

Would you pay for transparently useless advice? A test of boundaries of beliefs in the folly of predictions

Nattavudh Powdthavee*

Centre for Economic Performance, London School of Economics

Melbourne Institute of Applied Economic and Social Research, University of Melbourne

Yohanes E. Riyanto

Division of Economics, Nanyang Technological University

8th January 2014

Forthcoming in the *Review of Economics and Statistics*

* *Corresponding author:* Nattavudh Powdthavee

Address: Centre for Economic Performance, London School of Economics, Houghton Street, London, UK, WC2A 2AE.

Email: n.powdthavee@lse.ac.uk.

Tel: +44(0) 7990 815924.

Acknowledgment: We are grateful to Gigi Fosters, Paul Frijters, Daniel Houser, Jack Knetsch, Nuarpear Lekfuangfu, Pongsak Luangaram, Nicolas Motz, Andrew Oswald, Worrawat Sritrakul, Matthew Rabin, David Ribar, Jon Wongswan, Dimitri Vayanos, Marie-Claire Villeval, and participants at workshops in the University of Melbourne, the University of Wollongong, and the University of York for helpful comments on the draft. We also thank Liu Jia, Weerachart Kilenthong, and students and staffs at University of Thai Chamber of Commerce and Nanyang Technological University for their help in carrying out the experiments. The research was supported by the SUG grant provided by Nanyang Technological University, and was carried out whilst the first author was employed at Nanyang Technological University, Singapore.

Abstract

Standard economic models assume that the demand for expert predictions arise only under the conditions in which (i) individuals are uncertain about the underlying process generating the data, and (ii) there is a strong belief that past performances predict future performances. Thus, when it is transparently clear that the event in question is truly random, there should be no market for expert predictions. We set up the strongest possible test of these assumptions in the laboratory. In contrast to the theoretical suggestions made in the literature, people are willing to pay for predictions of truly random outcomes after witnessing only a short streak of accurate predictions *live* in the lab. Potential explanations and implications of such “irrational learning” are discussed in the contexts of economics and finance.

Keywords: behavioral finance; expertise; hot hand; dynamic-inference model; random streaks

JEL: C91; D03

“[Expert] intuition cannot be trusted in the absence of stable regularities in the environment” – Daniel Kahneman in *Thinking, Fast and Slow* (2011)

1. Introduction

Why do humans pay for expert predictions when most future events are predominantly unpredictable? What explains, for example, the significant amount of money spent in the finance industry on people who appear to be commenting about random walks,² payments for services by political and economic forecasters who are often only slightly better at forecasting the future than non-experts,³ or some other false-expert setting?⁴

Economists typically dismiss such behaviors as random error in decision-making. This is the notion that an average person is disinclined to commit such errors, and that people rationally pay for expert predictions only when they are *a priori* uncertain about the underlying data-generating process, yet maintain the belief that some systematic predictions are potentially possible. What this implies is that economists tend to set a high bar for the way people rationalize their actions *ex ante*, i.e., a significant degree of uncertainty would need to be present in our mind before we could be persuaded to pay for a prediction that is ultimately useless.

By contrast, the psychology literature assumes that human beings are hypersensitive at detecting agency, even when none exists, to help them to explain phenomena that cannot be easily explained. This implies that on average people will be happy to pay for advice that is generally counterintuitive if they believe there is an intelligent agent making the predictions for them. Such a divide between the two social-science disciplines in the beliefs of how people’s preferences for expert predictions are formed is scientifically unattractive.

² Hung et al. (2008), Chater et al. (2010).

³ Tetlock (2006).

⁴ http://en.wikipedia.org/wiki/Paul_the_Octopus.

While we cannot deny from innumerable observations that individuals can be induced to believe in the “hot hands” of experts (and consequently have been exploited by charlatans, witch-doctors, fortune tellers, casino operators, and mutual-fund managers the world over for centuries), we know much less about the *boundaries* of such seemingly irrational behaviors. For example, we do not know the extent to which people can be induced to believe in the hot hand of an expert if it is *a priori* clear that the event is truly random and if accurate expert predictions can be explained only by luck and not skills.

Our paper attempts to contribute to this research area by investigating whether people are willing to pay for predictions in a situation in which there is true randomness and predictions are blatantly useless. In a laboratory setting of clearly non-existent expertise, can an average person be challenged to switch from having the correct prior belief that “outcomes are determined by chance and predictions are inherently useless” to the false belief that “predictions provide useful information about the future” – thus leading the person to buy subsequent predictions at a fixed price – if she had recently observed an improbable streak of successful predictions being made in front of her *live*? In contrast to the literature, we found that the answer is yes and that the size of the error made systematically by people is large.

This paper is structured as follows: section 2 presents a brief review of the relevant literature, section 3 outlines the stylized dynamic-inference model, and section 4 describes the data and experimental procedures. Empirical strategy is discussed in section 5. Results are then laid out in section 6. Section 7 provides our discussions, the implications of our results, and the concluding remarks.

2. Background

There is little economic theory in this area. A few exceptions are the work of Matthew Rabin and Dimitri Vayanos. In their papers, Rabin (2002) and Rabin and Vayanos (2010) outline a

model in which believers of “the law of small numbers” – i.e., those who believe that a small sample of signals represents the parent population from which it is drawn (Tversky & Kahneman, 1971) – will be willing to pay for the services of financial analysts after observing randomly occurring streaks of profitable financial performances provided by mutual-fund managers. Such a belief in the hot hand of a financial expert arises because individuals over-infer the financial manager’s ability following a streak of successful stock performances. In other words, an investor who believes that the performance of a mutual fund is a combination of the manager’s skills and luck will, at first, underestimate the likelihood that a manager of average ability will exhibit a streak of above- or below-average performance. Following good or bad streaks, however, the investor will revert to overestimate the likelihood that the manager is above or below average, and so in turn will over-infer that the streak of unusual performance will continue (see Gilovich et al., 1985). Rabin and Vayanos’s model thus predicts that following a streak of “good” signals in settings where there is an element of skill involved in generating such a streak – e.g., a sequence of successful performances by stockbrokers or managers of actively managed funds – believers of the law of small numbers will be happy to pay for real-time price information provided by these financial “experts” even though it is well-documented that actively managed funds do not outperform their market benchmark on average (see, e.g., Fama, 1991; Cahart, 1997). The key assumption here is that, in order to form a belief in the hot hand of a manager, individuals must be unsure about the data-generating process of stock movements in the first place, but nevertheless hold the belief that past returns help to predict future returns (and that past management performances help to predict future performances).

Econometric evidence on the evolution of beliefs in expert predictions is also scarce. Much of the literature in this field tends to focus on situations in which no experts were present to generate predictions of future independent and identically distributed (i.i.d.)

signals. For example, Croson and Sundali (2005) show that, in a game of roulette, casino gamblers tend to bet against a sufficiently long streak rather than with a streak, which is consistent with the “gambler’s fallacy,” while at the same time gamblers tend to bet on more numbers after winning than after losing, which is a behavior consistent with the hot hand effect. Terrell and Farmer (1996) and Terrell (1998) find evidence of the gambler’s fallacy in horse and dog racing. The authors show that gamblers are less likely to bet on repeat winners by post position. For example, if the animal in post position 3 wins a race, then in the next race the (different) animal in post position 3 is significantly under-bet. Using a computerized roulette game, Ayton and Fischer (2004) show that subjects tend to believe in the gambler’s fallacy with respect to the sequence of outcomes of the roulette wheel. Yet when the subjects’ role is to predict the outcomes of the roulette wheel, they tend to over-predict how well or badly they would do at predicting based on their previous streak of predictions, thus exhibiting the hot hand fallacy.

More recently, Guryan and Kearney (2008) found unique evidence of the potential hot hand effect in stores that sell lottery numbers. The authors show that, in the week following the sale of a large-prize-winning ticket, the winning store experiences a significant increase in relative sales for the winning lottery game. Using a unique panel data of lottery players, Jørgensen et al. (2011) present evidence that, while most lottery players tend to pick the same set of numbers week after week without regard to the numbers drawn in the lottery in previous weeks, those people who do change the set do so in such a way that is consistent with the law of small numbers. On average, these “switchers” move away from numbers that have been recently drawn (the gambler’s fallacy), and move toward numbers that are on a streak (the hot hand fallacy). However, in both scenarios – “lucky” lottery stores and “lucky” lottery numbers – the switches in preferences that are consistent with the hot hand effect are often short-lived; for example, Clotfelter and Cook (1991, 1993) and Teller (1994) show that,

shortly after a lottery number wins, individuals are significantly more likely to bet on it. The effect soon diminishes: A few months later the winning number is as popular as the average number.

To the best of our knowledge, only one other study has explicitly tested the implications of expert predictions in a pure i.i.d. setting. In experiments ran by Huber, Kirchler, and Stockl (2010), participants were asked to decide between a risky and a risk-free investment. If they opt for the former they would be given an opportunity to either make the risky investment themselves or delegate the decision to an expert for some fees. There are two equally likely investment outcomes resembling coin flip outcomes. The procedure is implemented by asking participants to either bet on computer-generated outcomes of coin flips, or rely on randomized “expert” predictions. When randomized expert predictions were chosen, participants were told that the experts would make the investment decision by choosing “Head” or “Tail”. The outcome of this investment decision would be determined by a computer-generated coin flip. The researchers were able to show that people who rely on randomized experts tend to pick those experts who have been successful in the past, which is consistent with the hot hand fallacy. Those who decide the outcome of coin flips on their own tend to behave consistently with the gambler’s fallacy, as the frequency of betting on heads (tails) decreases after streaks of heads (tails).

However, there are some potential shortcomings of their study. Firstly, instead of real coins, a “virtual” coin was used in the experiment to generate the signals. The use of virtual coin could have induced the erroneous belief that the coin was not fair and that its outcomes were predictable. Furthermore, participants were specifically told that they are playing an investment game that is designed to replicate an actual asset market. The coin flip is used as an analogy to this investment game. Since the setting is primarily framed as an investment context, it is thus not ‘transparently’ clear whether participants would indeed consider the

setting as equivalent to a pure coin flip decision. It is possible that the investment framing might subconsciously anchor participants' cognitive thinking on the expertness of the experts and give participants a confounding impression that skills were involved in predicting the outcomes of the investment decision.

Second, in their setup participants must pay a surcharge fee and a management fee when they decide to engage an expert to make an investment on their behalf. If they decide to continue relying on the same expert to make investment, they would not need to pay the surcharge fee anymore. Consequently, the net earnings conditional on winning would be higher when they continue using the same expert. This may induce participants who have an inclination to rely on the expert in the first place to continue using the expert to save cost.

Third, the investment decision in their paper is a simple binary decision of whether or not to invest. Participants do not make any decision on the amount to invest and their earnings conditional on winning remain constant. In contrast, we allow participants to vary the amount of bet, thereby varying their earnings conditional on winning. By analyzing the betting amount, we are able to capture the willingness of participants to treat the coin toss predictions more seriously upon observing a streak of correct predictions.

Finally, in their paper, when participants choose not to rely on an expert to make the investment in any particular round, they would never know whether the expert would have made a correct decision in that round. This is because, the advice given by these experts would never be revealed to participants. In contrast, in our paper, even when participants decide not to rely on the coin flip prediction prior to making their bet, they would still be able to know ex-post whether the coin-flip prediction would have been correct. Such information would influence participants' decision to switch from making their own bet to relying on the coin flip prediction.

All in all, two main lessons may be learned from the literature. The first is that humans tend to behave in such a way that is consistent with the gambler's fallacy in situations where the underlying mechanism generating the streak of signals is transparently random. The second is that the hot hand fallacy normally arises in situations where human skills are *a priori* – albeit erroneously – perceived as part of the streak-generating process. Or to quote Huber, Kirchler, and Stockl (2010):

The hot hand belief is usually attributed to human skilled performance, whereas the gambler's fallacy is often attributed to inanimate chance mechanism. (p. 446)

The current article builds on this little literature and sets out to test one of the key assumptions in the model by Rabin and Vayanos (2010)⁵: In a truly random situation in which experts' predictions are *a priori* known to be transparently useless, people's behaviors will be influenced only by the gambler's fallacy (which is a luck-related perception) and not by the hot hand (which is a performance-related perception) of the non-existent expert.

3. Experimental framework

3.1. Data

To investigate the conditions under which people would be willing to pay for predictions of truly random events, a series of laboratory experiments was conducted on voluntary participants in Thailand and Singapore. We ran our first set of experiments in Thailand in December 2011, and the randomly selected participants were undergraduate students at the University of the Thai Chamber of Commerce and Chulalongkorn University in Bangkok ($N=177$). We then ran our second set of experiments in Singapore in March 2012. Here, the

⁵ The model by Rabin and Vayanos (2010) is a more tractable version of an earlier model by Rabin (2002) and a more suitable setup to analyze the hot hand fallacy. For this reason, this paper will adopt the specification of Rabin and Vayanos (2010) rather than that of Rabin (2002).

volunteer participants originated from randomly drawn undergraduate students at Nanyang Technological University ($N=201$). Overall, participants were from various schools and faculties, including humanities and social sciences, engineering, sciences, and business and accounting. We ran 12 sessions in total (4 in Thailand and 8 in Singapore) and were able to recruit approximately 45 people per session in Thailand and 30 people per session in Singapore.

The experiment conducted in Thailand was done manually using pen and paper, while the experiment conducted in Singapore was computerized and programmed with the software z-Tree (Fischbacher, 2007).⁶ This computerized experiment was done in two experimental labs. Upon arriving at the experiment venue, participants were randomly assigned to one of the two labs. Once they entered the labs, there were randomly assigned to cubicles.

Once seated, they were then asked to complete two tasks. The first task involved participants placing bets on the outcomes of five rounds of “fair” coin flips. To ensure the fairness of the coins used in the experiment, from the beginning we made explicitly clear to participants the following:

- (i) The coins will be supplied by the participants rather than the experimenters.
- (ii) The coins will be changed after the second and fourth flip.
- (iii) A volunteer participant in the experiment will be randomly chosen to step out in front of everyone and flip the coin.
- (iv) The coin-flipper will be changed in every round.

Participants were each given an initial endowment with which to make their bets in the five rounds of coin flips. There was a minimum bet of 10 tokens per round, and participants were not allowed to go bankrupt before the final round was reached. Participants in Thailand were given an initial endowment of 100 tokens. Since a few participants in Thailand went bankrupt

⁶ We outline the experimental procedures in Appendix B, which can be found in the online annex.

before the final round, when we ran it in Singapore we decided to give each one a higher endowment level at the start of the experiment of 300 tokens. Placing a correct bet was worth double, and an incorrect one was worth zero in return. Each participant was also given at the beginning of the experiment five numbered envelopes, which were taped on each cubicle's table. Contained within was a "prediction" of the coin flip that *had not happened yet* in each of the numbered rounds. In each round, participants were given an opportunity to pay a fixed price of 10 tokens to see the inside of the corresponding numbered envelope *before a bet was placed and the coin flipped*. All other participants who decided not to pay were instructed to open the corresponding numbered envelope for *free* immediately after the flip in order to view whether the prediction matched the outcome. In addition, great care was taken not to provide any misleading information, e.g., on who made the predictions, whether the predictions were made by an expert or a group of experts, or how the predictions were generated, which could have potentially primed participants into buying (or not buying) the predictions by evoking the impression that the underlying process generating accurate predictions was humanly possible. Participants were then told that the remaining endowment at the end of the fifth round would be converted to either Singapore dollars (SG\$) or Thai baht at the exchange rate of 50 points = SG\$1 (25 Thai baht) or approximately US\$0.9.

To guarantee a significant number of participants received at least four consecutive correct predictions in five rounds of fair coin flips, predictions were generated and assigned in such a way that approximately $\frac{1}{2}$ of N received one correct prediction after round 1, $\frac{1}{2} \times \frac{1}{2}$ of N received two correct predictions after round 2, etc. (see Figure 1). This method of **randomization-in-randomization** (R-in-R) – i.e., the process of randomizing people within the same session into control and treatment groups – ensured that at least 1 out of N participants per session would randomly receive all-correct predictions *irrespective of the actual outcomes of the coin flips*. The R-in-R design also ensures that we have a random split

of participants with divergent beliefs about the predictability of the coin in the control and treatment groups. Of the total of 378 participants from the two countries, 191 received a correct prediction in the first round of coin flips, 92 all-correct predictions after the first two rounds, 48 after the first three rounds, and 23 after the first four rounds.⁷

In order to ensure that participants did not “cheat” by either opening a different envelope to the one that was bought or opening one that had not been purchased in the corresponding round, a more than usual number of research assistants (at least five research assistants per session) were hired to help with the supervising of the experiment. We stationed these research assistants at various locations inside the labs. Given their locations and the size of the labs, it was relatively easy for them to spot any cheating, i.e. participants opened the envelopes without paying or participants opened envelopes meant for other rounds. Throughout the experiment, we did not find any incidence of cheating. Participants were also not forced to open the envelopes during the course of the experiment if they wished not to. All but one participant from Psychology Department chose not to open the envelopes to see any of the predictions. We dropped this participant from the sample.

The second task of the experiment involved participants completing a set of probability tests (which was incentivized with each correct answer given = SG\$0.20), as well as a set of standard control questionnaires. At the end of the experiment, all participants were debriefed on the nature of the experiment either immediately (Thailand) or later via email (Singapore).

3.2. The first-round buyers

Of the 378 participants, 55 (or 14.5%) were first-round prediction buyers, i.e., the prediction was bought by individuals before any sequence of correct or incorrect predictions was

⁷ The method was first seen on UK television in 2008, demonstrated by the British magician Derren Brown. In his show called *The System*, Brown used the technique to illustrate how he was able to predict, for one particular person, six consecutive wins at a horserace.

observed. Of these 55, 41 (74.55%) were Singaporeans and 14 (25.45%) were Thais. For these subjects, it may be that (a) they did not completely grasp all the facts that we had stated earlier in the experiment about the fairness of the coins, or (b) they confidently believed that coins are predictable, i.e., those with a positive prior on predictions being useful from the start. Since we cannot distinguish (a) from (b), our main focus will be on the 323 non-first-round buyers. We briefly return in section 6.1 to analyze the first-round buyers.

We report in Appendix A the descriptive statistics of the buying behaviors of first-round buyers – all and by nationality – and all non-first-round buyers.

3.3. Compliers *versus* defiers among prediction buyers (excluding first-round buyers)

Conditioning on buying, most people who bought predictions went on to bet according to that prediction. In round $j = 2$, 16 out of 19 (or 84%) went on to place the same bet as the prediction. For round $j = 3$, 16 out of 18 (or 89%) bet the same as the prediction, whilst 17 out of 22 (or 77%) bet according to the prediction in round $j = 4$. And finally, for round $j = 5$, 21 out of 29 (or 72%) bet the same as the prediction. What these figures imply is that although not everyone who bought predictions went on to bet according to that prediction, a large proportion of people did.

4. Empirical Strategy

4.1. The Stylized Model

To fix ideas for our empirical models, we initially assume that each subject in the experiment observes, just before placing a bet on the outcome of a coin flip in round j , two sequences of signals whose probability distributions depend on some underlying states (Rabin & Vayanos, 2010). The first signal s_t is the outcome of a *fair* coin flip in round j , which takes the value of 1 if the outcome is ‘head’ and 0 if it is ‘tail’. The second signal a_t represents the success rate

of the prediction given by the affixed envelope in period t , which takes the value of 1 if the prediction matched the outcome of the coin flip in the corresponding t and 0 otherwise.⁸ To the subjects, the signal s_t in rounds $j = 1, 2, \dots, 5$ is

$$s_j = \mu + \varepsilon_j, \quad (1)$$

and the signal a_j in rounds $j = 1, 2, \dots$, is

$$a_j = \varphi + v_j, \quad (2)$$

where μ is the long-run mean of the i.i.d. signals, which is fixed at 0.5 for a series of fair coin flips. To the individual, the envelopes' long-run predictability of the outcomes of fair coin flips, φ , is also *a priori* fixed at 0.5; and ε_j and v_j are i.i.d. normal shocks with means zero and variances $\sigma_\varepsilon^2, \sigma_v^2 > 0$. We can also interpret the shock v_j as the envelope' luck at predicting s_j in round j . More importantly, s_j is assumed to be determined independently from, i.e., an outcome of a coin flip is determined by the actual flip rather than by a prediction given by an envelope in the corresponding round.

According to Rabin and Vayanos (2010), behaviors of the gambler's fallacy nature arise when subjects have a mistaken belief that the sequences $\{\varepsilon_j, v_j\}_{j \geq 1}$ are not i.i.d. but exhibit systematic reversal.⁹ What this implies is that: (i) following a streak of s (i.e. heads) up to $j - 1$, subjects will develop an erroneous belief that outcome in round j is more likely to be $1 - s$ (i.e. tail), and (ii) following a streak of a (correct predictions) up to $j - 1$, subjects will develop an erroneous belief that the envelope's prediction in j is likely to be $1 - a$ (or incorrect).

⁸ It should be noted that the second set of signals is viewed privately and allowed to vary across subjects in our experiment.

⁹ The gambler's fallacy is also explained in Rabin (2002) albeit in a different modeling setup. In Rabin (2002), the gambler's fallacy arises due to people's false belief that the signals are drawn from an urn with no replacement. The absence of replacement creates negative autocorrelation between successive draws making people to believe that it is less likely that the same signal is drawn again in the next period. For example, in the case of two flips of a fair coin, the probability of head appearing again in the second flip after appearing in the first flip would be $(1/2) \cdot (1/3)$ instead of $(1/2) \cdot (1/2)$.

There are no physical costs that could be incurred on the subject if he/she decides to switch from betting s (head) in $j - 1$ to betting $1 - s$ (tail), and vice versa. There is, however, a fixed cost that each subject has to pay if he/she wishes to observe the prediction of the coin flip before and not after having to place a bet in the corresponding round. Assuming for simplicity that subjects face no budget constraints in their betting and purchasing decisions, we can write our first set of testable hypotheses on behaviors influenced by the gambler's fallacy as follows:

H1: Subjects' betting behaviors will generally be influenced by the gambler's fallacy after having observed a streak of s (heads) or $1 - s$ (tails) up to $j - 1$;

H2: Following a streak of a (correct predictions) up to $j - 1$, subjects will pay to see the prediction in the next period and then bet the opposite; and

H3: Following a streak of $1 - a$ (incorrect predictions) up to $j - 1$, subjects will pay to see the prediction in the next period and then bet according to the prediction provided by the envelope.

As previously stated, the envelopes' long-run predictability of a series of fair coin flips, φ , is *a priori* known and fixed at 0.5. This is because it should be transparent to anyone able to carry out simple probabilistic calculations that there is no systematic method to accurately predict pure i.i.d. outcomes, in our case coin flips. However, we will allow for subjects' perceptions about the nature of φ to be influenced by a streak of a (correct predictions) or $1 - a$ (incorrect predictions) up to $j - 1$. More specifically, we allow subjects' perceptions about the envelopes' long-run predictability to change from being fixed at 0.5 to one that evolves according to the auto-regressive process:

$$\varphi_j = 0.5 + \rho(\varphi_{j-1} - 0.5) + \eta_j, \quad (3)$$

where $\rho \in [0,1)$ is the reversion rate to the long-run average of 0.5, and η_j is an i.i.d. normal shock with mean zero and variance σ_η^2 and that is independent of v_j . If it is possible for

subjects to develop a belief in serially correlated variation in φ (i.e. $\rho > 0$), then, in principle, a belief in the hot hand (or cold hand, in the case of a sequence of incorrect predictions) can develop – overtaking the gambler’s fallacy – after observing a streak of a (correct predictions) or $1 - a$ (incorrect predictions) up to $j - 1$. This produces our second set of testable hypotheses on behaviors influenced by the hot/cold hand fallacy as follows:

H4: Following a streak of a (correct predictions), subjects will pay to see the prediction in the next period and then bet according to the prediction provided by the envelope; and

H5: Following a streak of $1 - a$ (incorrect predictions) up to $j - 1$, subjects will pay to see the prediction in the next period and then bet the opposite.

Finally, the hypotheses **H2-H4** also imply that:

H6: Subjects who pay for a prediction in j will treat the paid-for information seriously by placing a higher bet in the corresponding round.

4.2. Testing participant’s belief in gambler’s fallacy (H1)

To test whether participants’ betting behaviors are influenced by the gambler’s fallacy, we estimate – for round j – the following equation:

$$\begin{aligned}
 BH_{it} = & \alpha_j + \beta_{1j}SH_{j-1} + \beta_{2j}ST_{j-1} + \delta_j P_{ij} + \gamma_{1j}(SH_{j-1} \times P_{ij}) + \gamma_{2j}(ST_{j-1} \times P_{ij}) \\
 (4) & \\
 & + X_i' \theta + u_{ij},
 \end{aligned}$$

where i indexes individuals and j indexes experimental round; BH_{ij} is an indicator variable that takes the value of 1 if the individual bets “head” in round j , and 0 if the individual bets “tail”; SH_{j-1} is a dummy representing a streak of heads up to the round $j - 1$; ST_{j-1} is a dummy representing a streak of tails up to the previous round, i.e. round $j - 1$; P_{ij} is an indicator variable representing whether the subject i paid 10 tokens to see the prediction in

round j ; X'_i is a vector of control variables; and u_{ij} is the error term. We will also estimate equation (4) separately for, respectively, the sub-sample of observations with head and tail as the prediction outcome. More formally, we test whether:

- (i) Subjects are more likely to bet tail in round j (i.e. $\beta_{1j} < 0$) following a streak of heads up to round $j - 1$. Similarly, if a streak of tails is observed until round $j - 1$, then the individual is more likely to bet head in round j (i.e. $\beta_{2j} > 0$); and
- (ii) Depending on the prediction in round j , those who paid for it will place a bet that corresponds to the hot hand of the expert's prediction rather than be influenced by the gambler's fallacy. What this implies is that subjects will be more likely to bet "head" following a streak of heads up to $j - 1$ if they had bought the prediction and the prediction told them to bet "head", i.e., going with the streak. Similarly, subjects will be more likely to bet "tail" following a streak of tails up to $j - 1$ if they had bought the prediction and the prediction told them to bet "tail".

4.3. Testing the effects of past predictions on buying behaviors (H2-H5)

In order to test whether a streak of past predictions matters to the subject's purchasing decision in the current round of coin flip, we estimate the following buying equation:

$$P_{ij} = \alpha_j + \phi_{1j}SC_{ij-1} + \phi_{2j}SI_{ij-1} + \lambda_j P_{ij-1} + X'_{ij}\vartheta + \varepsilon_{ij}, \quad (5)$$

where SC_{ij-1} is a set of dummy variables representing a streak of correct predictions up to round $j - 1$; SI_{ij-1} is a set of dummy variables representing a streak of incorrect predictions up to round $j - 1$; P_{ij-1} is the decision to buy in round $j - 1$; and ε_{ij} is the error term. Our key parameters of interest here are ϕ_{1j} and ϕ_{2j} , which represent the average effects of observing successful and failed streaks of past predictions on the subject's buying decision in round j , holding buying decision in the previous round constant. Note that, for $j=2$, the estimated effect of obtaining a correct prediction in round 1 is the effect relative to obtaining

an incorrect prediction in round 1. For $j=\{3,4,5\}$, the estimated effect of obtaining all-correct, or all-incorrect, predictions prior to round j is thus the effect relative to obtaining some correct and some incorrect signals, which is typically the outcomes that subjects *a priori* expected to see for predictions made on truly random events. Here, we test whether subjects are more likely to buy the prediction after observing a streak of predictions up to round $j-1$ (i.e. $\phi_{1j}; \phi_{2j} > 0$).

4.4. Testing the gambler's fallacy *versus* the hot hand fallacy in behaviors towards the paid-for predictions (H2-H5)

In order to test whether subjects' behaviors towards the paid-for predictions are influenced more by the gambler's fallacy (H2-H3) or the hot hand fallacy (H4-H5), we estimate the following equation:

$$C_{ij} = \pi_j + \delta_{1j}SC_{ij-1} + \delta_{2j}SI_{ij-1} + \tau_j P_{ij} + \Phi_{1j}(SC_{ij-1} \times P_{ij}) + \Phi_{2j}(SI_{ij-1} \times P_{ij}) \quad (6)$$

$$+ X_i' \theta + \zeta_{ij},$$

where C_{ij} is a dummy variable that takes the value of 1 if the individual bets the same as the predicted envelope in round j , and 0 otherwise; and ζ_{ij} is the error term. Here, we test whether subjects bet according to the prediction or against the prediction following a streak of (i) successful and (ii) failed predictions up to round $j-1$. Provided that the gambler's fallacy dominates, then we should expect to see $\Phi_{1j} < 0$ and $\Phi_{2j} > 0$. By contrast, we should expect to see $\Phi_{1j} > 0$ and $\Phi_{2j} < 0$ in cases where the hot/cold hand fallacy dominates.

4.5. Testing the effect of buying predictions on betting behaviors (H6)

The final equation tests whether the subjects who bought the predictions ended up treating them seriously. Here, a natural question is whether the amount of endowment used in each

bet is larger on average among buyers than non-buyers. We can test this by estimating the following bet amount equation:

$$EB_{ij} = \mu_j + \pi_j P_{ij} + \varpi_j EB_{ij-1} + X_i' \omega + v_{ij}, \quad (7)$$

where EB_{ij} denotes the log of the endowment amount used to bet by individual i in round j ; and v_{ij} is the error term. The null hypothesis here is that buyers will tend to treat the paid-for predictions seriously and consequently place larger bets than non-buyers (i.e. $\pi_j > 0$).

All four equations are estimated using OLS (linear probability model) with robust standard errors.¹⁰

4.6. Control variables

Depending on the outcome variable, the control variables, X_i' , may consist of gender; country (Thailand *versus* Singapore); endowment in round j ; a dummy variable representing whether subject bought the prediction in the previous round; the proportion of correct answers in statistical test; a streak of coin coming up heads or tails up to round $j-1$; an indicator variable representing whether the subject had made a wrong bet in the previous round and its interaction with the decision to buy in $j-1$ variable; a dummy variable for whether the subject followed a fixed betting strategy by betting either heads or tails in every round; and a dummy variable for whether the subject acts according to the prediction or deviates from it in the previous round and its interaction with the decision to buy in $j-1$ variable.

5. Results

5.1. Are subjects' betting behaviors influenced by the gambler's fallacy?

¹⁰ Although qualitatively the same results can also be obtained using a non-linear model (marginal Probit); see Appendix C in the online annex.

Table 1 tests whether non-first-round buyers are subject to the gambler’s fallacy in their betting behaviors. Here, the dependent variable is an indicator variable that takes a value of 1 if the participant chooses to bet “head” in round j , and 0 if he chooses to bet “tail” in round j . Since the same sequence of coin flips is observed by all participants in a given session and there are 12 sessions in total, we have 118 observations of two heads after the first two rounds and 152 observations of two tails after the first two rounds (i.e., $j=1$ and $j=2$). The number of consecutive heads and tails becomes even smaller at the start of round 4; there are only 23 observations of three heads after the first three rounds and 48 observations of three tails after the first three rounds (i.e., $j=1$, $j=2$, and $j=3$). Our focus will be on the subject’s propensity to bet “head” following streaks of heads and tails in rounds $j=3$ and $j=4$.

Looking at columns 1 (All; Round $j=3$) and 4 (All; Round $j=4$) of Table 1, we can see – based on the specification in Eq. (4) – that participants generally behave according to the gambler’s fallacy in that:

- (i) Participants are approximately 28 to 31 percentage points *less* likely to bet “head” in round j if a streak of *heads* has been observed up to round $j-1$.
- (ii) Participants are approximately 17 percentage points *more* likely to bet “head” in round j if a streak of *tails* has been observed up to round $j-1$.

These estimated effects are also statistically robust at conventional confidence levels and are robust to controlling for subject’s gender and nationality.

However, a more interesting pattern emerges when the “bought prediction in j ” dummy and its interaction with the streak variables are introduced as independent variables in the “betting ‘head’ in round j ” regression equations. Conditioning on the prediction in round j being “head”, there is an offsetting effect to the gambler’s fallacy for participants who bought the prediction and recently observed two consecutive heads up to $j-1$: The net effects on betting “head” for these individuals are $(0.015 + 0.582 = 0.597; S.E. = 0.083)$ in round $j=3$

and $(0.067 + 0.536 = 0.602; S.E. = 0.209)$ in round $j=4$. We can also see from round $j=4$ that, conditioning on the prediction in round j being “tail”, those who “bought the prediction in j ” and observed “a streak of tails up to $j-1$ ” are statistically less likely to bet “head” in round j if the prediction for round j is “tail”; here, the net effect is $(-0.387 - 0.298 = -0.685; S.E. = 0.111)$, which means that we cannot reject the null of zero at conventional confidence levels.

The results of Table 1 imply that participants who paid for the prediction in round j – presumably because they had (randomly) received accurate predictions in the previous rounds – will likely use them to guide how they place their bets in the same round, thus providing us with partial evidence that betting behaviors of prediction buyers are influenced more by the hot hand of the envelopes rather than by the gambler’s fallacy of previous coin outcomes.

5.2. Would people pay for transparently useless information?

Do people who randomly receive correct (incorrect) predictions then perceive in the hot hand (cold hand) of the predicting envelopes and thus start paying for such transparently useless information? If so, how long is it before they start buying? We found the answers to be: yes, and not long.

Figure 2 first illustrates this using the raw data of the pooled Thai and Singaporean sample. Conditioning on non-first-round buyers, the ratio of people who paid for round j ’s prediction increases from around 9% in round $j=2$ to 43% in round $j=5$, providing that they had just previously observed a streak of correct predictions up to round $j-1$. The rise is also monotonic and statistically well-determined throughout. The ratios of buyers to non-buyers are much lower (~1% to 4%) among those who observed a mix of correct and incorrect predictions. It is only after observing four consecutive incorrect predictions that we observe a noticeable increase in the proportion of participants paying for the prediction in the final round, with the ratio of buyers to non-buyers at 17%. A similar picture is observed when we

split the sample by countries (see Figures 3A and 3B), although the aforementioned patterns are noticeably more robust for the Thai sample than for the Singaporean sample.

Table 2 presents the results more formally through the use of multivariate regressions. Based on Eq. (5), the control variables here include gender, nationality, proportion of correct answers in statistical test, endowment in round j , a streak of coin coming up heads or tails up to round $j-1$, a dummy representing whether the subject had made a wrong bet in the previous round and its interaction with the buying decision in round $j-1$, a dummy for whether the subject bet heads or tails in every round, and an indicator variable representing whether the subject act according to the prediction in round $j-1$ and its interaction with the buying decision in round $j-1$.

Here, we can see that the general patterns in our earlier figures are preserved here in our linear probability estimates, with or without additional control variables. For example, in the full specification, the estimated probability of buying round 2's prediction for those who previously received a correct prediction in the first round are 5.5 percentage points higher than for those who previously received an incorrect prediction. The positive gradient in the buying decision as we move through rounds is also noticeable and individually statistically well-determined; holding the decision to buy in round $j-1$ constant, probabilities of buying are 15 percentage points in round $j=3$, 21 percentage points in round $j=4$, and 27 percentage points in round $j=5$. By contrast, the coefficient *All predictions up to $j-1$ had been incorrect* is statistically significantly different from zero only in the final round of coin flip: For individuals whose previous predictions had been incorrect four times in a row the probability of buying is 16 percentage points higher than those whose previous predictions were a mixture of successes and failures.

The empirical evidence of a monotonic rise in the ratio of buyers to non-buyers for those whose previous predictions up to $j-1$ had been correct, and a sharp rise in the same

ratio in the final round for those whose previous predictions up to $j-1$ had been incorrect, suggest that a significant number of subjects had come to believe in the predictability of the envelopes after having observed either a streak of successful and failed predictions being made live in the lab. To traditional economists, this finding is probably intuitively difficult to explain. Economic models such as the dynamic-inference model developed by Rabin and Vayanos (2010) predict that, when people are confident that no human skill is involved in making accurate predictions of truly i.i.d. events, a prediction of a future coin flip is worth as much as a blank sheet of paper in an envelope, and people will not pay for such transparently useless advice. Our experiment however provides statistically convincing evidence that rejects the idea that when individuals are absolutely rational (recall that these are non-buyers in the first round) they will remain so throughout the repeated random trial.

Other results in Table 2 are also interesting. There is some evidence in the final round that the buying decision is also persistent in itself: The coefficient *Bought prediction in round $j-1$* is positive and statistically significant at conventional confidence levels in round $j=5$. There is no difference in the buying behaviors between men and women, Singaporean and Thai subjects, or endowment levels. The coefficients *Proportion of correct answers in statistical test* have the anticipated negative sign, although they are not statistically significantly different from zero in the first three rounds of coin flip. The relationship between observing a streak of heads or tails up to $j-1$ and the propensity to buy a prediction in round j is not statistically robust in either round $j=3$ or round $j=4$. The coefficients on *Fixed betting strategy* are negative albeit statistically significant only in round $j=4$. There is also no evidence to suggest that people who bought the prediction in round $j-1$ and acted accordingly are more or less likely to buy the prediction again in round j . Finally, people who bought predictions that later turned out to be “bogus” – i.e., the interaction term between

Bought prediction in $j-1$ and Made wrong bet in $j-1$ – are significantly less likely to buy in future rounds.

One question of interest is whether an individual's tendency to buy is different for different groups of people. Though not reported here, we found no statistically significant slope differences by gender or nationality or in the test score of buying behavior among subjects who randomly received correct as well as incorrect predictions. This is the case even when the initial endowment in Singapore is significantly larger than that in Thailand and, consequently, prediction purchases are less costly for Singaporeans than for Thai participants (3.33% and 10% of initial endowment, respectively).¹¹ This implies that the perceived hot hand in the envelopes predicting the future is not statistically more pronounced for males compared to females, or for the Singaporean sample compared to the Thai sample, or for those who scored better on average in statistics and probabilities. In short, there is no statistical evidence to support the notion that some people are systematically more (or less) susceptible to such an irrational behavior.

We next carry out a test on people's behaviors towards the paid-for predictions, to see whether they are influenced more by the gambler's fallacy or by the hot/cold hand fallacy of the predicting envelopes. Using the same control variables in Eq. (6) as in the previous table, we are able to show in Table 3 that prediction buyers tend to bet according to – rather than against – the paid-for prediction after observing a streak of correct predictions up to round $j - 1$. This is more consistent with the hot hand fallacy as prediction buyers who had previously observed a streak of correct predictions up to the previous round are 35.4 percent and 58.1 percent more likely than the average non-buyers to bet according to the prediction in rounds $j = 2$ and $j = 3$, respectively. The estimated figure continues to rise to around 70 percent in round $j = 4$, before dropping to 42.7 percent in the final round.

¹¹ Qualitatively similar results hold when we re-estimate the buying equation by participants' nationality; see the results reported in Appendix D in the online annex.

The evidence in support of the gambler's fallacy behaviors following a streak of incorrect predictions is, however, mixed. For example, we can see that prediction buyers who had previously observed a streak of incorrect predictions up to the previous round are 53 percent more likely to bet the same as the paid-for prediction in round $j = 3$, which is more consistent with the gambler's fallacy previously stated in H3. However, a further analysis also reveals that, following a streak of incorrect predictions, subjects are approximately 43 percent *less likely* to bet according to the paid-for prediction in round $j = 5$, which is more consistent with the hot hand effect previously stated in H5.

More generally, what these numbers are essentially telling us is that, for the most part, the betting behaviors of prediction buyers are influenced more strongly by the hot hand and not by the gambler's fallacy of the predicting envelopes.

What about the betting and subsequent buying behaviors of those with a (potentially) positive prior, i.e., the first-round buyers? Are they significantly different from the non-first-round buyers? Table 4, which re-estimates the specification of Table 1 on the much smaller sub-sample of first-time buyers ($N=55$), suggests that both first-round buyers and non-buyers exhibit similar propensities to the gambler's fallacy in their betting behaviors: The coefficient *A streak of coins coming up heads up to $j-1$* in the betting "head" equation is -0.342 [$S.E.=0.144$] in round $j=3$, which is similar to the -0.310 [$S.E.=0.061$] observed earlier for the non-first-round buyers.

On the other hand, we can see from the buying equation of Table 5 that the estimated hot hand effect in round $j=3$ is almost twice the size of the same coefficient obtained without additional control variables in Table 2: The coefficients *All predictions up to $j-1$ had been correct* are 0.295 [$S.E.=0.142$] in round $j=3$ for the first-round buyers and 0.153 [$S.E.=0.043$] in the same round for the non-first-round buyers. This suggests evidence that the first-round buyers converge more quickly toward the belief that the envelopes have some predictive

power, compared to the non-first-round buyers, i.e., those who have either a zero or a negative prior at the beginning of the experiment. This supports the earlier decision to exclude this smaller sample from our initial analysis.¹²

There are several potential objections to our results. The first is that participants only bought the predictions to please the experimenters. However, if this was the case, then we would have expected to see a similar buying pattern between the treated and those in the control group, but we did not. The second is that the “presence” of envelopes themselves may have added an unknown element to how each individual calculate his or her expected utility from placing such bets. For instance, one could imagine each participant thinking at the start of the experiment, *“I know that predictions contained within these envelopes are useless. But if they are really useless, then why would they be here in the first place?”* It is therefore easy to rationalize after seeing a streak of correct predictions being made by these envelopes that the experimenters are “up to something” – perhaps through some type of magic trick – and it would thus be better for the participants to buy the predictions even if they do not really believe in their predictability.¹³

While we cannot completely deny such beliefs in our volunteer participants, we can nevertheless shed further light on this issue. To do this, we asked our participants in Singapore to state their reasons in the post-experimental questionnaire for buying at least one of the predictions (if they did at some point) and for not buying the predictions (if they did not). The results provide interesting insights into people’s purchasing decisions.

¹² However, for the sake of completeness, similar conclusions are obtained from the pooled data analysis (combined first-round and non-first-round buyers). See Powdthavee and Riyanto (2012) for the pooled data estimates.

¹³ Similarly, one could also imagine investors in the stock market thinking along the same lines about financial experts: *“If stock prices really follow random walk, then why do financial experts exist? And not only do they exist, they also get paid incredible sums of money to comment on something that is i.i.d. in nature.”* And then when we see a streak of successful performances by these financial experts, we readily use it as justification for their existence and for the large paychecks they typically receive.

Participants whose predictions were of mixed successes gave reasons for *not buying* that are mostly rational, e.g., “*Coin tosses are random,*” “*For this experiment, the outcomes are not predictable,*” “*I believe in my own choices,*” “*Predictions are as good as my own guess,*” “*Predictions are random and a waste of money.*” For a larger proportion of participants whose predictions had been correct *four* times in a row, their reasons for *buying* are mainly consistent with the belief in the hot hand of an invisible agent, rather than explicitly state that the experimenters were up to some trickery, e.g., “*Predictions had been more successful than placing own bets,*” “*The past predictions were correct,*” “*Based on the accuracy of past predictions.*” Participants whose predictions had been correct up to round $j-1$ but not in j gave reasons for *buying* that are a mixture of *ex post* rationalizations, excuses, and sometimes regrets, e.g., “*I could not think of an outcome to bet on my own,*” “*Out of curiosity,*” “*I am stupid to buy it.*”

5.3. Do buyers treat paid-for predictions seriously?

One way of inferring whether buyers treated the paid-for predictions seriously is to examine the amount of a bet placed by buyers compared to non-buyers. The hypothesis here is that people who bought the predictions will feel – through the process of rationalization – more confident about future outcomes and subsequently place higher bets than non-buyers, i.e., those who still believe that coin flips are i.i.d. and therefore systematically unpredictable. We formally test this hypothesis in Table 6 by running, for each round j , a regression equation in which the dependent variable is the log of the endowment amount used to bet in round j , i.e., Eq. (7). The control variables here are the same as in Table 2.

On average, the bet amount placed by buyers is between 27 and 51 percentage points higher than non-buyers in the final three rounds. The estimated effects are statistically significant at conventional levels in rounds $j=3$ and 5, and are robust to controlling for the

previous bet amount, the current endowment level, gender and nationality, among other things. These results provide strong evidence – of both statistical and economic significance – that buyers place a significant level of trust on the “hot” envelopes rather than buy them to satisfy their own curiosity or simply for fun.

6. Discussions and concluding remarks

An experimental game in which people guess and bet on the outcomes of “fair” coin tosses is ostensibly simple. Yet behind its simplicity lies its unique strength. Given that the coin is fair (in that it had been proven fair by various explicit processes), it should be universally irrefutable that the exact sequence of future coin tosses is systematically unpredictable.¹⁴ Participants’ predictions should therefore be influenced by the gambler’s fallacy rather than by the random accuracy of the envelope’s predictions. Such a setting provides us with a perfect experimental setting in which we could test the boundaries of people’s beliefs in the folly of coin-flip predictions.

However, our findings take us and perhaps many others¹⁵ by surprise. When randomized predictions are introduced into the game as a potential information good, it takes only a few past realizations of correct predictions for individuals to start forming the belief that transparently i.i.d. outcomes are systematically predictable and that the prefixed envelopes contain information worth paying for. The influence of such “bogus” predictions is

¹⁴ This is unlike the belief that past returns do not predict future returns in the stock market, the truth of which many people spend years debating and researching. Prior beliefs about the predictability of randomly selected coin flips should be more absolute and universal.

¹⁵ This comment is made in part to reflect how the earlier draft of this paper has been surprisingly received in the media and by the general public. Since its online publication as an IZA discussion paper in May 2012 (Powdthavee & Riyanto, 2012), our results have been discussed at great length – despite having had no press release – on the Freakonomics blog www.freakonomics.com/blog/ (“Paying for ‘transparently useless advice’,” 6 June 2012), *New Statesman* (“If you’ve got lucky, it’s easy to convince people you’re a sage,” 6 June 2012), *The Economist* (“Buttonwood: Not so expert,” 9 June 2012), *The Financial Times* (“Heads or tails? Just don’t bet on it,” 15 June 2012), *The Wall Street Journal* (“People will pay for ‘transparently useless advice’ about chance events,” 18 June 2012), among others. Even we, the experimenters, nearly gave up with the idea of this project before we had even started simply because we had difficulty believing anybody would be willing to pay for predictions of obviously fair coin flips.

sufficiently large to offset the gambler's fallacy and induce buyers to place higher bets on average than non-buyers.

Indeed, we believe that our laboratory experiment provides results that are essentially predictable by the dynamic-inference model by Rabin and Vayanos (2010). Our contribution, however, is that these results are obtained even without the key assumption – i.e., that individuals are *a priori* uncertain about the underlying process generating the data – needing to hold. In other words, our experiment establishes a new lower bound of how people's beliefs in transparently useless predictions can be formed, i.e., participants are not entirely protected from “irrational learning” even when, at the beginning of the experiment, these individuals had started with a very strong prior about the ability of the predicting agent. Our results call for the need to look outside the economic discipline for an alternative explanation of such puzzling findings.

One potential explanation is in the work of the psychologist Justin Barrett. In his Hypersensitive Agency Detection Device theory, humans are hardwired to detect patterns in otherwise unrelated events, details that defy straightforward explanations, or consequences that seem out of proportion to the alleged cause. Such sensitiveness to detecting agents even when none exists has several evolutionary advantages. For example, spotting and understanding other agents could have been key to survival for early man and thus continuing to reproduce: It is far better to avoid several imaginary predators than be eaten by a real one. Since there is such a high cost of failing to detect agents and a low cost of wrongly detecting them, evolution will select an inheritable tendency to over-detect agents, even when we do not see them, as a survival strategy (Barrett, 2004; Grey & Wegner, 2010). The way we set up our experiment, which randomly allowed some participants to experience implausible streaks of accurate predictions, may have helped to trigger the hypersensitive agency

detection device for these individuals, thus leading them to believe in the hot hand of an invisible agent when there was none.

Our results also open up a new discussion that perhaps people buy expert predictions (which, remember, are transparently useless in our case) not for their predictability but rather for other psychological reasons.¹⁶ For example, our participants may have bought the predictions to:

- **Delegate decision-making.** Participants know that their choice is no better than a random envelope's choice. Nevertheless, if the decision is wrong, at least they can blame the envelope and not themselves.
- **Avoid regret.** Participants know that they will learn the information in the envelope at the end of each coin flip for free. However, they may not want to regret not having that information if it turns out to be correct. Since regret is powerful, they may instead want to pay the small fee to avoid it.
- **Feel in control of the situation.** Participants may buy the predictions simply to feel in control of a situation in which they have no control over the outcomes. Like regrets, the ability to feel in control of an uncontrollable situation is a powerful emotion, and participants may be willing to pay a small fee to be in the possession of it (see, e.g., Langer, 1975).
- **Psychic hedge.** It may be the case that participants formally go with the predictions in the envelope even when they do not formally believe that it has any predictive power.

This way, regardless of what happens, they win (at least psychically).

Equally interesting are the potential implications of our findings across social-science disciplines. What we have learned, other than that people are not completely immune to irrational learning, is that it is unimaginably easy and costs almost nothing for “experts” to

¹⁶ We are especially grateful to Dan Houser for giving us advice on likely alternative psychological explanations of our results.

manufacture fallacious beliefs that even truly i.i.d. events are predictable. This is primarily because, according to Philip Tetlock (2006), experts are not punished sufficiently when they are wrong, nor will they admit to being wrong when they are. On the contrary, the benefit of predicting unpredictable events correctly, even if rarely, significantly outweighs the cost of not getting them right. Since experts are usually paid for their services, paid better when they give “good” advice, and are not severely punished when incorrect ones are made, we are essentially rewarding bad judgments over good ones at no cost to the overall demand for expert advice.

Hence, with the existing system of expertise, financial firms could in principle provide their customers with various mutual funds containing randomly chosen stocks, and then build their advertisements around those that outperform the market by chance. Since people also buy transparently useless information for many psychological reasons other than for their predictability, people’s demands for financial advice are likely to be inelastic with respect to the prices that financial firms typically charge their clients for their services.¹⁷ This is the case even when it may be clear to everyone involved that market evaluations revert to the mean over the long term and that such advice is not needed very often in reality. Moreover, economists and political forecasters can hand out random predictions about future economic and political crises, and would be tantalizingly rewarded when they are right and yet would not be held accountable for misprediction when they are wrong.

Our experiment highlights these flaws in the current system of expertise and argues that it may not be sufficient to leave it to individuals to judge for themselves whether an expensive expertise is worth paying for. This is the case even in situations where it should be transparent to everyone that outcomes are dominated by pure randomness. Our results also underscore the problem that years of using statistics cannot offset the erroneous intuitions we

¹⁷ A good reference for this comes in the review by Inderst and Ottaviani (2012). According to a recent survey of the market, the majority of US retail investors (approximately 73%) consult a financial adviser before purchasing shares.

sometimes have about the hot hand of an expert after streaks of accurate predictions are observed.

This raises an important question: If teaching people about statistics do not help to minimize these mistakes, how can we create a mechanism that will? One potential solution to this problem may come in the form of prediction markets (see, e.g., Wolfer & Zitzewitz, 2006; Arrow et al., 2008). Prediction markets, or information markets, allow participants to generate various contracts and price them on the basis of an aggregation of collective predictions of outcomes of some future events made by market participants and then trade them in the market at a price. The wisdom of the crowd underlying prediction markets offers a better alternative to the prediction provided by single (or few) expert(s). A typical example of a contract in the prediction market is a forecast that candidate X will win the next presidential election, which could be traded in the market at, for example, US\$1. If the market price of such a contract is currently 40 cents, an interpretation is that the market “believes” candidate X has a 40% chance of winning the election. Provided that the market is efficient, then the price of the contract perfectly aggregates dispersed information about the probability of candidate X being elected.

The promise of prediction markets is that, owing to the law of large numbers, the market prices will often produce aggregated forecasts of event outcomes that have lower prediction errors than predictions generated by a single or a few “experts.” Participants in the prediction markets also have a clear financial incentive for truthful revelation of true beliefs as losses will be made if predictions turn out to be wrong. More importantly, prices in the prediction market can also be used as an aggregated indicator of whether the event in question is random or predictable – i.e., aggregate prices should be lower for events that are significantly more random in nature – which may help to improve investment decisions for individuals looking to invest.

We began this paper by noting a divide between the economics and psychology literature. Our experimental results seem to provide novel evidence in favor of psychological explanations for the apparent demands for useless predictions. It appears that economists may need to readjust their prior about where the lower bound – between rational and irrational beliefs – actually lies. Future research may need to return to this to construct a general quantitative model of such “irrational learning” that can be applied across a wide variety of settings, and to determine cost-effective ways (other than the promise of prediction markets) to help minimize these mistakes in the market.

References

- Arrow, Kenneth J., and 21 other colleagues. 2008. The promise of prediction markets. *Science*, 877-878.
- Asparouhova, Elena, Michael Hertzel and Michael Lemmon. 2009. Inference from streaks in random outcomes: experimental evidence on beliefs in regime shifting and the law of small numbers. *Management Science*, 55(11), 1766-1782.
- Ayton, Peter and Ilan Fischer. 2004. The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? *Memory and Cognition*, 32(8), 1369-1378.
- Barrett, Justin L. 2004. *Why would anyone believe in God?* Walnut Creek, CA: AltaMira Press.
- Burns, Bruce D. and Bryan Corpus. 2004. Randomness and inductions from streaks: "gambler's fallacy" versus "hot hand". *Psychonomic Bulletin and Review*, 11(1), 179-184.
- Cahart, Mark. 1997. On persistence in mutual fund performance. *Journal of Finance*, 52, 57-82.
- Chater, Nick, Steffen Huck, and Roman Inderst. 2010. Consumer Decision-Making in Retail Investment Services: A Behavioural Economics Perspective. Report to the European Commission/SANCO.
- Clotfelter, Charles and Phil Cook. 1991. Lotteries in the real world. *Journal of Risk and Uncertainty*, 4, 227-232.
- Clotfelter, Charles and Phil Cook. 1993. The gambler's fallacy in lottery play. *Management Science*, 39,
- Croson, Rachel and James Sundali. 2005. The gambler's fallacy and the hot hand: empirical data from casinos. *Journal of Risk and Uncertainty*, 30(3), 195-209.
- Fama, Eugene F. 1991. Efficient capital markets: II. *Journal of Finance*, 46, 1575-1617.

- Fischbacher, Urs. 2007. Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.
- Gilovich, Thomas, Robert Vallone and Amos Tversky. 1985. The hot hand in basketball: on the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- Gray, Kurt and Daniel M. Wegner. 2010. Blaming God for our pain: human suffering and the divine mind. *Personality and Social Psychology Review*, 14(1), 7-16.
- Guryan, Jonathan and Melissa S. Kearney. 2008. Gambling at lucky stores: empirical evidence from state lottery sales. *American Economic Review*, 98(1), 458-473.
- Huber, Jurgen, Michael Kirchler, and Thomas Stockl. 2010. The hot hand belief and the gambler's fallacy in investment decisions under risk. *Theory and Decision*, 68(4), 445-462.
- Hung, Angela A., Noreen Clancy, Jeff Dominitz, Eric Talley, Claude Berrebi, and Farrukh Suvankulov. 2008. Investor and Industry Perspectives on Investment Advisers and Broker-Dealers. RAND Institute for Civil Justice Technical Report.
- Inderst, Roman and Marco Ottaviani. 2012. Financial advice. *Journal of Economic Literature*, 50(2), 494-512.
- Jørgensen, Claus Bjørn, Sigrid Suetens and Jean-Robert Tyran. 2011. Predicting lotto numbers. CEPR Discussion Paper 8314.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Allen Lane: Penguin Books.
- Langer, Ellen. 1975. The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311-328.
- Powdthavee, Nattavudh, and Yohanes E. Riyanto. 2012. Why do people pay for useless information? Implications of gambler's and hot-hand fallacies in false expert setting. IZA Discussion Paper Series No. 6557.
- Rabin, Matthew. 2002. Inferences by believers in the law of small numbers. *Quarterly Journal of Economics*, 117(3), 775-816.

Rabin, Matthew; and Dimitri Vayanos. 2010. The gambler's and hot hand fallacies: theories and applications. *Review of Economic Studies*, 77(2), 730-778.

Teller, Dek. 1994. A test of the gambler's fallacy: evidence from Pari-Mutuel games. *Journal of Risk and Uncertainty*, 17, 151-166.

Tetlock, Philip. 2006. *Expert Political Judgment: How Good is It? How Can We Know?* Princeton University Press.

Tversky, Amos and Daniel Kahneman. 1971. Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105-110.

Wolfers, Justin and Eric Zitzewitz. 2006. Prediction markets in theory and practice. NBER Working Paper 12083.

Additional Supporting information may be found in the online version of this article:

Appendix B: Experimental instructions and screen shots.

Appendix C: Marginal effects probit model.

Appendix D: Buying regression equation by nationality, OLS regressions.

Figure 1: The coin tree

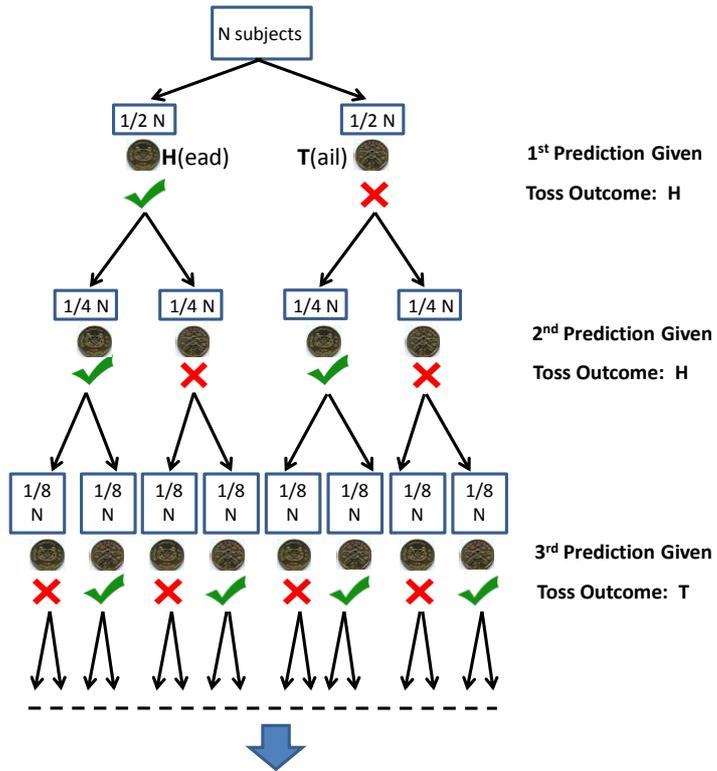


Table 1: Testing the gambler's fallacy in people's betting behaviors

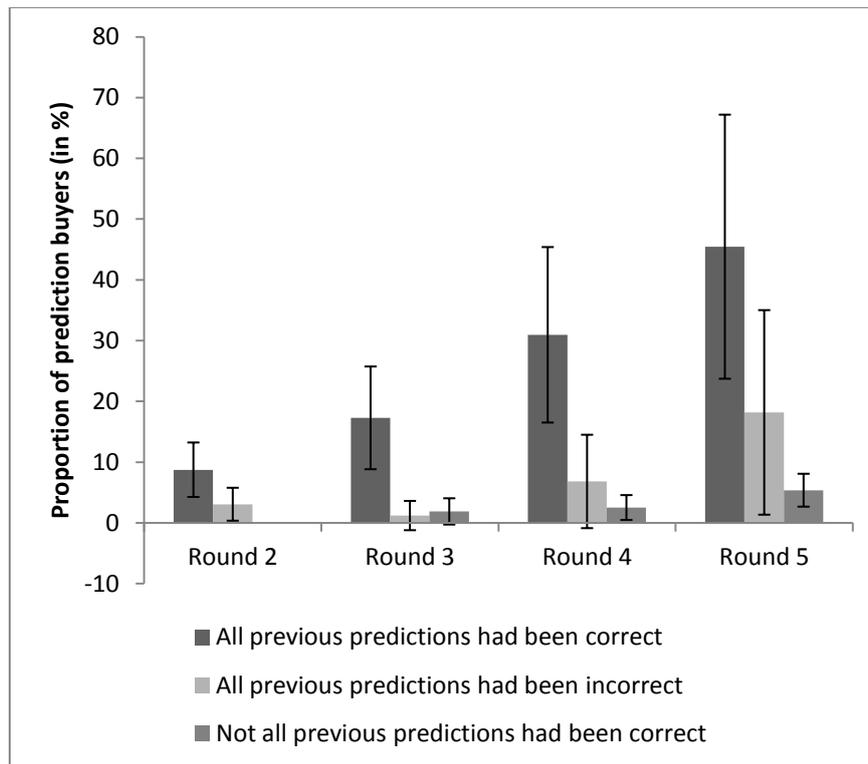
Dependent variable: Betting head in round j	Round $j = 3$			Round $j = 4$		
	All	Prediction = H	Prediction = T	All	Prediction = H	Prediction = T
A streak of coin coming up heads up to $j-1$	-0.310*** [0.0858]	-0.277** [0.122]	-0.375*** [0.129]	-0.284*** [0.100]	-0.317 [0.217]	-0.408*** [0.152]
A streak of coin coming up tails up to $j-1$	0.180*** [0.0608]	0.166* [0.0878]	0.174* [0.0911]	0.120* [0.0717]	0.203* [0.113]	0.150 [0.128]
Bought the prediction in round j		0.0151 [0.307]	-0.702*** [0.0795]		0.0671 [0.161]	-0.387*** [0.138]
A streak of heads \times bought prediction in j		0.582* [0.318]			0.536** [0.263]	0.0948 [0.202]
A streak of tails \times bought prediction in j		0.167 [0.312]	0.0942 [0.239]			-0.298* [0.178]
Male	0.0645 [0.0494]	0.0852 [0.0702]	0.00982 [0.0703]		0.192** [0.0770]	0.0355 [0.0792]
Singaporean	-0.0282 [0.0648]	-0.0118 [0.0898]	-0.0588 [0.0955]		0.0489 [0.0881]	0.147* [0.0885]
Constant	0.623*** [0.0559]	0.607*** [0.0812]	0.698*** [0.0820]	0.588*** [0.0282]	0.474*** [0.0726]	0.518*** [0.0714]
Implied net effects on betting head in round j						
<i>Streak of heads up to $j-1$ and bought prediction</i>		0.597*** [0.083]			0.602*** [0.209]	-0.292*** [0.143]
<i>Streak of tails up to $j-1$ and bought prediction</i>		0.182*** [0.054]	-0.607*** [0.224]			-0.685*** [0.111]

Observations	322	164	158	322	165	157
R-squared	0.196	0.175	0.310	0.028	0.075	0.135

Note: * $<10\%$; ** $<5\%$; *** $<1\%$. Robust standard errors are in parentheses.

In round $j=3$, there were 118 observations (31.2%) with a streak of coin coming up heads up to $j=2$, and 152 observations (40.2%) with a streak of coin coming up tails up to $j=2$. In round $j=4$, there were 23 observations (6.1%) with a streak of coin coming up heads up to $j=3$, and 48 observations (12.7%) with a streak of coin coming up tails up to $j=3$. All regressions are estimated using a linear probability model and conditioning on non-prediction-buyers in the first round.

Figure 2: Proportion of prediction buyers by types of prediction streaks



Note: These are raw means and conditioning on non-prediction-buyers in the first round. 2-standard-error bands (95% C.I.) are reported, i.e. 2 above and 2 below se.

Figures 3A-3B: Proportion of prediction buyers by types of prediction streaks and by nationality

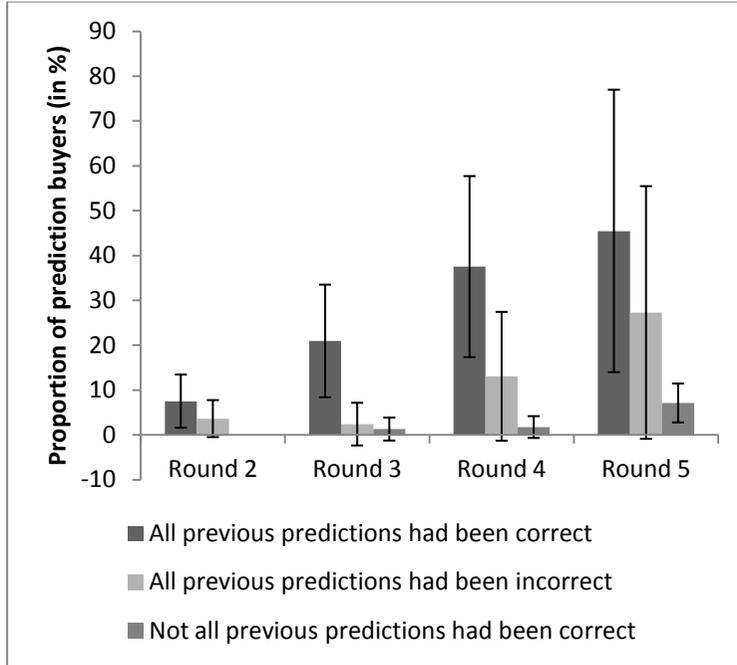


Figure 3A: *Thai sample*

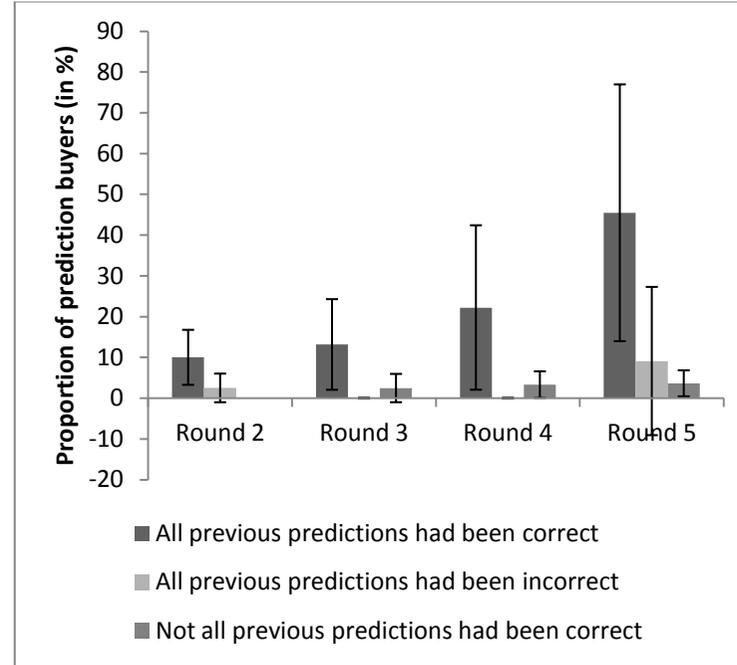


Figure 3B: *Singaporean sample*

Note: These are raw means and conditioning on non-prediction-buyers in the first round. 2-standard-error bands (95% C.I.) are reported, i.e., 2 above and 2 below se.

Table 2: Linear probability model estimates of factors determining the decision to buy prediction in each round

Dependent variable: Buy prediction in round j	$j=2$	$j=3$	$j=4$	$j=5$
a) Without controls				
All predictions up to $j-1$ had been correct	0.0568** [0.0262]	0.153*** [0.0434]	0.265*** [0.0688]	0.341*** [0.102]
All predictions up to $j-1$ had been incorrect		-0.00242 [0.0160]	0.0458 [0.0400]	0.144* [0.0837]
Bought prediction in round $j-1$		0.0976 [0.0811]	0.153* [0.0839]	0.278*** [0.0977]
Constant	0.0307** [0.0135]	0.0121 [0.0106]	0.0188* [0.0111]	0.0378*** [0.0126]
Observations	323	323	323	323
R-squared	0.015	0.097	0.159	0.189
b) With controls				
All predictions up to $j-1$ had been correct	0.0552** [0.0262]	0.153*** [0.0428]	0.211*** [0.0729]	0.273*** [0.104]
All predictions up to $j-1$ had been incorrect		0.00643 [0.0158]	0.0427 [0.0385]	0.156* [0.0798]
Bought prediction in round $j-1$		0.326 [0.289]	0.199 [0.180]	0.540*** [0.190]
Male	-0.0119 [0.0254]	0.0151 [0.0248]	-0.00503 [0.0244]	-0.0231 [0.0261]
Singaporean	0.0373 [0.0626]	0.0561 [0.0540]	0.0119 [0.0375]	-0.0475 [0.0401]
Proportion of correct answers in statistical test	-0.00609 [0.0501]	-0.0689 [0.0668]	-0.0131 [0.0518]	-0.104 [0.0652]
Endowment in round j	-0.000139 [0.000229]	-0.000161 [0.000148]	-0.000174 [0.000150]	-0.00016* [9.02e-05]
A streak of coin coming up heads up to $j-1$		-0.0683 [0.0495]	0.0437 [0.0749]	
A streak of coin coming up tails up to $j-1$		0.00786 [0.0359]	-0.0341 [0.0427]	-0.124*** [0.0387]
Made wrong bet in $j-1$	0.0182 [0.0331]	-0.00383 [0.0280]	0.0327 [0.0343]	-0.0712** [0.0316]
Bought prediction in $j-1$ × Made wrong bet in $j-1$		-0.0406 [0.152]	-0.472*** [0.163]	-0.493*** [0.190]

Fixed betting strategy, i.e. bet H or T in all j s	-0.0390 [0.0249]	-0.0379 [0.0275]	-0.0622** [0.0289]	-0.0110 [0.0251]
Bet the same as the predicted envelope in $j-1$	-0.0108 [0.0266]	-0.000208 [0.0232]	-0.0128 [0.0280]	0.0338 [0.0266]
Bought prediction in $j-1$ × Bet the same as envelope in $j-1$		-0.258 [0.295]	0.247 [0.182]	-0.0530 [0.153]
Constant	0.0528 [0.0506]	0.0900 [0.0640]	0.0786 [0.0508]	0.226*** [0.0710]
Observations	323	323	323	323
R-squared	0.030	0.137	0.227	0.293

Note: * $<10\%$; ** $<5\%$; *** $<1\%$. Robust standard errors are in parentheses.

All regressions are estimated using a linear probability model and conditioning on non-prediction-buyers in the first round. Dependent variable is a binary variable that takes a value of 1 if the subject paid to see the prediction in the corresponding numbered envelope, and 0 otherwise. Reference groups include: female; Thai students; no streak up to $j-1$; and made correct bet in $j-1$.

Table 3: Testing the gambler's fallacy versus the hot hand fallacy in behaviors towards the paid-for predictions

Dependent variable:				
Bet the same as the envelope in round j	$j=2$	$j=3$	$j=4$	$j=5$
All predictions up to $j-1$ had been correct	-0.0359 [0.0584]	-0.0619 [0.0756]	-0.171* [0.0972]	0.0249 [0.147]
All predictions up to $j-1$ had been incorrect		0.0124 [0.0700]	-0.0318 [0.0876]	-0.0943 [0.124]
Bought the prediction in round j	0.271* [0.157]	-0.217 [0.340]	0.0756 [0.198]	0.297** [0.126]
All predictions up to $j-1$ had been correct × Bought prediction in j	0.0830 [0.190]	0.799** [0.350]	0.623*** [0.222]	0.130 [0.196]
All predictions up to $j-1$ had been incorrect × Bought prediction in j		0.744** [0.350]	-0.228 [0.387]	-0.743*** [0.175]
The implied effect of betting the same as the prediction				
<i>Bought the prediction in j and all predictions had been correct up to $j-1$</i>	0.354*** [0.105]	0.581*** [0.072]	0.698*** [0.096]	0.427*** [0.170]
<i>Bought the prediction in j and all predictions had been incorrect up to $j-1$</i>		0.526*** [0.088]	-0.152 [0.330]	-0.445*** [0.127]
Observations	323	323	323	323
R-squared	0.052	0.083	0.072	0.068

Note: * $<10\%$; ** $<5\%$; *** $<1\%$. Robust standard errors are in parentheses.

Dependent variable takes the value of 1 if the subject bet according to the prediction in the corresponding round. Same control variables as in Table 2.

Table 4: The gambler's fallacy in first-round buyers

Dependent variable: Betting head in round j	Round $j = 3$		
	All	Prediction = H	Prediction = T
A streak of coin coming up heads up to $j-1$	-0.342* [0.144]	-0.500* [0.179]	-0.338 [0.229]
A streak of coin coming up tails up to $j-1$	-0.0921 [0.141]	-0.167 [0.172]	-0.131 [0.181]
Bought the prediction in round j		-1.03E-15 [1.18e-08]	-0.576* [0.317]
A streak of heads \times bought prediction in j		-1.31E-15 [0.438]	0.00433 [0.379]
A streak of tails \times bought prediction in j			-0.202 [0.352]
Constant	0.842*** [0.0860]	1.000*** [3.73e-09]	0.909*** [0.0962]
Observations	55	23	32
R-squared	0.104	0.213	0.295

Note: * $<10\%$; ** $<5\%$; *** $<1\%$. Robust standard errors are in parentheses.

See Table 1. Because of the small sample size, no additional controls were included.

Table 5: Buying behaviors of first-round buyers

Dependent variable: Buy prediction in round j	$j=2$	$j=3$
All predictions up to $j-1$ had been correct	0.165 [0.108]	0.295** [0.142]
All predictions up to $j-1$ had been incorrect		0.0356 [0.0942]
Bought prediction in round $j-1$		0.260* [0.131]
Constant	0.125* [0.0688]	0.0213 [0.0489]
Observations	55	55
R-squared	0.039	0.185

Note: * $<10\%$; ** $<5\%$. Robust standard errors are in parentheses.

See Table 2.

Table 6: OLS estimates of the log of bet amount placed in each round

Dependent variable: ln(bet amount in round j)	$j=2$	$j=3$	$j=4$	$j=5$
Bought prediction in round j	-0.00114 [0.156]	0.511** [0.208]	0.273 [0.196]	0.375** [0.145]
Bought prediction in round $j-1$		0.483 [0.731]	-0.678 [0.446]	0.198 [0.253]
Ln(bet amount in round $j-1$)	0.577** [0.0651]	0.725** [0.0590]	0.738** [0.0729]	0.736** [0.0703]
Male	0.000994 [0.0658]	0.0832 [0.0673]	-0.0338 [0.0763]	0.116 [0.0902]
Singaporean	1.087*** [0.215]	0.492*** [0.135]	0.404*** [0.116]	0.127 [0.127]
Proportion of correct answers in statistical test	0.0943 [0.153]	0.147 [0.137]	0.162 [0.155]	0.0716 [0.177]
Endowment in round j	-0.00427*** [0.000991]	-0.00181*** [0.000501]	0.273 [0.196]	-0.00108** [0.000478]
A streak of coin coming up heads up to $j-1$		-0.0214 [0.105]	0.156 [0.133]	
A streak of coin coming up tails up to $j-1$		-0.165** [0.0824]	-0.0115 [0.123]	-0.309** [0.140]
Made wrong bet in $j-1$	-0.946*** [0.0937]	-1.093*** [0.0833]	-0.710*** [0.0977]	-0.838*** [0.104]
Bought prediction in $j-1$ × Made wrong bet in $j-1$		-0.0603 [0.367]	-0.467 [0.446]	-0.863*** [0.277]
Fixed betting strategy, i.e. bet H or T in all j s	0.0551 [0.0836]	0.125 [0.0772]	-0.158* [0.0883]	-0.135 [0.101]
Bet the same as the predicted envelope in $j-1$	-0.0671 [0.0628]	-0.0274 [0.0632]	0.0851 [0.0723]	-0.0893 [0.0879]
Bought prediction in $j-1$ × Bet the same as envelope in $j-1$		-0.564 [0.711]	0.743*** [0.225]	-0.0526 [0.212]
Constant	2.160** [0.224]	1.408** [0.195]	1.146** [0.187]	1.618** [0.243]
Observations	321	320	321	319
R-squared	0.396	0.546	0.469	0.452

Note: *<10%; **<5%; ***<1%. Robust standard errors are in parentheses.

Dependent variable is log of the bet amount placed in round j . Robust standard errors are reported in parentheses.

Appendix A: Descriptive statistics

	Non-first-round buyers			First-round buyers
	All	Thais	Singaporeans	All
Round $j = 2$				
Correct predictions up to $j-1$	160	80	80	31
% of all predictions	50.56%	49.08%	50.00%	56.36%
Bought prediction in $j = 2$	19	6	8	9
% of all correct predictions up to $j-1$	8.75%	7.50%	10.00%	29.03%
Round $j = 3$				
Correct predictions up to $j-1$	81	43	38	11
% of all predictions	25.08%	26.38%	23.75%	20.00%
Bought prediction in $j = 3$	14	9	5	4
% of all correct predictions up to $j-1$	17.28%	20.93%	13.16%	36.36%
Round $j = 4$				
Correct predictions up to $j-1$	42	24	18	6
% of all predictions	13.00%	14.72%	11.25%	10.91%
Bought prediction in $j = 4$	13	9	4	1
% of all correct predictions up to $j-1$	30.95%	37.50%	22.22%	16.67%
Round $j = 5$				
Correct predictions up to $j-1$	22	11	11	1
% of all predictions	6.81%	6.75%	6.88%	1.82%
Bought prediction in $j = 5$	10	5	5	0
% of all correct predictions up to $j-1$	45.45%	45.45%	45.45%	0.00%
N	323	163	160	55

Note: Total N=378 (Thai: N = 177; Singaporean: N=201).