ROI of 17.3%, which is generated

[11] A refinement to this robustness check could less conservatively exclude only matches that exactly included qualifiers, after collecting data on the qualifying events, not least because some 'wildcard' players with a ranking greater than 100 could have entered the tournament directly.

Table 4
Out-of-sample betting strategy results for WTA match results, 2019–2020.

	Average	Best	Bet365			
	(I)	(II)	(III)	w/out rank (IV)	Elo (V)	W-Elo (VI)
N odds ($2 \times J$ matches)	5190	5188	5156	5156	4796	4362
Number of bets placed	221	2350	312	276	1778	2058
Mean overround (%)	5.33	−0.23	6.46	6.46	6.48	6.49
Investment (×per bet budget)	4.30	76.63	7.15	4.99	295.36	941.39
Absolute return (×per bet budget)	−0.27	2.34	1.24	1.44	−36.16	−116.50
Return on investment (%)	−6.37	3.05	17.26	28.82	−12.24	−12.38

Notes. 'Out-of-sample' uses the model from column (IV) of Table 1 estimated on matches up to the end of the 2018 season, then uses it to predict match outcomes and apply the Kelly criterion for the 2019 and 2020 seasons. Average odds are always used to estimate the models and generate forecasts, but the odds used in the Kelly criterion are varied.
Column (I): uses the reported average pre-match available odds from oddsportal.com.
Column (II): uses the reported best available pre-match odds from oddsportal.com.
Column (III): uses pre-match odds from Bet365.
Column (IV): uses Bet365 odds but with a version of the preferred model estimated without the rank distance variable.
Column (V): uses Bet365 odds but with the standard Elo predicted probability forecast of the match outcome.
Column (VI): uses Bet365 odds but with the W-Elo predicted probability forecast of the match outcome, as per (Angelini et al., 2021a).

from placing bets according to the criterion on 12% of the main draw WTA matches between the beginning of 2019 and March 2020, equivalent to investing 7.15 times the per-bet budget. To check whether these profitable opportunities are driven by the Wiki Relative Buzz Factor, we drop the rank distance measure from the model estimation, with the results shown in column (IV). In this case, fewer matches are bet on and less money is invested according to the Kelly criterion, but the ROI is increased to 28.8% and the absolute return is also greater. To provide a meaningful benchmark ROI using an alternative probability forecasting model of match results, also applied with the Kelly criterion, the same samples of matches, and the Bet365 odds, column (V) of Table 4 shows results using the standard (Elo, 1978) ratings model described in Section 2.3. The ROI from applying the betting strategy with this alternative set of probability forecasts is −12.2%. This model would also have led to substantial amounts of betting activity and absolute losses over the sample period, because of the frequency and magnitude of differences between the simple Elo-predicted probabilities of match outcomes and what bookmaker odds imply, particularly leading to over-betting on longshots.[12] As a further comparison, column (VI) shows betting results using W-Elo, which is a more sophisticated Elo forecasting model of tennis match results that reflects contributions by Kovalchik (2016) and Angelini et al. (2021a). This model gives greater weight to past match wins at prestigious tournaments and takes into account the margins of victory that players achieved.[13] However, the W-Elo model predictions, applied with the Kelly criterion and Bet365 odds, generate a marginally worse ROI and over three times greater absolute losses in our out-of-sample period compared to the standard Elo model in column (V).

Finally, we check whether the betting returns from using the Wiki Relative Buzz Factor are driven by subsets of matches expected to be more or less competitive by bookmakers. We estimate the same model over the 2015–2018 seasons and follow the same betting strategy as in column (IV) of Table 4, which yielded an out-of-sample ROI of 28.8%, except we consider matches in particular odds ranges. The results in Table 5 show that applying the model and betting strategy over matches with intermediate odds, i.e., matches expected to be relatively competitive, generates a marginally higher ROI and a substantially higher absolute return than applying it over all matches, and a substantially higher ROI than applying it over matches expected to be relatively uncompetitive. In this way, the Wiki Relative Buzz factor tends to be a stronger predictor of bookmaker mispricing when matches are expected to be more competitive, and the players involved are by implication more similar in their ability or form.

In summary, a buzz factor about tennis players, constructed from their Wikipedia profile page-view data, provides relevant information that is not being fully incorporated into the match-result prices offered by bookmakers. This information can be used to generate sustained and substantial profits when used in a relatively simple betting strategy.

4. Conclusion

In this paper, we constructed a measure of relative pre-match buzz about tennis players using Wikipedia profile page-view data. We found that this Wikipedia Relative Buzz Factor can predict bookmaker odds-implied forecast errors and the significant mispricing of outcomes, suggesting profitable opportunities for bettors who back players with relatively greater buzz than their opponents going into a match. Using these results to forecast outcome probabilities and the Kelly criterion to select how

[12] This is not necessarily an indictment of Elo ratings for tennis forecasting and betting. Angelini et al. (2021a) show that standard Elo ratings, a more conservative and sophisticated betting strategy, and the best available odds from a sample of bookmakers, can be used to generate positive betting returns for elite tennis matches.

[13] To generate these ratings, we use an R package associated with Angelini et al. (2021a), *welo* (Candila, 2021). When calculating the W-Elo ratings, we restrict the data to only players who played at least 10 WTA matches since the beginning of 2007, hence the reduced number of odds considered in the betting strategy analysis. The parameters are set to those preferred by Angelini et al. (2021a): player starting points of 1500, Kovalchik (2016) scale factors, and weights based on the number of games won rather than sets.

Table 5
Out-of-sample betting strategy results for WTA match result, 2019–2020: selecting sample odds based on match competitiveness.

	Bet365 odds			
	(I)	(II)	(III)	(IV)
N odds ($2 \times J$ matches)	732	3459	4424	1697
Number of bets placed	4	87	363	263
Mean overround (%)	5.71	6.02	6.58	7.01
Investment (×per bet budget)	0.05	1.03	7.25	9.27
Absolute return (×per bet budget)	−0.002	0.008	1.46	2.72
Return on investment (%)	−3.02	0.81	20.11	29.38

Notes. Betting strategy results equivalent to column (IV) of Table 4, varying the sample of match odds used in estimating the model and considered for bets by column. Average odds are always used to estimate the models and generate forecasts, but Bet365 odds are used in the Kelly criterion.
Column (I): uses only odds in the sample which imply a match win probability of $p \in (0, 0.2) \cup (0.8, 1)$.
Column (II): uses only odds in the sample which imply a match win probability of $p \in (0, 0.4) \cup (0.6, 1)$.
Column (III): uses only odds in the sample which imply a match win probability of $p \in [0.2, 0.8]$.
Column (IV): uses only odds in the sample which imply a match win probability of $p \in [0.4, 0.6]$.

much to bet on what matches, we found that tennis result betting markets are inefficient. Prices do not fully incorporate the information contained in the buzz factor. The returns on investment from applying the model and betting strategy were sustained and substantial, including when using only the odds of Bet365, the world's highest revenue online sportsbook. Two previous studies also found that online information representing the wisdom of crowds can be used to form profitable betting strategies, though with much smaller rates of return than we found in tennis markets (Brown & Reade, 2019; Peeters, 2018). However, it is unclear whether correcting these sources of inefficiency would result in greater profits for bookmakers. What we labelled as mispricing may correlate with unobserved biases and heterogeneity among bettors that bookmakers exploit when setting odds.

There are two natural extensions to this research. The 'wisdom of crowds' might explain why a measure constructed from Wikipedia page-view data can predict bookmaker mispricing. While this is an appealing and plausible explanation, we have done nothing here to prove it. This would require complementary data sources that capture explicit predictions about tennis match outcomes or evaluations of the players, like the crowd-sourced football transfer market values used by Peeters (2018). The Wikipedia Relative Buzz Factor may only be capturing relative changes in the media interest in tennis players before matches. If that were the case, then our results could perhaps be described more accurately as being driven by the 'wisdom of the media', or by a small number of tennis commentators and pundits who selectively draw attention to some players over others. Second, we can think of no good reason why the betting market inefficiencies found here would be constrained to the top level of women's professional tennis. It would be interesting for others to check whether these results apply to tennis below the WTA level, men's tennis, or entirely different sports. To this end, we have provided readily adaptable replication code and instructions for all our results on a GitHub page: https://github.com/philiprami/betting_on_a_buzz.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abinzano, I., Muga, L., & Santamaria, R. (2016). Game, set and match: the favourite-long shot bias in tennis betting exchanges. *Applied Economics Letters*, *23*(8), 605–608.
Abinzano, I., Muga, L., & Santamaria, R. (2019). Hidden power of trading activity: The FLB in tennis betting exchanges. *Journal of Sports Economics*, *20*(2), 261–285.
Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, *85*(4), 803–815.
Angelini, G., Candila, V., & L., De Angelis. (2021a). Weighted Elo rating for tennis match predictions. *European Journal of Operational Research*, *297*(1), 120–132.
Angelini, G., De Angelis, L., & C., Singleton. (2021b). Informational efficiency and behaviour within in-play prediction markets. *International Journal of Forecasting*, *38*(1), 282–299.
Angelini, G., & L., De Angelis. (2019). Efficiency of online football betting markets. *International Journal of Forecasting*, *35*(2), 712–721.
Avery, C. N., Chevalier, J. A., & Zeckhauser, R. J. (2016). The CAPS prediction system and stock market returns. *Review of Finance*, *20*(4), 1363–1381.
Barnett, T., & Clarke, S. R. (2005). Combining player statistics to predict outcomes of tennis matches. *IMA Journal of Management Mathematics*, *16*(2), 113–120.
Behrendt, S., Peter, F. J., & Zimmermann, D. J. (2020). An encyclopedia for stock markets? Wikipedia searches and stock returns. *International Review of Financial Analysis*, *72*, Article 101563.
Brown, A., Rambaccussing, D., & Rossi, G. (2018). Forecasting with social media: Evidence from tweets on soccer matches. *Economic Inquiry*, *56*(3), 1748–1763.
Brown, A., & Reade, J. J. (2019). The wisdom of amateur crowds: Evidence from an online community of sports tipsters. *European Journal of Operational Research*, *272*(3), 1073–1081.
Candila, V. (2021). *Welo: Weighted and standard Elo rates. R package version 0.1.0*.
Candila, V., & Scognamillo, A. (2018). Estimating the implied probabilities in the tennis betting market: A new normalization procedure. *International Journal of Sport Finance*, *13*(3), 225–242.
Chen, H., De, P., Hu, Y. J., & 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.

del Corral, J., & Prieto-Rodríguez, J. (2010). Are differences in ranks good predictors for Grand Slam tennis matches? *International Journal of Forecasting*, *26*(3), 551–563.
Easton, S., & Uylangco, K. (2010). Forecasting outcomes in tennis matches using within-match betting markets. *International Journal of Forecasting*, *26*(3), 564–575.
Elaad, G., Reade, J. J., & Singleton, C. (2020). Information, prices and efficiency in an online betting market. *Finance Research Letters*, *35*, Article 101291.
Elo, A. E. (1978). *The rating of chessplayers, past and present*. London Batsford.
Fama, E. F. (1965). The behavior of stock-market prices. *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.
Forrest, D. (2008). Soccer betting in britain. In D. B. Hausch, & W. T. Ziemba (Eds.), *Handbook of sports and lottery markets* (pp. 421–446). San Diego Elsevier.
Forrest, D., & Mchale, I. (2007). Anyone for tennis (betting)? *The European Journal of Finance*, *13*(8), 751–768.
Galton, F. (1907). *Vox populi*.
He, X.-Z., & Treich, N. (2017). Prediction market prices under risk aversion and heterogeneous beliefs. *Journal of Mathematical Economics*, *70*(C), 105–114.
Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, *26*(3), 460–470.
J., Lahvička (2014). What causes the favourite-longshot bias? Further evidence from tennis. *Applied Economics Letters*, *21*(2), 90–92.
Kelly, J. L. (1956). A new interpretation of information rate. *The Bell System Technical Journal*, *35*(4), 917–926.
Knottenbelt, W. J., Spanias, D., & Madurska, A. M. (2012). A common-opponent stochastic model for predicting the outcome of professional tennis matches. *Computers & Mathematics with Applications*, *64*(12), 3820–3827.
Kovalchik, S. A. (2016). Searching for the GOAT of tennis win prediction. *Journal of Quantitative Analysis in Sports*, *12*(3), 127–138.
Kovalchik, S. (2020). Extension of the elo rating system to margin of victory. *International Journal of Forecasting*, *36*(4), 1329–1341.
Kovalchik, S., & Reid, M. (2019). A calibration method with dynamic updates for within-match forecasting of wins in tennis. *International Journal of Forecasting*, *35*(2), 756–766.
Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, *65*(C).
Lyócsa, Štefan, & Výrost, T. (2018). To bet or not to bet: A reality check for tennis betting market efficiency. *Applied Economics*, *50*(20), 2251–2272.
Manski, C. (2006). Interpreting the predictions of prediction markets. *Economics Letters*, *91*(3), 425–429.
McHale, I., & Morton, A. (2011). A Bradley–Terry type model for forecasting tennis match results. *International Journal of Forecasting*, *27*(2), 619–630.
Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 1–46). NBER.
Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves. *Scientific Reports*, *3*(1), 1–5.
Newall, P. W. S., & Cortis, D. (2021). Are sports bettors biased toward longshots, favorites, or both? A literature review. *Risks*, *9*(1), 22.
Ottaviani, M., & Sørensen, P. N. (2008). The favorite-longshot bias: An overview of the main explanations. In D. B. Hausch, & W. T. Ziemba (Eds.), *Handbook of sports and lottery markets* (pp. 83–101). San Diego Elsevier.
Ottaviani, M., & Sørensen, P. N. (2015). Price reaction to information with heterogeneous beliefs and wealth effects: Underreaction, momentum, and reversal. *American Economic Review*, *105*(1), 1–34.
Peeters, T. (2018). Testing the wisdom of crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, *34*(1), 17–29.
Scheibehenne, B., & Broder, A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting*, *23*(3), 415–426.
Snowberg, E., & Wolfers, J. (2010). Explaining the favorite-long shot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, *118*(4), 723–746.
Spanias, D., & Knottenbelt, W. J. (2013). Predicting the outcomes of tennis matches using a low-level point model. *IMA Journal of Management Mathematics*, *24*(3), 311–320.
Sprenger, T. O., Tumasjan, A., & Welpe, I. M. (2014). Tweets and trades: the information content of stock microblogs. *European Financial Management*, *20*(5), 926–957.
Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. Brown Little.
Vaughan Williams, L., Sung, M., Fraser-Mackenzie, P. A. F., Peirson, J., & Johnson, J. E. V. (2018). Towards an understanding of the origins of the favourite–longshot bias: Evidence from online poker markets, a real-money natural laboratory. *Economica*, *85*(338), 360–382.
Williams, Vaughan. and , L. . (1999). Information efficiency in betting markets: A survey. *Bulletin of Economic Research*, *51*(1), 1–30.
Ziemba, W. T. (2020). *Parimutuel betting markets: Racetracks and lotteries revisited*: *SRC discussion paper 103*, Systemic Risk Centre, London School of Economics.