Bayesian Reasoning: Evidence from the Field

Constantinos Antoniou and Christos P. Mavis*

March, 2016

We conduct a field test of Bayesian reasoning, examining whether agents form expectations by placing a larger weight on cues that are more informative with lower process variance. To test this notion we analyze subjective probabilities inferred from odds on the outcomes of tennis matches, exploiting *exogenous* variation in process variance related to the format with which tennis matches are played. Our results are consistent with "*process variance neglect*", i.e., agents are not adjusting their ex-ante probabilities sufficiently according to process variance, and thus violate Bayes Rule. This result is robust to inferring subjective probabilities from odds offered by professional bookmakers or odds achieved on a person-to-person betting exchange. The resulting biases in expectations are costly.

JEL Classifications: D8, G1

Keywords: Bayes Rule, Behavioral Biases, Field Experiments

* *This version*: November 2016. Antoniou: Warwick Business School, University of Warwick, U.K. E-mail: constantinos.antoniou@wbs.ac.uk. Mavis: Surrey Business School, University of Surrey, U.K., E-mail: c.mavis@surrey.ac.uk. We have benefited from discussions with Nick Chater, Jerker Denrell, Arie Gozluklu, Andrea Isoni, Chris Starmer, Neil Stewart, Nathaniel Wilcox and particularly Graham Loomes, as well as seminar and conference participants at the Decision Research at Warwick Forum, the 54th Edwards Bayesian Research Conference, the 2016 Foundations of Utility and Risk Conference, Nottingham University, Warwick Business School and Surrey Business School. All remaining errors are our own.

1 Introduction

Do economic agents form expectations according to Bayes Rule? This question is fundamental for the analysis of economic decisions under uncertainty. Evidence from controlled laboratory experiments show that subjects do not engage in Bayesian reasoning (e.g., Kahneman, Slovic and Tversky, 1982; Grether, 1980) . However, because conditions in laboratory experiments differ from those in real-life in important ways, to further our understanding on the applicability of Bayes Rule, it is vital to also test Bayesian reasoning in the "field" (Harrison and List, 2004). We take a step in this direction, utilizing feature of professional tennis matches to test an aspect of Bayes Rule using subjective probabilities inferred from betting data.

Bayes Rule requires that agents correctly asses the "process variance" of each signal that affects the likelihood of interest, and attach a larger weight on those signals that are more informative with lower process variance. We examine whether economic agents adjust their expectations in this way, by analyzing subjective probabilities inferred from odds offered by bookmakers on the outcomes of tennis matches, exploiting *exogenous* variation in process variance that is related to the format with which tennis matches are played.

Men's singles tennis matches are played in two formats: A best- out-of-three format (BO3), where a player must win two out of possible three sets to win a match, and a best-out-of-five format (BO5), where a player must win three out of possible five sets to win a match.[1] In our sample, the BO5 matches are called "Grand Slams" (GS), and the BO3 matches are called "ATP World Tour Masters 1000" (MS). In each of these match types bookmakers observe signals that capture a player's skill *on the day*, calculate the subjective probability that this player will win the match, and set their odds accordingly.

[1] Each set is comprised by individual games. To win one game a player must win at least four points. To win a set a player must win at least 6 games. For more information on the rules of tennis see http://www.atpworldtour.com.

One such important signal that relates to a player's skill on the day is their official ranking, which is based on their cumulative performance during the previous 52 weeks.[2] Higher ranked players are generally more skilful than lower ranked players, and therefore more likely to win a match. If ranking is a perfect indicator of skill on the day, then process variance is zero and the probability that the high-ranked player wins is one. However, because skill on the day is affected by random components, ranking is an imperfect indicator of skill on the day, therefore the probability that the high-ranked player wins is less than one. How much less depends on the variance of the process, i.e., how likely it is that the low-ranked player produces a surprise and wins the match. If a surprise is less likely, then the probability that the high-ranked player wins the match should be adjusted upward.

A surprise is *less* likely in the GS format because the low-ranked player must win *more* sets to win the match.[3] To illustrate this idea with an analogue from coin spins, assume that we have a biased coin with probability of heads in a single toss equal to 60%. The probability that we receive two heads out of a possible three spins is 64.8%, and the probability that we receive three heads out of a possible five spins is 68.3%. The increase in the probability in each case reflects a reduction in process variance in the sense that it becomes more difficult for the low probability event (i.e., tails in this example or the low-ranked player winning in our framework) to "win" when more successes are required. Indeed, in line with this logic, we find that high-ranked players are roughly 7% more likely to win a GS match, as opposed to an MS match. Therefore, Bayes Rule dictates that

[2]Throughout the paper we refer to the player with the highest ranking (i.e., a smaller ranking number), as the high-ranked player. Players' rankings are based on their immediate past 52 weeks performance from a total of 18 tournaments (or 19, if a player participates in the Barclays ATP World Tour Finals). A player's total ranking points are calculated from the four Grand Slams, eight compulsory Masters 1000 tournaments (out of nine), and his best six results (in terms of points) from all other ATP and Futures tournaments that he participates. For more information see http://www.atpworldtour.com/en/corporate/rulebook.

[3]Klaassen and Magnus (2014) also make this point. They develop a model to estimate the probability of a player winning a specific point, and show that the higher skill player is more likely to win a longer, as opposed to a shorter, match.

assigned probabilities to the high-ranked player winning must also be adjusted upward for GS matches to reflect this reduction in process variance. Our objective is to test whether such an adjustment takes place.

For our main analysis we use subjective probabilities inferred from fixed decimal odds offered by several major betting houses on professional men's tennis matches for the period 2005 to 2014. For each match we infer the subjective probability that the high-ranked player wins the match from bookmaker odds, π, estimate the "objective" probability as the fitted value from a logit model, $\hat{p}$, and define *bias* as $\pi - \hat{p}$. We test our hypothesis based on the difference in the average *bias* between MS and GS.[4] If bookmakers upwardly adjust π for GS to reflect the increase in $\hat{p}$ due to the longer match format (BO5), then this difference should be 0. Our results indicate that bookmakers adjust π upwards for GS, but by an *insufficient margin*, with *bias* being lower in GS by 3.3%. This finding is consistent with "*process variance neglect*", i.e., the notion that bookmakers subjective probabilities do not properly reflect changes in process variance in GS, and thus violate Bayes Rule.

Could these results reflect strategic behavior by rational bookmakers to exploit punters? For example, Levitt (2004) finds that, in spread-betting markets, bookmakers sometimes offer actuarially biased prices to exploit the preference of punters to bet on the favorite. We do various calculations to illustrate that the bias we document is suboptimal, but the most direct test we do to address strategic behavior is to re-do our analysis using a different data set where odds are determined in a person-to-person betting exchange market, called Betfair. In this setting strategic incentives do not exist because the odds on the two competing players are set by via the interactions of *different* agents on a competitive basis. An additional interesting feature of this dataset, is that it allows us

[4]By testing the hypothesis based on the difference in *bias* across MS and GS our conclusions are not affected by any systematic errors in the estimation of π or $\hat{p}$ that are constant across the two match formats. We discuss our methods in detail in Section 2.

to examine the "wisdom of crowds", since Betfair prices aggregate the information sets of many different individuals.[5]

Analyzing tennis matches for the period 2009-2014, we find results that are consistent with our baseline analysis using bookmaker data. Specifically, *bias* is lower for GS by 2.6%, and this difference is highly statistically significant and comparable in magnitude to that found with bookmaker data. This finding suggests that strategic behavior is an unlikely explanation for our findings with bookmaker odds. In additional analysis we examine whether high-volume matches, which aggregate information from a larger number of punters, price the change in process variance in GS more accurately. We find that they do, with the difference in *bias* between MS and GS for those matches being closer to 0. However, high-volume markets still entail a significant bias due to process variance neglect.

Errors in expectations should be costly to decision makers. To put this notion in context, our findings show that *bias* is significantly more *negative* in GS relative to MS, indicating that for GS matches bookmakers are offering excessively high odds for the high-ranked player. Therefore, bookmakers should be earning less in GS matches. Using different indicators of *actual* profitability we indeed confirm this assertion. We also analyze bookmakers *expected* profits, and again find that a lower *bias* in GS is suboptimal for a rational bookmaker. Overall, the analysis of profits suggests that the biases we document due to process variance neglect are costly.

Despite our effort to create a sample where the BO3 matches are as similar as possible to the BO5 matches, some differences remain between the two match formats; for example GS matches are more prestigious, offering higher prize money to players and attracting more attention from punters. Could these differences drive (or contribute to) the results we document? For our final test we conduct a placebo test to address such

[5]For an early examination of the wisdom of crowds using laboratory generated data see Camerer (1987).

concerns, using data for professional women's tennis matches. For women, the differences across GS and MS matches are generally the same as for the men, but for women GS matches *are played in a BO3 format, exactly like the MS tournaments*. Hence, women's matches preserve the differences across the two match formats *except* for the change process variance. If the biases we document are driven by factors other than process variance neglect, we should still observe them in the women's data.

Logistic analysis shows that the type of match (MS vs GS) *does not* affect the probability that the high-ranked player wins for women's matches, consistent with no difference in process variance across match formats. Therefore, bookmakers should not adjust π upward for GS. When we compare *bias* calculated from bookmaker odds for MS and GS we find that the difference is *insignificant*. When we do the same for *bias* calculated from betting exchange odds, we find that *bias* is *higher* in GS by 1.6%. This is a form of process variance neglect that is opposite to that found in the men data, i.e., punters are increasing π for GS matches, even though these matches do not entail reductions in process variance.

Overall, results from women's matches are in stark contrast from those obtained from men's matches, which provides support to the claim that the patterns we observe in our baseline analysis reflect process variance neglect.

We conduct various robustness checks. Firstly, we show that our results are not sensitive to different sample specifications or econometric models. Secondly, we consider various alternative explanations for our findings, such as noise trading or bookmaker preference differentials across match formats and competition between rational bookmakers, and conclude that such alternative explanations cannot offer a parsimonious interpretation for all our results across bookmaker *and* betting exchange markets, for both men's and women's matches.

Tests of Bayes Rule have been mainly conducted in controlled laboratory experiments (Grether, 1980, 1992; Camerer, 1987; Griffin and Tversky, 1992; Holt and Smith, 2009; Antoniou, Harrison, Lu and Read, 2015). The findings from these experiments generally show that subjects in experimental conditions are unable to engage in Bayesian reasoning. However, because the extent to which laboratory findings carry over into the field is unclear, it is important to analyze both laboratory *and* field generated data to properly understand the process of belief formation.[6]

Along these lines, DeBondt and Thaler (1990) present evidence that the earnings forecasts of professional security analysts are affected by representativeness. Chen, Moskowitz and Shue (2016) show that the decisions of asylum judges, loan officers and baseball umpires are affected by the gambler's fallacy. Various empirical studies in behavioral finance can also be seen as indirect field tests of Bayes Rule, although their conclusions hinge on relatively strong assumptions.[7] We contribute to this literature conducting a field test of Bayesian reasoning with three attractive features: Firstly, subjective probabilities are inferred from the decisions of expert agents who are pricing securities in their natural habitat with significant monetary consequences. Secondly, uncertainty is fully resolved when the match is finished, which allows us to test for a bias in subjective probabilities with relatively weak assumptions. Thirdly, and most importantly, in the tennis data the variation in process variance that allows us to test for Bayesian reasoning

[6] As discussed in Harrison and List (2004) the field differs from the laboratory in at least two fundamental ways: Firstly, the field can help create heuristics that affect decision making, which are probably not present in the "sterile" environment of the laboratory where subjects encounter artificial tasks for the first time (e.g., List, 2004 and Haigh and List, 2005). And secondly, incentives in the laboratory tend to be weak, therefore behavior inferred from this domain may not be directly applicable to real-life economic decisions where the stakes are significantly higher (e.g., Andersen, Ertac, Gneezy, Hoffman and List, 2011). Other studies that point to differences between lab and field behavior include Levitt and List (2007a,b) and Levitt, List and Reiley (2010).

[7] For example, DeBondt and Thaler (1985) present evidence consistent with investors forming expectations according to representativeness. Other studies that draw similar conclusions include Lakonishok, Shleifer and Vishny, (1994), Daniel and Titman (2006), Li and Yu (2012), and Antoniou, Doukas, and Subrahmanyam (2016). However, such tests rely on assumptions regarding the process that links risk with expected stock returns, which make identification of errors in expectations difficult (Fama, 1998). An interesting exception in this literature is Moskowitz (2015), who tests for the existence of asset pricing anomalies in sports betting markets where risks are completely idiosyncratic, and therefore Fama's (1998) critique does not apply.

is completely exogenous, in the sense that it is governed solely by the rules of the game. Our findings suggest that decision makers do not combine cues optimally when forming expectations and support the claim that violations of Bayes Rule affect real-life economic decisions.

2 Data and Methods

2.1 Data

For our baseline analysis we obtain data from www.tennis-data.co.uk.[8] For every match this database contains data the name of the tournament, the date of the match, the names of the two competing players, their official ATP rankings prior to the tournament, the winner of the match, as well as fixed decimal odds from various international betting houses on both players.[9] In our analysis, we average the odds offered by the various bookmakers on the two players to infer subjective probabilities.

We include in our sample Grand Slam (GS) matches, which are played in a BO5 format, and ATP World Tour Masters 1000 matches (MS), which are played in a BO3 format. GS tournaments are the most prestigious, with the winner receiving 2,000 ranking points, and on average collecting 2.25 million dollars (in 2015). For comparison, the winner of an MS tournament earns 1,000 ranking points and, on average, 0.7 million dollars (in 2015). There are other tournaments that are played in a BO3 format, which yield, for example, 500 or 250 ranking points to the winner of the tournament, and offer less prize money. Such tournaments are significantly less prestigious, involving on average lower-ranked players, and attracting less attention from punters. In our analysis, to ensure that the BO3 matches are as similar as possible to the BO5 GS matches, we

[8]Data from this database has been used previously by academic work on tennis matches (e.g., Forrest and McHale, 2007; Del Corral and Prieto-Rodriguez, 2010).

[9]This dataset contains odds from eight different bookmakers; Bet365, Centrebet, Expekt, Ladbrokes, Interwetten, Pinnacles Sports, Stan & James, and Unibet.

focus on the more prestigious tournaments from the BO3 class.

We apply the following criteria to the initial dataset (n=10,790) to create our final sample: (i) we retain only completed matches (n=10,334), (ii) drop matches with missing rankings information (n=10,327), (iii) matches with no odds for either player (n=10,295), and (iv) matches that entail a negative Vig,[10] (n=10,266). Moreover, (v) we drop matches were both players are ranked outside the top 100 players in the world, as indicated by their ranking prior at the start of the tournament (n=10,013). Because skill in tennis has a pyramid structure, as noted by Klaassen and Magnus (2014), rankings are less informative among lower-ranked players. Lastly, (vi) we drop matches where the high-ranked player is indicated as an outsider by bookmakers even though he is ranked by at least 15 places higher than his opponent at the beggining of the tournament (n=9,046). Such cases are likely to reflect recent developments like injuries, which are not yet incorporated in the rankings thus making them less informative.[11] Our final sample consists of 9,046 tennis matches, covering the period 2005-2014. Table 1 shows a breakdown of the observations by year and tournament. We have data for 4 GS tournaments, and for 12 MS tournaments. Some MS tournaments are discontinued and others are introduced at various point in time.

[Insert Table 1 here]

2.2 Methods

Assume a tennis match between players X and Y. The bookmaker offers fixed decimals odds for player X to win equal to d_X, and for player Y equal to d_Y, where d_X and d_Y

[10] Vig stands for "vigorish" or housetake, reflecting a type of commission collected by the bookie. Matches with negative vig are likely data errors.

[11] Because our analysis aims to examine how skill-related signals (i.e., rankings) are priced across MS and GS, filters (v) and (vi) are aimed to remove cases where these signals are unreliable indicators of skill. In Section 3.4 we test the robustness of our results in a different sample that does not incorporate these filters and find that our conclusions remain the same.

are greater than 1. To obtain subjective probabilities, we first invert the quoted odds for X, $O_X = \frac{1}{d_X}$, and for Y, $O_Y = \frac{1}{d_Y}$. In a perfectly competitive and frictionless market with a risk-neutral bookmaker O_X and O_Y correspond to true subjective beliefs. However, typically $O_X+O_Y > 1$, which reflects the vigorish or "vig", a form of commission collected by the bookie. To obtain subjective probabilities we normalize O_X and O_Y to sum to 1, using $\pi_X = \frac{O_X}{O_X+O_Y}$ and $\pi_Y = \frac{O_Y}{O_X+O_Y}$. The vig is thus split proportionally between the two players, depending on their relative odds.[12] Throught the analysis we refer to π_i as the subjective probability that the high-ranked player wins match i.

To examine whether process variance changes across match format (MS vs. GS) we use the logistic model, shown below:

$$Pr(Y_i = 1|GS_i, Rskill_i) = F(\alpha + \beta_1 GS_i + \beta_2 Rskill_i) \tag{1}$$

The dependent variable, Y_i is a binary indicator taking the value of 1 if the high-ranked player wins match i, and 0 otherwise. GS_i is a dummy variable that equals 1 if match i is Grand Slam, and 0 otherwise. $RSkill_i$ captures differences in player rankings for match i and is calculated as log(low-ranked player ranking) - log (high-ranked player), following the specification in Klaasen and Magnus (2001).[13] F is the logistic distribution.

As we discuss in more detail in the next section, GS matches entail more players in the draw, and therefore entail higher average ranking differences between the two players. The model in Equation (1) captures the effect of the change in process variance across MS and GS, whilst controlling for these ranking differences. That is, the coefficient on GS should be positive and significant, reflecting the increase in the probability that the

[12] This assumption is typical in the literature (e.g., Croxson and Reade, 2014; Smith, Paton and Williams, 2009).

[13] Note that the ranking of the low-ranked player is a larger number than the ranking of the high-ranked player. As discussed in Klaasen and Magnus (2001) player skill in tennis resembles a pyramid (i.e., the difference in skill between players ranked #1 and #10 is higher than the skill difference between players ranked #80 and #90), therefore a logarithmic transformation is appropriate.

high-ranked player wins a GS match due to the longer match format (BO5). To estimate the objective probability that the high-ranked player wins we use the fitted value from this model, $\hat{p_i}$.

Having estimates for both objective and subjective probabilities, we estimate the bias in expectations for match i as $bias_i = \pi_i - \hat{p_i}$. A strict version of the null hypothesis under Bayes Rule is that $bias_i = 0$, for both MS and GS. This formulation, however, requires the strong assumption that subjective and objective probabilities have been recovered without any systematic error.[14] To avoid making this assumption, we test our hypothesis based on the *difference* in average *bias* between the two match formats $\Delta_{bias} = \overline{bias_{GS}} - \overline{bias_{MS}}$. With this formulation our conclusions are not affected by any systematic errors in π_i or $\hat{p_i}$ that are symmetric across the two match formats.[15]

Under the null hypothesis of Bayesian reasoning, where subjective probabilities are properly adjusted to changes in process variance between MS and GS, $\Delta_{bias} = 0$. The alternative hypothesis, under "*process variance neglect"*, where bookmakers do not properly adjust their subjective probabilities, is $\Delta_{bias} < 0$. We test these hypotheses by examining the sign and significance of β_1 in the ordinary least square regression shown below:[16]

$$bias_i = \alpha + \beta_1 GS_i + \beta_2 RSkill_i + \epsilon_i \tag{2}$$

What cognitive mechanism can lead to process variance neglect? In a seminal pa-

[14] Various factors could lead to a systematic error in $bias_i$, for example, the effect of bookmaker's preferences toward risk on subjective probabilities (i.e., Savage, 1971), and/or the effect of variables besides GS_i and $RSkill_i$ on objective probabilities.

[15] In Sections 3.4 and 3.7 we discuss sources of potential assymetric recovery errors, and when possible address them econometrically.

[16] An alternative way to test our hypotesis is to estimate Equation (1) as $\hat{p_i} = Pr(Y_i = 1|Rskill_i, RSkill_i \times GS_i) = F(\alpha + \beta_1 Rskill_i + \beta_2 RSkill_i \times GS_i)$ and then test the hypothesis using $bias_i = \alpha + \beta_1 RSkill_i + \beta_2 RSkill_i \times GS_i + \epsilon_i$, expecting that β_2 is negative and significant. In untabulated analysis, which is availiable from the authors upon request, we have tried this specification and our conclusions in relation to process variance neglect remain very similar.

per in Bayesian Updating, Edwards (1968) shows that people underweight new evidence, forming posteriors that are significantly below the correct Bayesian probability. In contrast to Edwards (1968), Kahneman and Tversky (1972) show that subjects overweight the importance of new evidence, resulting in posteriors that overshoot the Bayesian probability. Griffin and Tversky (1992) explain these contradictory findings, suggesting that people form expectations by under weighting the statistical validity of available evidence, or "weight" (i.e., process variance in our context), and overweighting its "strength", or how saliently it supports a specific hypothesis.[17] Griffin and Tversky (1992) argue that the Kahneman and Tversky (1972) experiments entail signals of low weight (a small sample) whereas the Edwards (1968) experiments entail signals of high weight (a large sample). In our case the signal (i.e., player ranking) is more predictive of outcomes in GS matches (BO5), therefore, in the language of Griffin and Tversky (1992), it is of higher weight and will thus elicit more underreaction.[18]

[Insert Table 4 here]

[17]To illustrate consider the example discussed in Griffin and Tversky (1992). A graduate student has two reference letters, one from a young assistant professor (AP) and one from a seasoned full professor (P). The AP describes the student's potential as "outstanding", whereas P as "satisfactory". The letter from AP is high in strength (salient tone) but low in weight (AP has limited experience in assesing student quality). The opposite is true for P. Griffin and Tversky (1992) suggest that those who read the letters are likely to overreact to AP's letter and underreact to P's letter. Such a strength-weight bias may arise because strength is easily observed by the decision maker (i.e., the tone of the letter is directly visible), whereas weight (the experience of the referee) requires more effort to be processed.

[18]Such a bias can be illustrated as follows: Assume an agent who is updating his expectation after observing a signal, S, with variance equal to σ_S^2. This agent has priors with variance σ_P^2. Assuming normal distributions, the weight put on the signal according to Bayes Rule is equal to $\frac{\sigma_P^2}{\sigma_P^2+\sigma_S^2}$. Because the agent exhibits process variance neglect he will instead use a weight on S equal to $\frac{\sigma_P^2}{\sigma_P^2+\lambda\times\sigma_S^2}$. In this expression, $\lambda = f(\sigma_S^2) > 1$ captures the bias due to process variance neglect, with $\frac{\partial\lambda}{\partial\sigma_S^2} < 0$. All else equal, this agent underreacts to S, thus his posterior is below the Bayesian probability. If the variance of the signal is lower (as in GS) λ is higher, which leads to stronger underreaction.

2.3 Descriptive Statistics

Table 2 presents descriptive statistics for our main sample, separately for GS (Panel A) and MS (Panel B) matches. The average posted odds offered by bookmakers that the high-ranked player wins the match (HR_{Odds}) are much higher than those for the low-ranked player (LR_{Odds}) to win the match (1.34 vs. 6.17 for GS and 1.48 vs. 4.11 for MS), which shows that player ranking is indeed a metric for the relative skill of the two players on the day that is used by bookmakers. The *vig* is roughly equal in the two match formats (0.05 in GS and 0.06 in MS) and has low volatility, as noted by other authors (Forrest and McHale, 2007). For GS the average ranking between the high- and low-ranked players (HR_{Rank} and LR_{Rank}) is 24.9 and 95.42 respectively, for an average difference in rankings ($diff_{Rank}$) equal to 70.5 positions. For MS the corresponding rankings are 20.56 and 65.56, for an average difference of 45 positions. This difference occurs because GS tournaments have more players in the draw and is important because it contributes to the probability that the high-ranked wins a GS match without being related to changes in process variance. Therefore, our models test the hypothesis whilst controlling for $RSkill$.

The next two rows show the average estimate of the objective probability obtained as the fitted value from the model in Equation (1), ($\hat{p}$), as well as the average of a dummy variable that equals 1 if the high-ranked player has won, and 0 otherwise (D_{HR}). As it can be seen, $\hat{p}$ is higher for GS than MS (0.79 vs 0.69), consistent with view that ranking is a signal with lower process variance in GS. The penultimate row in the table shows the average subjective probability that the high-ranked player wins the match (π), which is higher for GS compared to MS (0.75 vs. 0.68). This means that bookmakers are adjusting their subjective probabilities for GS matches relative to MS matches in the direction predicted by Bayes Rule. However, as shown in Figure 1, this adjustment seems insufficient, i.e., it does not completely reflect the increase $\hat{p}$ for GS. The last row in Table

2 shows the average prize money (in dollars) received by the winner of an MS and a GS tournament, using 2015 prize money. GS tournaments offer almost three times higher prizes than MS tournaments.

[Insert Table 2 and Figure 1 here]

3 Analysis

3.1 Changes in Process Variance

We start our analysis by showing more formally that the high-ranked player is more likely to win a GS match due to the longer format (BO3 vs BO5) using the model shown in Equation (1). The marginal effect associated with the dummy variable *GS* quantifies the effect of this change in process variance from MS to GS.

The results are shown in Table 3. We run specifications with and without *RSkill*, and a specification that additionally includes fixed effects related to the surface that the match is played on (clay vs. hard),[19] the round of the match (1 to 7), and the year that it takes place.[20]

From Column (1) we observe that the marginal effect associated with GS is 10.1% and highly statistically significant, indicating that high-ranked players are more likely to win a GS match. Once we control for *RSkill* in Column (2) the marginal effect associated with GS reduces to 7.4%, but remains statistically significant. *RSkill* is positive and also highly significant. From Column (3), we observe that the additional fixed effects do not

[19]Matches are played in either hard, clay, grass or carpet surfaces. The surface relates to the speed that the ball loses after it bounces, and some players prefer faster surfaces (i.e., grass), whereas others prefer slower surfaces (i.e., clay). In our sample, the only tournament played on grass is the Wimbledon (GS), and the only tournament played on carpet is Paris (MS) between 2005 and 2008. Because in terms of speed, grass and carpet surfaces are more similar to hard than clay surfaces, we include them in the hard category.

[20]The year fixed effect is aimed to capture the influence of year to year variations in variables that are ommited from our model, such as betting activity or the number of bookmakers.

materially change the findings.

Overall, this analysis shows that the reduction in process variance in GS due to to longer match format exerts a significant and positive effect on the probability that the high-ranked player wins a GS match, thus should be considered by Bayesian bookmakers.

[Insert Table 3 here]

3.2 Biases in Subjective Probabilities

Table 4 shows our main results. In a univariate setting in Panel A, we find that $\overline{bias_{GS}}$ is equal to -3.8% and $\overline{bias_{MS}}$ is equal to -0.4%, making Δ_{bias} equal to -3.5%, and highly statistically significant.

In Panel B, we more formally test the hypothesis in a multivariate setting. We find that the coefficient on *GS* is -3.3% and highly statistically significant. This is marginally lower than in the univariate setting, reflecting that a small portion of the bias documented in Panel A is driven by *RSkill,* which is also negative and statistically significant. Adding surface, round and year fixed effects in Column (2) does not materially change our findings.

Overall, the results in Table 4 are consistent with the alternative hypothesis of *process variance neglect.* Even though bookmakers upwardly adjust their subjective probabilities that the high-ranked player wins a GS match relative to an MS match (68.2% vs. 74.8%), their adjustment is not sufficient.

[Insert Table 4 here]

3.3 Betting Exchange Data

In this section, we test our hypothesis using subjective probabilities inferred from odds achieved on a person-to-person betting exchange called Betfair.[21] Analysis of this dataset allows us to test the hypothesis in an environment where odds are unaffected by any strategic incentives of bookmakers to exploit punters, and also examine whether markets are less biased than individual experts.[22] This data set also contains information on the volume bet ($) on each player, which allows us to provide an estimate for bookmaker's actual profits. Our sample contains 4,802 observations for the period 2009-2014.

In Panels A and B of Table 5 we present descriptive statistics for this alternative sample. The majority of the variables are very similar as those in Table 2. One noteworthy difference is that the vig in the Betfair data is much lower than in the bookmaker data, which occurs because Betfair odds on the two players are competitively determined by many different agents, as opposed to being set by a single bookmaker who seeks to earn the vig regardless of the outcome.

Table 5 contains descriptive statistics on the total volume bet on each match $TotVol$ and also the proportion of $TotVol$ that backs the high-ranked player, $RVol$ for the two match formats (MS vs. GS). As shown by $TotVol$, GS matches attract higher betting activity (roughly by 40% on average), which reflects the fact that they are higher profile events. Moreover, as shown by $RVol$, punters show a very strong preference to bet on the high-ranked player (i.e., the favorite) for both MS and GS matches, consistent with the results in Levitt (2004).[23]

[21]Betfair is by far the largest person-to-person betting exchange, with almost one million active users (Croxson and Reade, 2014). The Betfair data set was purchased from Fracsoft, available at http://www.fracsoft.com/. The dataset has incomplete coverage of the Australian Open (GS tournament), with no observations for 2009 and 2010. Moreover, it does not include data for matches that were completed in more than one day (for example due to rain delays). We use "back" odds on the two players to calculate subjective probabilities (i.e., odds available to punters who want to bet that a specific player wins).

[22]Prediction markets, such as the betting exchange, have been shown to forecast uncertain future events better than individual experts (Wolfers and Zitzewitz, 2004).

[23]This tendency of punters to prefer the favorite is in contrast to the well known favorite-longshot

[Insert Table 5 here]

The results are shown in Table 6. In Panel A the univariate analysis shows that in MS tournaments π is equal to 70.2% and p is equal to 69.7%. The corresponding figures for GS are 77% and 78.9% respectively, for a significant Δ_{bias} of -2.4%.

In Panel B, we test the hypothesis using multivariate analysis. Similar to the analysis in Table 4 we observe that controlling for *Rskill* does not change our findings, as the coefficient on the *GS* dummy is equal to -2.4% and statistically significant. Adding surface, year and round fixed effects does not influence the result either.

When the trading volume in prediction markets like Betfair is higher, resulting prices may be more efficient, as they are likely to incorporate information from a larger number of traders. Thus, the odds achieved in high-volume matches may reflect changes in process variance in GS more accurately. We use data on the volume that backs the high-ranked player (Vol_{HR}) to test this hypothesis. Specifically, for each match type (MS vs. GS) and round of match (1 to 7), we rank according to Vol_{HR}, noting the median of the distribution. If, for a specific match, Vol_{HR} is higher than the corresponding median then $VolD$ equals 1, otherwise it is 0.[24] We re-estimate the model in Equation (2), by including $VolD$ and the interaction between $VolD \times GS$ as additional regressors. The coefficient on $VolD \times GS$ tests whether the bias related to process variance neglect is different for high-volume matches.

The results are shown in Table 6, Panel B, Column (3). We find that the coefficient on GS for low-volume matches is equal to -2.9% and highly significant. The coefficient on the interaction $VolD \times GS$ is *positive* and significant at 1.3%, indicating that the bias related to process variance neglect for high-volume matches is reduced by roughly 45%

bias, which implies that the long-shot event is overbet (e.g., Griffith, 1949; Snowberg and Wolfers, 2010). In the tennis data it is the short event that is overbet, i.e., the proportion of volume that backs the high-ranked player is higher than the corresponding probability.

[24] We note the median after we sort on the basis of match type and round of match because GS matches attract higher volume, as do later rounds in a tournament. In this way, the dummy $VolD$ is not capturing mechanically such cases.

(-2.9% vs -1.6%). However, even though high-volume markets are more efficient, they still entail a significant bias due to process variance neglect.[25]

Overall, this analysis shows that a bias reflecting process variance neglect exists in the betting exchange data, where strategic incentives do not exist. Moreover, the bias is smaller in matches that attract higher betting volume.

[Insert Table 6 here]

3.4 Robustness Checks

In this section, we conduct various tests of robustness, presenting the results in Table 7. In Panel A, we define $bias_i$ using subjective probabilities inferred from bookmaker odds and in Panel B using betting exchange odds.

In Column (1), we estimate the model in Equation (2) by defining $bias$ as $\pi - D_{HR}$, where D_{HR} is a dummy that equals 1 if the high-ranked player won the match, and 0 otherwise. This specification is unaffected by any measurement error in $\hat{p}$, and transfers all the uncertainty associated with its estimation to the residual. Column (1) of Panels A and B shows that, even though standard errors are higher, the coefficient of GS remains highly statistically significant.

In Column (2), we estimate a quantile regression model to control for the confounding effect of any outliers. In Panel A the coefficient on GS equals -2.3% and in Panel B -1.1%, both highly statistically significant.

In Column (3), we re-estimate the model in Equation (2) using a sample that does not include filters (v) and (vi), discussed in Section 2.1.[26] In Panel A, the coefficient on GS is equal to -2.5% and in Panel B -2.0%, both statistically significant. The magnitude

[25]Note that the $bias$ of -1.6% in high-volume markets is significant at the 1% level (unreported result).

[26]For this model we estimate $\hat{p}$ using the logit in (1) applied to the sample that does not include filters (v) and (vi).

of the effect is somewhat reduced in this case, but this is expected since we are adding in our sample matches where rankings are more noisy indicators of skill differentials.

Because GS matches offer more prize money it is possible that high-ranked players "time" their form to peak at GS matches, and therefore our findings reflect such timing effects and not changes in process variance. In Column (4) we address this issue, by adding an additional regressor in the model in Column (3), $RStreak$, calculated as the difference in the proportion of games won by the high- and low-ranked player in the two previous tournaments. The results in Column (4) of Table 7, show that the coefficient on GS remains negative and significant in this specification, equalling -2.8% in Panel A and -2.9% in Panel B.[27]

Overall, the results in this section show that our findings our robust to different model and sample specifications.

[Insert Table 7 here]

3.5 Profits

In this section, we analyze profits of bookmakers to examine the economic implications of the biases documented in Section 3.2. Our first measure of profitability is the return earned by the bookmaker for each unit of currency bet on on the high-ranked player, r_{HR}. Recall that D_{HR} is an indicator variable, equal to 1 if the high-ranked player has won the match, and 0 otherwise. The return earned by the bookmaker in match i for the high-ranked player, r_{HR}, is thus:

[27]For the model in Column (4), we estimate $\hat{p}$ using a logit model that includes $RStreak$ as an additional independent variable. The results from this model show that the marginal effect associated with $RStreak$ is positive and significant. In addition, $RStreak$ is higher for GS by roughly 2%. The results in Column (4) show that agents are not properly pricing the effects of $RStreak$ on the probability that the high-ranked player wins, as its coefficient is negative and significant. Note that this model entails a smaller number of observations due to missing values in $RStreak$, which occur when a player did not compete in the two previous tournaments.

$$r_{HR} = \begin{cases} 1 & \text{if} \quad D_{HR} = 0 \\ -(O_{HR} - 1) & \text{if} \quad D_{HR} = 1 \end{cases} \tag{3}$$

Our second measure of profitability provides an estimate for the bookmaker's actual profits, as a proportion of the total volume staked for match i, $R\Pi_i$, using the volume information from the betting exchange as below:

$$R\Pi_i = 1 - RVol_W \times O_W \tag{4}$$

where $RVol_w$is the proportion of the total volume staked that backs the winning player, and O_W are the fixed decimal odds offered by the bookmaker on the winning player. An implicit assumption in the analysis of $R\Pi_i$ is that $RVol_w$ is similar in the two different betting platforms (bookmakers vs. betting exchange). We estimate r_{HR} and $R\Pi_i$ using both bookmaker and betting exchange odds.[28]

To examine whether profitability varies across MS and GS we use the model in Equation (2) with the profitability indicators as the independent variable. Since bookmakers offer excessively high odds for the high-ranked player to win a GS match (as shown by the more negative *bias* in GS in Table 4) we expect that both r_{HR} and $R\Pi_i$ are lower for GS.

As seen from Panel A1 in Table 8 (bookmaker odds), r_{HR} is lower in GS by a significant -4.1%, showing that bookmakers are earning proportionately less on bets on the high-ranked player for GS. From Panel B1 (betting exchange odds) we observe that the coefficient on GS is negative (-2.5%), but statistically insignificant.

The analysis of r_{HR} is only indicative since it does not incorporate returns earned by bookmakers on bets on the low-ranked player. Since bookmakers are offering excessively high odds on the high-ranked player for GS, by analogy, they must be offering excessively low odds on the low-ranked player increasing their returns on those bets. So it is

[28]However, $R\Pi_i$ does not capture the profits of punters who offer odds on the betting exchange. The calculation of $R\Pi_i$ with betting exchange odds corresponds to the profits earned by a hypothetical bookmaker who offers odds on both players, equal to those achieved on the exchange.

possible that their overall profitability increases (although this seems unlikely since the overwhelming majority of bets backs the high-ranked player, as shown by Table 5).

In Column (2), in Panels A and B, we present results for $R\Pi_i$, which gives a more complete picture of profits. The results show that the coefficient on GS is negative and significant in both cases (bookmaker and betting exchange), equal to -2.5% in Panel A2 and -2.1% in Panel B2. This shows that bookmakers are earning a smaller proportion of the volume staked in GS relative to MS, consistent with the notion that biases due to process variance neglect are costly.

The Appendix further illustrates that the biases we document are suboptimal by deriving an expression for the *expected* profits earned by a Bayesian and risk-neutral bookmaker in GS, with and without a bias due to process variance neglect. This analysis shows that a rational bookmaker would be better off in the absence of bias, which further suggests that the biases we document are costly.

[Insert Table 8 here]

3.6 A Placebo Test

Our tests implicitly assume that the only difference across MS and GS matches is match format, i.e., BO3 vs. BO5. However, as mentioned, GS matches are more prestigious than MS tournaments, offering more prize money and attracting higher betting volumes. Such differences may affect our findings. For example, one could argue that in GS matches players are more motivated to win due to the higher prize money, thus high-ranked players (who possess more skill) win more often. And perhaps bookmakers are not properly adjusting their subjective probabilities to such variations in incentives across MS and GS.

To examine whether our conclusions are affected by any unobserved heterogeneity across MS and GS that is unrelated to process variance we conduct a placebo test using data for women's tennis matches. This data provide an ideal setting for such a test as they preserve the key differences across MS and GS that are observed in the men's sample (i.e., GS are more prestigious, offer more prize money, and attract more betting volume), but for women there is *no change in process variance because both MS and GS are played in a BO3 format.* Therefore, if our results are driven by other factors and not changes in process variance, Δ_{bias} should also be negative and significant in the women data.

We construct the women's sample using an approach similar to that used in our baseline analysis with the men's data. GS tournaments are the same for women as for the men, and for MS tournaments we again focus on the more prestigious tournaments.[29] After applying the same filters to the initial sample as those for the men we end up with 2,491 MS matches and 3,532 GS matches from 2007-2014 for the bookmaker sample, and 1,408 MS matches and 2,210 GS matches from 2009-2014 for the betting exchange sample.

Table 9 shows descriptive statistics for the women bookmaker sample. The vig for women is similar to that for men. The difference in player rankings (*diff*) is very similar in the women data set, equal to 46.9 for MS and 66.06 for GS. The latter tournaments offer the same prize money to women as they do to men. As in the ATP, GS matches for women always offered more prize money than MS matches. Similar results are shown in Table 10 for the betting exchange women's sample. Overall, the data from Tables 9 and 10 show that the women samples are broadly similar as the men's.

[Insert Tables 9 and 10 here]

In Table 11, we present analysis from the logistic regressions shown in Equation (1) using the women data (Panel A bookmaker, Panel B betting exchange). As shown from Column (1), when the only regressor is the *GS* dummy, the marginal effect associated *GS* is positive and significant. However, from Column (2), we observe that this significance is only capturing *Rskill* differentials; once we control for *Rskill* the marginal effect associated

[29] Pre-2009, the top tier masters tournaments were called Tier-I and post-2009 they are called Premier Mandatory. For the bookmaker sample we have 12 MS tournaments prior to 2009 (Berlin, Charleston, Doha, Indian Wells, Miami, Montreal, Moscow, Rome, San Diego, Tokyo, Toronto, and Zurich) and the 4 Premier Mandatory tournaments after 2009 (Beijing, Indian Wells, Madrid, Miami). The Betting exchange sample starts in 2009 so we have data on the 4 Premier Mandatory tournaments. The GS matches for women are the same as for the men (Australian Open, French Open, Wimbledon, and US Open).

with *GS* is *insignificant,* so the probability of the high-ranked player winning an MS or a GS match is the same. Therefore, since process variance does not change across MS and GS matches for women, agents *should not* upwardly adjust their subjective probabilities of the high-ranked player winning a GS match.

We continue by testing for biases in subjective probabilities, using the model in Equation (2). The results are shown in Table 12 (Panel A bookmaker, Panel B betting exchange). From Panel A, Column (1), we observe that, in contrast to the men's sample, the coefficient on *GS* is positive but *insignificant*, equal to 0.3%. After we control for fixed effects, it becomes significant at the 10% level. We obtain similar results in Panel B, where the coefficient on *GS* is positive and significant, equal to 1.7%.[30]

The results in this section are in stark contrast to those obtained in the men's data. For women's matches, especially for the betting exchange sample, agents are adjusting their subjective probabilities for the high-ranked player winning a GS match upward, as they do for men's matches, even though in this case this adjustment is unwarranted, as shown in Table 11. This creates a bias due to process variance neglect, but in the *opposite* direction from that found in men's data, i.e., $\Delta_{bias} > 0$.

Overall, the placebo test with the women's data provides support to the claim that our baseline results in Tables 4 and 6 reflect biases due to process variance neglect and not other differences between MS and GS.

[Insert Table 11 here]

3.7 Alternative Explanations

In this section, we discuss whether our findings may be related to alternative explanations other than process variance neglect.

Our baseline findings show that because bookmakers are offering overly attractive odds to the high-ranked player winning a GS match, they earn proportionately less.

[30]In untabulated analysis, available from the authors upon request, we examine whether bookmaker's profits differ across MS and GS for women's matches. We find that, in the case where bookmakers actual profits ($R\Pi$) are calculated using betting exchange odds, $R\Pi$ is higher for GS (significant on the 10% level), consistent with Δ_{bias} being more positive in those cases (Table 12, Panel B).

However, their average profitability for GS matches remains positive (unreported result). Could competition between rational bookmakers to attract punters drive them to this equilibrium where they offer overly attractive odds to the high-ranked player for GS? Assuming that these competitive forces differ across MS and GS, this explanation could explain the findings using the bookmaker data.[31] However, it cannot explain the findings in the betting exchange data, where the odds on the two players are set through the trades of many different punters, who would only accept bets at favorable odds.[32] Moreover, this explanation requires that these competitive forces are absent in the women's bookmaker sample, where $\Delta_{bias} = 0$.

GS matches are more prestigious events, thus more highly publicized than MS matches. Could biases in GS matches be larger ($\Delta_{bias} < 0$) because these matches attract relatively more unsophisticated bettors who hold more biased expectations? This "noise-trader" explanation could explain the findings with the betting exchange data for men, but it cannot explain the findings with the bookmaker data where the same agents are setting odds for both MS and GS matches.[33] In addition, the biases found in the women's betting exchange data, where GS matches are also more prestigious and more publisized, are opposite to those found in the men data ($\Delta_{bias} > 0$).

Another alternative explanation is that there is a systematic error in π that is related to agent's preferences, which is asymmetric across MS and GS. For example, if bookmaker's risk attitude changes for GS matches due to, for example, the higher betting volumes, they will distort the odds offered for these matches further from risk-neutral probabilities. Thus, it may be the case that the differentials in *bias* across the two match formats reflect such distortions due to shifts in preferences for GS matches. This explanation could explain our findings for the men's bookmaker data, however it cannot

[31]Because bookmakers set the odds for *both* players and charge a vig, they have some room to strategically manoeuvre the odds to attract more volume.

[32]For example, a punter who has a subjective probability that the high-ranked player wins equal to π would only accept a bet if the odds up to that point in the market imply a probability that is less than π. Otherwise, this punter will choose to stay out of the market and will not trade.

[33]Moreover, as shown by the analysis in the Appendix, if bettors were more unsophisticated in GS matches, then rational bookmakers would have a stronger incentive to do the opposite from what we observe in the data, i.e., offer excessively low odds on the high-ranked player to win a GS match.

explain the findings with the betting exchange data, where odds are set competitively in a market setting.[34] Moreover, for bookmaker data, this explanation requires that these shifts in preferences do not occur for women's GS matches ($\Delta_{bias} = 0$ in this sample), which also entail higher betting volumes.

Overall, these alternative explanations can explain some of our findings, but require additional assumptions to explain all of them. Although we cannot conclusively rule them out with our data, we believe it is unlikely that our findings are driven by these explanations.

4 Concluding Remarks

We conduct a field test of Bayesian reasoning by examining whether agents form expectations by placing a larger weight on cues that are more informative with lower process variance. We use subjective probabilities inferred from odds on the outcomes of tennis matches, exploiting *exogenous* variation in process variance related to the format with which tennis matches are played.

Our findings are consistent with *"process variance neglect,"* i.e., bookmakers are not adjusting their subjective probabilities sufficiently to reflect changes in process variance. This result is robust to inferring subjective probabilities from odds offered by professional bookmakers or odds achieved on a person-to-person betting exchange. Moreover, the bias is costly, i.e., the tendency of bookmakers to insufficinely adjust their probabilities lowers their profitability.

In the tennis data the change in process variance between MS and GS matches is relatively intuitive, and easy to observe from historical data. Moreover, clear feedback on the quality of the decicisions is availiable after the match is finished and uncertainty is resolved. Our findings show that agents are unable to engage in Bayesian reasoning even in these simple conditions.

[34]The risks faced by individual punters on the betting exchange are the same across MS and GS, so there is no reason to expect that risk preferences change for GS matches.

How wide-spread are biases related to process variance neglect in other domains of decisions under uncertainty like, for example, when agents calculate the expected returns of different investments? In such cases, the relevant cues - and therefore the process variance considerations - are largely case-specific, making them more difficult to process. Moreover, feedback on the quality of the decisions is extremely noisy. We suspect that biases in expectations related to process variance neglect are more severe in such situations.

References

[1] S. Andersen, S. Ertaç, U. Gneezy, M. Hoffman, and J. A. List, "Stakes matter in ultimatum games," *American Economic Review*, pp. 3427–3439, 2011.

[2] C. Antoniou, J. A. Doukas, and A. Subrahmanyam, "Investor sentiment, beta, and the cost of equity capital," *Management Science*, vol. 62, no. 2, pp. 347–367, 2016.

[3] C. Antoniou, G. W. Harrison, M. I. Lau, and D. Read, "Information characteristics and errors in expectations: Experimental evidence," *Journal of Financial and Quantitative Analysis*, 2015.

[4] C. F. Camerer, "Do biases in probability judgment matter in markets? experimental evidence," *American Economic Review*, vol. 77, no. 5, pp. 981–997, 1987.

[5] D. L. Chen, T. J. Moskowitz, and K. Shue, "Decision-making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires," *Chicago Booth Research Paper*, 2016.

[6] K. Croxson and J. James Reade, "Information and efficiency: Goal arrival in soccer betting," *Economic Journal*, vol. 124, no. 575, pp. 62–91, 2014.

[7] K. Daniel and S. Titman, "Market reactions to tangible and intangible information," *Journal of Finance*, vol. 61, no. 4, pp. 1605–1643, 2006.

[8] W. F. De Bondt and R. H. Thaler, "Do security analysts overreact?" *The American Economic Review*, pp. 52–57, 1990.

[9] J. Del Corral and J. Prieto-Rodríguez, "Are differences in ranks good predictors for grand slam tennis matches?" *International Journal of Forecasting*, vol. 26, no. 3, pp. 551–563, 2010.

[10] W. Edwards, "Conservatism in human information processing," *Formal representation of human judgment*, vol. 17, p. 51, 1968.

[11] E. F. Fama, "Market efficiency, long-term returns, and behavioral finance," *Journal of Financial Economics*, vol. 49, no. 3, pp. 283–306, 1998.

[12] D. Forrest and I. McHale, "Anyone for tennis (betting)?" *European Journal of Finance*, vol. 13, no. 8, pp. 751–768, 2007.

[13] D. M. Grether, "Bayes rule as a descriptive model: The representativeness heuristic," *Quarterly Journal of Economics*, pp. 537–557, 1980.

[14] ——, "Testing bayes rule and the representativeness heuristic: Some experimental evidence," *Journal of Economic Behavior & Organization*, vol. 17, no. 1, pp. 31–57, 1992.

[15] D. Griffin and A. Tversky, "The weighing of evidence and the determinants of confidence," *Cognitive Psychology*, vol. 24, no. 3, pp. 411–435, 1992.

[16] R. M. Griffith, "Odds adjustments by american horse-race bettors," *The American Journal of Psychology*, vol. 62, no. 2, pp. 290–294, 1949.

[17] M. S. Haigh and J. A. List, "Do professional traders exhibit myopic loss aversion? an experimental analysis," *Journal of Finance*, vol. 60, no. 1, pp. 523–534, 2005.

[18] G. W. Harrison and J. A. List, "Field experiments," *Journal of Economic literature*, pp. 1009–1055, 2004.

[19] C. A. Holt and A. M. Smith, "An update on bayesian updating," *Journal of Economic Behavior & Organization*, vol. 69, no. 2, pp. 125–134, 2009.

[20] D. Kahneman, P. Slovic, and E. Tversky, Amos, *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press, 1982.

[21] D. Kahneman and A. Tversky, "Subjective probability: A judgment of representativeness," *Cognitive psychology*, vol. 3, no. 3, pp. 430–454, 1972.

[22] F. Klaassen and J. R. Magnus, *Analyzing Wimbledon: The power of statistics.* Oxford University Press, 2014.

[23] F. J. Klaassen and J. R. Magnus, "Are points in tennis independent and identically distributed? evidence from a dynamic binary panel data model," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 500–509, 2001.

[24] J. Lakonishok, A. Shleifer, and R. W. Vishny, "Contrarian investment, extrapolation, and risk," *Journal of Finance*, vol. 49, no. 5, pp. 1541–1578, 1994.

[25] S. D. Levitt, "Why are gambling markets organised so differently from financial markets?*," *Economic Journal*, vol. 114, no. 495, pp. 223–246, 2004.

[26] S. D. Levitt and J. A. List, "Viewpoint: On the generalizability of lab behaviour to the field," *Canadian Journal of Economics/Revue canadienne d'économique*, vol. 40, no. 2, pp. 347–370, 2007.

[27] ——, "What do laboratory experiments measuring social preferences reveal about the real world?" *Journal of Economic Perspectives*, pp. 153–174, 2007.

[28] S. D. Levitt, J. A. List, and D. H. Reiley, "What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments," *Econometrica*, vol. 78, no. 4, pp. 1413–1434, 2010.

[29] J. Li and J. Yu, "Investor attention, psychological anchors, and stock return predictability," *Journal of Financial Economics*, vol. 104, no. 2, pp. 401–419, 2012.

[30] J. A. List, "Neoclassical theory versus prospect theory: Evidence from the marketplace," *Econometrica*, vol. 72, no. 2, pp. 615–625, 2004.

[31] T. J. Moskowitz, "Asset pricing and sports betting," *Chicago Booth Research Paper*, no. 15-26, 2015.

[32] L. J. Savage, "Elicitation of personal probabilities and expectations," *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 783–801, 1971.

[33] M. A. Smith, D. Paton, and L. V. Williams, "Do bookmakers possess superior skills to bettors in predicting outcomes?" *Journal of Economic Behavior & Organization*, vol. 71, no. 2, pp. 539–549, 2009.

[34] E. Snowberg and J. Wolfers, "Explaining the favorite-longshot bias: Is it risk-love or misperceptions?" *Journal of Political Economy*, vol. 118, no. 4, pp. 723–746, 2010.

[35] J. Wolfers and E. Zitzewitz, "Prediction markets," *The Journal of Economic Perspectives*, vol. 18, no. 2, pp. 107–126, 2004.

Appendix

Expected Profits in GS

In this section, we calculate the expected profit to a *rational* and *risk-neutral* bookmaker in GS with and without bias. This bookmaker understands the correct probability that each player wins the match, but chooses to "salt" this probability with a bias. This is analysis is useful in order to illustrate the conditions under which a bias is suboptimal.

Assume that the correct probability that the high-ranked player wins the match is p_{HR}, and correspondingly for the low ranked player is $p_{LR} = 1 - p_{HR}$. The bookmaker starts from this probability, and adjusts it to reflect a vigorish, v, and a bias, b, arriving at p^*_{HR} using the expression below:

$$p^*_{HR} = (p_{HR} + b) \times (1 + v)$$

Correspondingly, the adjusted probability for the low ranked player is:

$$p^*_{LR} = ((1 - p_{HR}) - b) \times (1 + v)$$

The odds offered for the high- and low-ranked player are thus $O_{HR} = \frac{1}{p^*_{HR}}$ and $O_{LR} = \frac{1}{p^*_{LR}}$, respectively. Furthermore, Vol_{HR} and Vol_{HR} are the volumes that back the high- and low-ranked player respectively, and $TotVol = Vol_{HR} + Vol_{LR}$ is the total volume staked. The expected profit for the rational bookmaker is therefore:

$$E(\mathit{\Pi}) = TotVol - p_{HR} \times Vol_{HR\times} O_{HR} - (1 - p_{HR}) \times (Vol_{HR}) \times O_{LR}$$

Dividing through by $TotVol$ to get expected profits as a proportion of the volume staked:

$$E(R\mathit{\Pi}) = 1 - \frac{p_{HR} \times RVol_{HR}}{(p_{HR}+b)\times(1+v)} - \frac{(1-p_{HR})\times(1-RVol_{HR})}{((1-p_{HR})-b)\times(1+v)}$$

where $RVol_{HR} = \frac{Vol_{HR}}{TotVol}$, and $RVol_{LR} = 1 - RVol_{HR}$.

We conduct a simple calibration to show how $E(R\mathit{\Pi})$ varies with b setting $p_{HR} = 0.79$ and $v = 0.05$ (Table 2, GS), and plotting $E(R\mathit{\Pi})$ for different values of b and $RVol_{HR}$.

The results are shown in Figure 2. Starting with the black solid line where $RVol_{HR} = p_{HR}$, we observe that $E(R\Pi)$ is inverse U-shaped with a unique maximum at $b = 0$.[35] When $RVol_{HR} < p_{HR}$ $E(R\Pi)$ increases as b becomes smaller, and this relationship is steeper as $RVol_{HR}$ decreases. Conversely, when $RVol_{HR} > p_{HR}$ $E(\Pi)$ decreases as b becomes smaller, and this relationship is steeper as $RVol_{HR}$ increases.

Are bookmakers acting optimally by setting odds at a negative b, as we show in Table 4? From the betting exchange data in Table 5, we see that punters show a strong preference for the high-ranked player with $RVol_{HR} = 0.87 > p_{HR} = 0.79$. Under this scenario (as shown by Figure 1), it is suboptimal for a rational bookmaker to move away from fair odds and offer "salted" odds at a lower b. A strategic bookmaker should do exactly the opposite, i.e., offer salted odds at a higher b.

The analysis in Figure 2 assumes that $RVol_{HR}$ does not change in response to variations in b. However, since b influences offered odds it influences $RVol_{HR}$. Our last calculation in this section estimates changes in expected profits in GS matches when the rational bookmaker moves from fair odds to biased odds with a negative b, by taking into account changes in $RVol_{HR}$. Setting $b = -3.5\%$, $p_{HR} = 0.79$ (Table 4) and $RVol_{HR} = 0.87$ (Table 5) we find that $E(R\Pi) = 3.06\%$. Now, suppose a hypothetical scenario where the bookmaker sets $b = 0$. This would affect the stated odds and would thus influence $RVol_{HR}$. To get an estimate of $RVol_{HR}$ in this hypothetical scenario we can use the model below, estimated using the betting exchange data for GS matches:

$$RVol_{HR,i} = \alpha + \beta \times O_{HR,i} + \epsilon_i$$

We find that α is equal to 1.36 and β is equal to -0.35, both highly statistically significant. This shows that when the odds on the high-ranked player increase (i.e., when he becomes less favorite), $RVol_{HR}$ decrases, i.e., volume is bet on the two players more equally. This result is sensibe, given the strong preference of punters to bet on the favorite. Using the model above we find that when b is set to 0 and the odds oferred

[35]Moreover, this is the only point at which the bookmaker faces no risk, i.e., the standard deviation of relative profits is 0. Of course this is insignificant for the risk-neutral bookmaker.

to the high-ranked player decrease (i.e., he becomes more favorite), $RVol_{HR}$ increases to 93.2%. Expected profits in this hypothetical scenario (with $p_{HR} = 0.79$) equal 5.12%, up by 2.06% from the scenario with a b of -3.5%.

[Insert Figure 2 here]

Figure 1: Probabilities in MS vs GS

This figure depicts average objective probabilities (blue bars) and average subjective probabilities (red bars) for MS and GS matches.

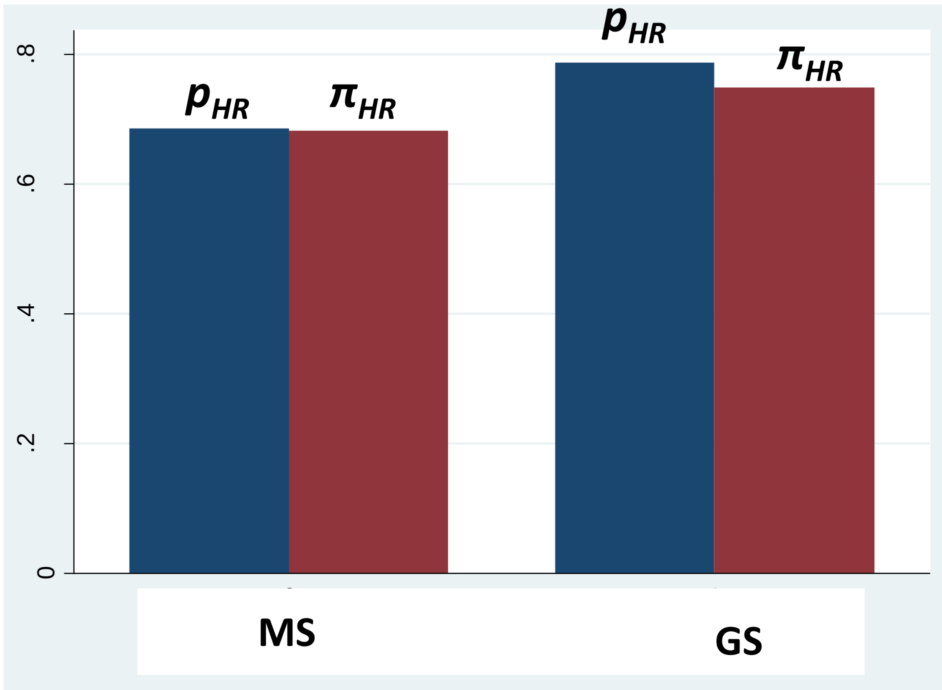

Figure 2: Expected Profits and Bias

This figure depicts Expected Profits (Y-axis) for different levels of bias (X-axis), following the procedure explained in Appendix 1.

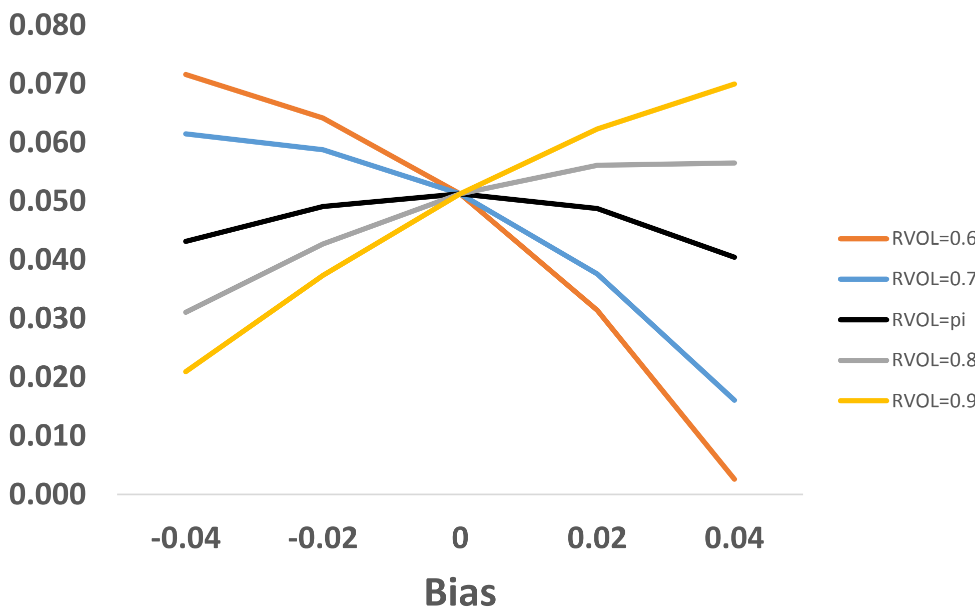

Table 1: Observations by Tournament and Year

This table shows a breakdown of the data used in the analysis sorted by tournament and year. The data are retrieved from www.tennis-data.co.uk. Panel A contains data for Grand Slam (GS) matches which are played in a best-out-of-five format, and Panel B for ATP World Tour Masters 1000 (MS) matches, which are played in a best-out-of-three format. We apply the following filters to create our final sample: (*i*) we retain only completed matches, and drop matches with (*ii*) missing rankings information, (*iii*) missing odds for either player, or (*iv*) matches that entail a negative vig. We also drop matches if (*iv*) both players ranked outside the top 100 players in the world at the start of the tournament, and (*vi*) matches where the high-ranked player is indicated as an outsider by the bookmakers even though he is ranked by at least 15 places higher than the low-ranked player. Our final sample contains 9,046 matches from 2005-2014.

	Year										
Tournament	05	06	07	08	09	10	11	12	13	14	Total
Panel A: GS											
Australian Open	106	106	107	112	109	108	107	104	113	95	1,067
French Open	100	99	98	107	108	110	111	110	105	112	1,060
US Open	107	106	106	111	104	105	102	110	105	106	1,062
Wimbledon	102	102	95	97	99	105	105	104	100	105	1,014
Other	6	6	2	0	0	0	0	0	0	0	14
Total	421	419	408	427	420	428	425	428	423	418	4,217
Panel B: MS											
Cincinatti	57	59	47	42	46	44	47	51	46	51	490
Hamburg	53	53	48	47	0	0	0	0	0	0	201
Indian Wells	79	75	81	77	78	76	70	79	74	81	770
Madrid	40	36	43	41	43	45	49	49	53	50	449
Miami	74	68	76	79	79	78	77	80	73	74	758
Monte Carlo	50	59	46	45	50	48	48	50	47	48	491
Montreal	55	0	48	0	49	0	50	0	48	0	250
Paris	37	38	43	38	41	43	41	42	43	40	406
Rome	54	46	46	42	51	49	45	49	51	47	480
Shanghai	0	0	0	0	44	47	48	52	49	50	290
Toronto	0	51	0	51	0	47	0	44	0	51	244
Total	499	485	478	462	481	477	475	496	484	492	4,829

Table 2: Descriptive Statistics

This table shows descriptive statistics for the main variables used in our analysis. HR_{Odds} and LR_{Odds} are the average decimal odds offered by the betting houses that the high-ranked (low-ranked) player win a match. *Vig* is the housetake, which is obtained by summing the inverse of the odds for the high and low ranked players and subtracting one. HR_{Rank} and LR_{Rank} are the rankings for the high and low ranked player, respectively, and $diff_{Rank}$ is the difference in the ranking between the two players. $\hat{p}$ is the estimated "objective" probability, obtained from averaging the predicted values from the logit model shown in Equation (1) for MS and GS. D_{HR} is a dummy that equals 1 if the high-ranked player has won the match, and 0 otherwise. π_{HR} is the subjective probability that the high-ranked player wins the match, as derived from the bookmaker odds. The sample consists of 9,046 matches that satisfy the criteria outlined in Table 1. The last row in the table shows the average prize money collected by the winner of an *MS* and a GS tournament using 2015 prize money.

Variable	Mean	σ	Min	Q1	Median	Q3	Max
Panel A: GS							
HR_{Odds}	1.34	0.41	1.00	1.09	1.23	1.45	6.79
LR_{Odds}	6.17	5.79	1.09	2.70	4.12	7.33	60.00
Vig	0.05	0.01	0.00	0.05	0.05	0.06	0.08
HR_{Rank}	24.9	23.12	1	7	18	35	100
LR_{Rank}	95.42	93.11	2	47	78	113	1,370
$diff_{Rank}$	70.52	89.41	1	22	49	87	1,369
$\hat{p}$	0.79	0.10	0.60	0.70	0.79	0.87	0.99
D_{HR}	0.79	0.41					
π_{HR}	0.75	0.15	0.14	0.65	0.77	0.87	0.98
Prize Money ($)	2,525,000						
Panel B: MS							
HR_{Odds}	1.48	0.46	1.01	1.19	1.38	1.61	7.45
LR_{Odds}	4.11	3.26	1.09	2.28	3.02	4.59	32.80
Vig	0.06	0.01	0.01	0.05	0.06	0.06	0.09
HR_{Rank}	20.56	19.01	1	6	15	30	100
HR_{Rank}	65.56	78.05	2	29	49	77	1,517
$diff_{Rank}$	45.00	73.85	1	12	27	53	1,460
$\hat{p}$	0.69	0.12	0.50	0.58	0.67	0.78	0.98
D_{HR}	0.69	0.46					
π_{HR}	0.68	0.15	0.13	0.59	0.69	0.79	0.97
Prize Money ($)	790,000						

Table 3: Likelihood that high-ranked player wins an MS and a GS Match

This table presents results from logit models, where the dependent variable is takes the value of 1 if the high-ranked player has won the match, and 0 otherwise. *GS* is a dummy variable that equals 1 if the match is GS, and 0 otherwise. *Rskill* is calculated as the log (low-ranked player ranking) - log (high-ranked player ranking) for MS (GS) matches. The table reports marginal effects associated with each of the independent variables. The sample consists of 9,046 matches that satisfy the criteria outlined in Table 1. The robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Variable	(1)	(2)	(3)
GS	0.101***	0.074***	0.072***
	[0.009]	[0.009]	[0.009]
Rskill		0.121***	0.123***
		[0.005]	[0.005]
Surface F.E.	NO	NO	YES
Year F.E.	NO	NO	YES
Round F.E.	NO	NO	YES
N	9,046	9,046	9,046
pseudo-R^2	0.011	0.071	0.073

Table 4: Biases in Subjective Probabilities in MS vs. GS

This table reports biases in subjective probabilities for GS and MS tennis matches. In Panel A we present univariate analysis and in Panel B multivariate analysis. In Panel A π and $\hat{p}$ denote the average subjective and the average (estimated) objective probabilities that the high-ranked player wins an MS or a GS match. $\hat{p}$ is obtained by averaging the predicted values from the logit model shown in Equation (1) for MS and GS, and π is derived from the average odds on the high-ranked player to win the match offered by the bookmakers. *Bias* is the average difference between π and $\hat{p}$ for MS and GS. In Panel B we present results from OLS regressions with an intercept. The dependent variable is *Bias* and the independent variables are *GS* and *Rskill*, defined as in Table 3. The sample consists of 9,046 matches from 2005 to 2014 that satisfy the criteria outlined in Table 1. In Panel B the robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Panel A: Univariate			
	MS	GS	GS - MS
π	0.682	0.748	
$\hat{p}$	0.686	0.786	
Bias	-0.004	-0.038	-0.034***
			[0.002]
N	4,829	4,217	

Panel B: Multivariate		
Variable	(1)	(2)
GS	-0.033***	-0.035***
	[0.002]	[0.002]
Rskill	-0.007***	-0.005***
	[0.001]	[0.001]
N	9,046	9,046
R^2	0.029	0.054
Surface F.E.	NO	YES
Year F.E.	NO	YES
Round F.E.	NO	YES

Table 5: Descriptive Statistics: Betting Exchange

This table shows descriptive statistics for the main variables used in the analysis with the Betting Exchange data. HR_{Odds} and LR_{Odds} are the average decimal odds offered by the betting houses that the high-ranked (low-ranked) player win a match. *Vig* is the housetake, which is obtained by summing the inverse of the odds for the high and low ranked players and subtracting one. HR_{Rank} and LR_{Rank} are the rankings for the high and low ranked player, respectively, and $diff_{Rank}$ is the difference in the ranking between the two players. *Totvol* is the total volume that backs the high-ranked player ($000's), and *Rvol* is the proportion of the total volume bet in the match on both players that backs the high-ranked player. $\hat{p}$ is the estimated "objective" probability, obtained from averaging the predicted values from the logit model shown in Equation (1) for MS and GS. D_{HR} is a dummy that equals 1 if the high rank player has won the match, and 0 otherwise.π_{HR} is the subjective probability that the high-ranked player wins the match, as derived from the bookmaker odds. The sample consists of 2,728 MS matches and 2,074 GS matches that satisfy the criteria outlined in Table 1 except criterion *(iv)*, since in this sample there are no matches with a negative vig. For this sample we drop matches where the vig is greater than 0.05.

Variable	Mean	σ	Min	Q1	Median	Q3	Max
Panel A: GS							
HR_{Odds}	1.38	0.47	1.01	1.09	1.24	1.50	8.60
LR_{Odds}	10.30	14.42	1.12	2.94	4.90	11.00	100.00
Vig	0.01	0.01	0.00	0.00	0.01	0.01	0.05
HR_{Rank}	24.01	22.72	1	6	17	34	100
LR_{Rank}	92.43	88.5	2	44	76	112	1,120
$diff_{Rank}$	68.42	83.97	1	21	48	86	1,063
$TotVol$	240	460	4	47	88	220	6,200
$RVol$	0.87	0.19	0.02	0.86	0.94	0.97	1
$\hat{p}$	0.79	0.10	0.59	0.71	0.80	0.87	0.99
D_{HR}	0.79	0.41					
π_{HR}	0.77	0.16	0.12	0.66	0.80	0.91	0.99
Panel B: MS							
HR_{Odds}	1.52	0.54	1.01	1.20	1.40	1.66	9.00
LR_{Odds}	5.72	7.40	1.12	2.46	3.45	5.60	90.00
Vig	0.01	0.01	0.01	0.00	0.01	0.01	0.05
HR_{Rank}	19.82	18.89	1	5	14	29	99
LR_{Rank}	62.58	68.37	2	27	46	76	1,147
$diff_{Rank}$	42.76	64.22	1	11	25	51	1,074
$TotVol$	170	220	2.3	40	86	210	2,200
$RVol$	0.82	0.23	0.02	0.81	0.91	0.96	1
$\hat{p}$	0.70	0.12	0.50	0.59	0.69	0.80	0.99
D_{HR}	0.70	0.46					
π_{HR}	0.70	0.16	0.11	0.60	0.71	0.82	0.99

Table 6: Biases in Subjective Probabilities in MS vs. GS: Betting Exchange Data

This table reports biases in subjective probabilities for GS and MS tennis matches. In Panel A we present univariate analysis and in Panel B multivariate analysis. In Panel A π and $\hat{p}$ denote the average subjective and the average (estimated) objective probabilities that the high-ranked player wins an MS or a GS match. $\hat{p}$ is obtained by averaging the predicted values from the logit model shown in Equation (1) for MS and GS, and π is derived from the odds on the high-ranked player to win the match as obtained from the betting exchange. *Bias* is the average difference between π and $\hat{p}$ for MS and GS. In Panel B we present results from OLS regressions that include an intercept. The dependent variable is *Bias* and the independent variables are *GS* and *Rskill*,defined as in Table 3. In Column (3) of Panel B $VolD$ is a dummy calculated as follows: We first rank all matches according to type of match (MS vs. GS) and match round (1 to 7). Within each of these 14 groups we sort all matches according to the volume that backs the high ranked player, noting the median of the distribution. $VolD$ equals to one if the specific match is above this median, and 0 otherwise. The sample consists of 4,802 matches from 2009 to 2014 that satisfy the criteria outlined in Table 5. In Panel B the robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Panel A: Univariate

	MS	GS	GS - MS
π	0.702	0.770	
$\hat{p}$	0.697	0.789	
Bias	0.005	-0.019	-0.024***
			[0.003]
N	2,728	2,074	

Panel B: Multivariate

Variable	(1)	(2)	(3)
GS	-0.024***	-0.027***	-0.029***
	[0.003]	[0.003]	[0.005]
$Rskill$	-0.001	0.002	-0.014***
	[0.001]	[0.002]	[0.002]
$VolD$			0.067***
			[0.004]
$VolD \times GS$			0.013**
			[0.006]
N	4,802	4,802	4,802
R^2	0.011	0.024	0.107
Surface F.E.	NO	YES	YES
Year F.E.	NO	YES	YES
Round F.E.	NO	YES	YES

Table 7: Robustness Checks

In this table we conduct various robustness checks. In Column (1) the dependent variable is $Bias = \pi_{HR} - D_{HR}$, where D_{HR} is a dummy that equals 1 if the high ranked player has won the match, and 0 otherwise. In Column (2) we estimate a quantile regression model, expressing the median of the conditional distribution as a linear function of the indepedent variables. In Column (3) we use OLS regressions as in Panel B of Table 4 but without dropping observations where both players are ranked outside top 100 players of the world, or where the difference in rankings is greater than 15 positions but the high-ranked player is indicated as an outsider by the odds. In Column (4) we add an additional regressor in the model, *RStreak*, calculated as the difference in the proportion of games won by the high and low ranked player for the two previous tournaments. For this test we estimate $\hat{p}$ using an expanded version of the logit model in (1) that includes *RStreak* as an additional independent variable. In all columns the independent variables *GS* and *Rskill* are defined as in Table 3. All models include an intercept term. In Panel A (B) we derive subjective probabilities using bookmaker (betting exchange) odds. The robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Panel A: Bookmaker				
Variable	(1)	(2)	(3)	(4)
GS	-0.033***	-0.023***	-0.025***	-0.028***
	[0.009]	[0.003]	[0.003]	[0.003]
Rskill	-0.007**	-0.003**	-0.010***	-0.007***
	[0.003]	[0.001]	[0.001]	[0.001]
RStreak				-0.018***
				[0.003]
N	9,046	9,046	10,266	6,243
R^2	0.002	0.007	0.015	0.025
Panel B: Betting Exchange				
Variable	(1)	(2)	(3)	(4)
GS	-0.024**	-0.011***	-0.020***	-0.029***
	[0.012]	[0.004]	[0.004]	[0.004]
Rskill	-0.001	0.001	0.001	0.004**
	[0.005]	[0.001]	[0.002]	[0.002]
RStreak				-0.026***
				[0.006]
N	4,802	4,802	5,405	3,199
R^2	0.001	0.002	0.005	0.019

Table 8: Bookmaker Profits in MS vs GS

This table reports analysis of the profits earned by bookmakers, presenting results from OLS regressions that include an intercept (not reported). In Panel A we present results using bookmaker odds and in Panel B betting exchange odds. On the top row we show the depedent variable in each regression. r_{HR} is the return made by the bookmaker for each uit of currency bet on the high ranked player, and is calculated according to Equation (3). $R\Pi$ is an estimate of the actual profit earned by the bookmakers, using volume information from the betting exchange, and is calculated according to Equation (5). The independent variables are *GS* and *Rskill*, defined as in Table 3. The robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	A:Bookmaker		**B: Betting Exchange**	
Variable	1:r_{HR}	2:$R\Pi$	1:r_{HR}	2:$R\Pi$
GS	-0.041***	-0.025**	-0.025	-0.021**
	[0.014]	[0.011]	[0.020]	[0.011]
Rskill	-0.010*	-0.005	0.006	0.002
	[0.006]	[0.005]	[0.009]	[0.004]
Surface F.E.	YES	YES	YES	YES
Year F.E.	YES	YES	YES	YES
Round F.E.	YES	YES	YES	YES
N	9,046	9,046	4,802	4,802
R^2	0.004	0.005	0.005	0.004

Table 9: Descriptive Statistics: Women's Bookmaker Data

This table shows descriptive statistics for the main variables used in the analysis of the women bookmaker data. *HROdds* (*LROdds*) are the average decimal odds offered by the betting houses that the high-ranked (low-ranked) player win a match. *Vig* is for the over-run, which is obtained by summing the inverse of the odds for the high- and low-ranked player and subtracting one. $Rank_H$ and $Rank_L$ are the rankings for the high- and low-ranked player, respectively, and *diff* is the difference in the ranking between the two players. $\hat{p}$ is the estimated "objective" probability, obtained from averaging the predicted values from the logit model shown in Equation (1) for MS and GS. D_{HR} is a dummy that equals 1 if the high rank player has won the match, and 0 otherwise. π_{HR} is the subjective probability that the high-ranked player wins the match, as derived from the bookmaker odds. The sample consists of 2,491 MS matches and 3,532 GS matches from 2007 to 2014 that satisfy the criteria outlined in Table 1. The last row in the table shows the average prize money collected by the winner of an *MS* and a GS tournament using 2015 prize money.

Variable	Mean	σ	Min	Q1	Median	Q3	Max
			Panel A: GS				
HR_{Odds}	1.37	0.44	1.01	1.11	1.26	1.49	9.40
LR_{Odds}	5.24	4.08	1.06	2.58	3.78	6.42	35.40
Vig	0.06	0.01	0.00	0.05	0.06	0.06	0.08
$Rank_H$	25.38	23.02	1	8	18	35	100
$Rank_L$	91.44	82.63	2	48	78	110	1,208
$diff$	66.06	80.09	1	22	49	85	1,195
$\hat{p}$	0.75	0.11	0.55	0.66	0.75	0.83	0.98
D_{HR}	0.75						
π_{HR}	0.73	0.15	0.10	0.63	0.75	0.85	0.97
Prize Money ($)	2,525,000						
			Panel B: MS				
HR_{Odds}	1.46	0.42	1.01	1.19	1.35	1.59	5.47
LR_{Odds}	3.93	2.53	1.14	2.32	3.14	4.54	22.00
Vig	0.06	0.01	0.02	0.05	0.06	0.06	0.09
$Rank_H$	22.88	20.80	1	7	16	34	100
$Rank_L$	69.78	73.94	2	31	53	84	1,063
$diff$	46.90	68.84	1	13	29	56	988
$\hat{p}$	0.70	0.11	0.53	0.61	0.70	0.79	0.98
D_{HR}	0.70						
π_{HR}	0.69	0.14	0.17	0.59	0.70	0.79	0.96
Prize Money ($)	665,000						

Table 10: Descriptive Statistics: Women's Betting Exchange Data

This table shows descriptive statistics for the main variables used in the analysis of the women betting exchange data. *HROdds* (*LROdds*) are the odds obtained in the betting exchange that the high-ranked (low-ranked) player win the match. *Vig* is for the over-run, which is obtained by summing the inverse of the odds for the high and low ranked player and subtracting one. $Rank_H$ and $Rank_L$ are the rankings for the high and low ranked player, respectively, and *diff* is the difference in the ranking between the two players. *Totvol* is the total volume that backs the high-ranked player ($000's), and *Rvol* is the proportion of the total volume bet in the match on both players that backs the high-ranked player. $\hat{p}$ is the estimated "objective" probability, obtained from averaging the predicted values from the logit model shown in Equation (1) for MS and GS.D_{HR} is a dummy that equals 1 if the high rank player has won the match, and 0 otherwise. π_{HR} is the subjective probability that the high-ranked player wins the match, as derived from the bookmaker odds. The sample consists of 1,408 MS matches and 2,210 GS matches that satisfy the criteria outlined in Table 5.

Variable	Mean	σ	Min	Q1	Median	Q3	Max
			Panel A: GS				
HR_{Odds}	1.43	0.53	1.01	1.15	1.32	1.55	11.50
LR_{Odds}	6.98	8.68	1.09	2.74	4	7.20	100
Vig	0.01	0.01	0.00	0.01	0.01	0.01	0.05
$Rank_H$	25.30	22.98	1	8	18	35	100
$Rank_L$	90.61	81.05	2	48	77	110	1,208
$diff$	65.30	78.04	1	22	49	85	1,195
$TotVol$	130	250	0.36	20.06	44.40	120	2,900
$RVol$	0.86	0.20	0.01	0.85	0.93	0.97	1
$\hat{p}$	0.74	0.11	0.55	0.65	0.74	0.82	0.97
D_{HR}	0.74	0.44					
π_{HR}	0.74	0.16	0.09	0.64	0.75	0.86	0.99
			Panel B: MS				
HR_{Odds}	1.52	0.47	1.02	1.24	1.41	1.67	6.20
LR_{Odds}	4.57	4.20	1.17	2.44	3.30	4.85	46.00
Vig	0.01	0.01	0.00	0.01	0.01	0.01	0.05
$Rank_H$	22.28	20.48	1	6	15	34	100
$Rank_L$	68.41	71.09	2	32	52	80	821
$diff$	46.13	65.70	1	13	29	54	748
$TotVol$	56.03	82.84	0.41	12.83	26.06	59.19	690
$RVol$	0.82	0.23	0.02	0.79	0.91	0.96	1
$\hat{p}$	0.71	0.11	0.54	0.62	0.70	0.79	0.97
D_{HR}	0.71	0.45					
π_{HR}	0.69	0.15	0.16	0.60	0.70	0.80	0.98

Table 11: Likelihood that high-ranked player wins an MS and a GS Match: Women's Data

This table presents results from logit models, where the dependent variable is takes the value of 1 if the high-ranked player has won the match, and 0 otherwise. *GS* is a dummy variable that equals 1 if the match is GS, and 0 otherwise. *Rskill* is calculated as the log (low-ranked player ranking) - log (high-ranked player ranking) for MS (GS) matches. The table reports marginal effects associated with each of the independent variables. In Panel A we present results for the bookmaker sample that consists of matches that satisfy the criteria in Table 1, and in Panel B for the betting exchange sample that consists of matches that satisfy the criteria in Table 5. The robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Variable	(1)	(2)	(3)
Panel A: Bookmaker			
GS	0.043***	0.018	0.016
	[0.012]	[0.011]	[0.011]
Rskill		0.121***	0.126***
		[0.006]	[0.006]
Surface F.E.	NO	NO	YES
Year F.E.	NO	NO	YES
Round F.E.	NO	NO	YES
N	6,023	6,023	6,023
pseudo-R^2	0.002	0.060	0.064
Panel B: Betting Exchange			
GS	0.030**	0.009	0.004
	[0.015]	[0.015]	[0.015]
Rskill		0.117***	0.121***
		[0.008]	[0.008]
Surface F.E.	NO	NO	YES
Year F.E.	NO	NO	YES
Round F.E.	NO	NO	YES
N	3,618	3,618	3,618
pseudo-R^2	0.001	0.055	0.059

Table 12: Biases in Subjective Probabilities in MS vs. GS: Women's data

This table reports biases in subjective probabilities for GS and MS women's tennis matches, presenting results from OLS regressions that include an intercept. The dependent variable is *Bias* and the independent variables are *GS* and *Rskill*, defined as in Table 3. In columns (1) and (2) we present results when subjective probabiities are derived from bookmaker odds, and in columns (3) and (4) Panel B when subjective probabiities are derived from odds achieved on the betting exchange. The sample consists of matches that satisfy the criteria outlined in Table 1. The robust standard errors shown in brackets are calculated using the Huber-White estimator. *,**,*** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

	A:Bookmaker		**B: Betting Exchange**	
Variable	(1)	(2)	(3)	(4)
GS	0.003	0.005*	0.017***	0.017***
	[0.003]	[0.003]	[0.004]	[0.004]
Rskill	-0.008***	-0.008***	-0.000	-0.001
	[0.001]	[0.001]	[0.002]	[0.002]
N	6,023	6,023	3,618	3,618
R^2	0.007	0.007	0.006	0.009
Surface F.E.	NO	YES	NO	YES
Year F.E.	NO	YES	NO	YES
Round F.E.	NO	YES	NO	YES