Payoff-Based Belief Distortion∗

Peiran Jiao†

Abstract

Heterogeneous beliefs often arise among people with the same information but different personally experienced payoffs. To explain this, I propose a mechanism in which experienced payoffs distort beliefs: gains lead an agent to relatively underweight negative new signals and thus to become overoptimistic, whereas losses do the opposite. I experimentally test this mechanism and find behaviour consistent with its predictions. The experiment created a setting where payoffs carried no informational value for Bayesian updating, and thus offered a strong test for the effect of payoffs on beliefs. The findings were robust, and distinct from alternative mechanisms in important ways.

Keywords: Experience, Payoffs, Belief Distortion, Reinforcement

JEL Classification: D81, D83, C91

∗The author thanks Peyton Young, Vincent Crawford, Raymond Duch, Johannes Abeler, Peter Wakker, Nick Chater, Olivier l’Haridon, Aurlian Baillon, Paul Zak, Joshua Tasoff, Henrich Nax, Carlos Cueva, John Jensenius III, and seminar/session participants at the CESS-Nuffield, Oxford, ETH Zurich, Erasmus, University of East Anglia, Fudan, Shanghai University of Finance and Economics, University of Glasgow, Leicester University, Maastricht University, Nankai University, Colby College, the Tinbergen Institute, University of Gothenburg, CESS-Santiago, SEF 2016, M-BEES 2016 for helpful discussions, suggestions and comments. The author also thanks Kaixin Bao, Patrick Lu, and Taotao Zhou for their excellent research assistantship. This project was completed when the author was a Postdoctoral Research Fellow at Nuffield College and Department of Economics, University of Oxford. This research was supported by the George Webb Medley/OEP Fund, the John Fell OUP Research Fund under project code CUD09320, and the Horizon 2020 Marie Curie Individual Fellowship under project code 797396. The experimental procedure was ethically approved by the Central University Research Ethics Committee at the University of Oxford under the project code ECONCIA13-039.

†Address: Tongersestraat 53, 6211 LM, Maastricht, the Netherlands. Telephone Number: +31(0)433884602. Email: p.jiao@maastrichtuniversity.nl.
1 Introduction

A growing body of research in economics suggests that personal experience has a significant impact on subsequent decisions.¹ Experience from an event typically consists of the observation of relevant information and the obtainment of payoffs. However, to date little has been done to address the issues that (1) payoff differences can lead to different personal experience from the same event, and that (2) these payoff differences can give rise to belief heterogeneity. This paper directly contributes to the understanding of these two issues, and demonstrates that people with different experienced payoffs may entertain different beliefs based on the same information, even when payoffs carry no useful information for belief updating, that is when they are information-free.

For instance, people who lived through the same earthquake may have suffered different amounts of losses, and thus hold different beliefs about future earthquakes; investors may have different ownership experience with the same stock, and thus hold different beliefs about that stock’s future performance. What they fail to realise is that they overweight personal payoffs, which may be determined by factors that are irrelevant for updating beliefs about future outcomes: such as their houses’ relative location to the epicentre of the earthquake, or their own buying and selling of the stock.²

This paper proposes a behavioural mechanism, the payoff-based belief distortion (PBBD), in which information-free experienced payoffs bias beliefs by distorting information processing. In the model, the agent observes realised outcomes to update belief about the underlying state, and can obtain payoffs from these outcomes. Each state corresponds to a risk profile that assigns a probability to each possible payoff, and the agent has clear preference ranking over the states. For instance, the model could apply to an entrepreneur who updates belief about own skill or the underlying economic condition after success or failure of own business, or to an investor who updates belief about own abilities or private information quality after gains or losses from own portfolio. Based on payoffs experienced, a reinforcement value can be calculated, which is a weighted sum of past payoffs as in reinforcement learning (e.g.²

¹Relevant literature will be discussed in more detail shortly.
²Admittedly, in reality payoffs may contain informational value: for instance, investors could make inference about their own skills (e.g. Gervais and Odean, 2001; Seru et al., 2010). Therefore, the experimental setup in this paper guarantees that the documented effect is free from such confounds.
Thorndike, 1898; Erev and Roth, 1998). This reinforcement value then influences the agent’s belief updating. Particularly, gains make the agent relatively underweight bad news and become overoptimistic about the more preferred states, while losses do the opposite. Therefore, heterogeneous beliefs can arise based on the same information and different payoffs experienced.

The model’s prediction on belief differences and the behavioural mechanism of biased information processing are then demonstrated to be valid in two complementary experiments. In both experiments, experienced payoffs were information-free, with no value for Bayesian updating, so the experiments offered a robust test of the effect of experienced payoffs on beliefs. In the first experiment, subjects viewed real-world stock price sequences with gains/losses exogenously manipulated between treatments. This experiment used a relatively natural task and established that subjects who gained were significantly more optimistic than those who lost after observing the same stock. The second experiment further investigated the behavioural mechanism: whether payoffs lead to distorted information processing. In this experiment, subjects observed ball drawings with replacement from one of two potential urns containing two types of balls, and predicted future draws. Gains/losses from the ball drawings were exogenously manipulated between treatments. Gains, compared with losses, made subjects more optimistic about balls associated with future gains. Additionally, relative to the Bayesian benchmarks, subjects who gained (lost) significantly overweighted signals that are associated with further gains (losses).

The effect of experienced payoffs on beliefs has not been formally tested so far, but there is some suggestive empirical evidence. Malmendier and Nagel (2011) show that witnessing higher stock returns correlates with more optimistic beliefs about future returns, although they did not specifically look at personally obtained gains and losses. More relatedly, Hoffmann and Post (2015) find investors extrapolate experienced return and risk. Kuhnen (2015) experimentally show that experiencing losses made subjects more pessimistic when updating beliefs.

---

3In the field, it is difficult to determine the objectively correct value of personal payoffs for subsequent belief updating and decision making.
4Yet the belief deviations from the Bayesian benchmark were not completely consistent with the model’s predictions, potentially due to some residual desirability bias.
5A third experiment reported in Supplementary Materials Section 5 further demonstrates that the beliefs results are valid when payoffs are updated every period due to subjects’ own decisions, and that experienced payoffs also exert significant impact on actions.
6However, compared to the experimental tests in this paper, Hoffmann and Post (2015) could
The above examples, as well as PBBD itself, are about how experienced payoffs influence subsequent beliefs, not choices. It is difficult to find direct evidence for PBBD in the field, because beliefs are not observable. However, the PBBD-driven belief channel can potentially reconcile some empirical findings about choice behaviour. First, people who lived through negative events, such as financial crises, tend to become less willing to take financial risks subsequently (e.g. Malmendier and Nagel, 2011; Malmendier et al., 2011; Bucher-Koenen and Ziegelmeyer, 2013; Guiso et al., 2013; Knüpfer et al., 2017; Giannetti and Wang, 2016). Second, personal experience of a natural disaster made people overestimate future occurrences of the disaster (e.g. Helweg-Larsen, 1999; Gallagher, 2014). Third, gains from IPO auctions, 401(k) portfolios, or common stocks made investors participate more in these markets (e.g. Kaustia and Knüpfer, 2008; Chiang et al., 2011; Choi et al., 2009), even when those gains were the result of luck (Anagol et al., 2019). Fourth, equity markets sometimes positively responded to bad macroeconomic news during expansionary periods (Andersen et al., 2007), which is potentially due to misperception of bad news after gains experience. These findings are potentially consistent with beliefs that are biased upward after gains and downward after losses.

Additionally, a novel prediction of PBBD under the depression babies context (Malmendier and Nagel, 2011) is that heterogeneous beliefs may arise not just between those who lived through an event and those who indirectly observed it, but also among those who obtained different payoffs from living through the event. This has not been formally tested in empirical works, but recent evidence shows that (1) experience may influence behaviour through the belief channel (e.g. Malmendier and Nagel, 2011, 2016; Koudijs and Voth, 2016), and that (2) payoff differences (or intensity of experience) matter (e.g. Bernile et al., 2017; Hanaoka et al., 2018).

The behavioural mechanism of PBBD could emerge from several neurobiological and psychological foundations. For instance, gains and losses activate different brain areas (e.g. Knutson and Bossaerts, 2007), which could lead to differences in

---

7This is by having appropriate assumptions about preferences (e.g. a risk-neutral expected utility maximizer) or about the causal link between beliefs and actions. See Costa-Gomes et al. (2014) for an example that confirms such a link in an experimental study.

8This is part of their results. As the authors put it, “bad macroeconomic news has the traditionally-expected negative equity market impact during contractions, but a positive impact during expansions.” (p.276)

9The nucleus accumbens (NAcc) is related to positive emotions and the processing of gains (e.g.
information processing. Kuhnen and Knutson (2011) posit that gains (losses) trigger positive (negative) emotional states, leading to optimism (pessimism) due to a ‘self-preservation motive’. Moreover, availability bias (e.g. Tversky and Kahneman, 1973) may play a role. While the availability heuristic does not specify what is considered more available in a decision maker’s mind, PBBD could offer an instantiation of the availability heuristic when payoffs are considered, and are used to update belief about the state of the world. Both require the decision maker to believe in some positive correlation between past and future outcomes.\(^\text{10}\)

It is noteworthy that PBBD is closely related to but distinct from the following mechanisms in important ways. These alternative mechanisms will be tested in my experimental data, whereas the PBBD results hold even after considering them. (1) Pure reinforcement learning (e.g. Erev and Roth, 1998): In PBBD, past payoffs influence the weighting of new signals, and thus new information processed in the light of old payoffs can lead to larger biases in belief and behaviour; but in reinforcement learning, the arrival of new information that does not influence payoffs should not matter. (2) Confirmation bias (e.g. Rabin and Schrag, 1999): While both are about belief consistency with some prior, this prior is experienced payoffs in PBBD and prior belief (typically exogenously given) in confirmation bias models. The two could be connected if prior beliefs are formed based on payoff information. However, they could work in opposite directions if prior beliefs are based on all observed outcomes, without a bias towards the subset from which payoffs are obtained. The contrast is larger if payoffs come from a limited unrepresentative subset of realised outcomes. (3) Hot hand fallacy (e.g. Gilovich et al., 1985): While both are beliefs in continuation, PBBD can be thought of as a belief in the continuation of similar payoff experiences, and hot hand fallacy relies on streaks in observed outcomes that are usually payoff-independent.

The paper is structured in the following way. Section 2 presents the model. Section 3 contains the design, procedure and results of both experiments, whereas robustness checks and tests of alternative mechanisms are in Section 4. Section 5 discusses the practical and policy implications of PBBD with examples from financial decision-making. Section 6 concludes.

---

\(^{10}\)Psychologists have studied how different experienced payoffs may drive belief differences (e.g. Helweg-Larsen, 1999).

Breiter et al., 2001), whereas the anterior insula is related to negative emotions and the processing of losses (e.g. Chua et al., 1999).
2 A Model of Payoff-Based Belief Distortion

2.1 The Model

This section proposes a behavioural model in which experienced payoffs bias beliefs by distorting information processing. The model involves belief updating under uncertainty in repeated decision-making with payoff feedback. I first present the general model, and then provide a simple example that will be more directly applicable in the experimental tests. I will discuss implications of the model and generate two testable hypotheses.

In this model, uncertainty is represented by a finite set of possible states \( \mathcal{A} = \{A_1, A_2, ..., A_M\} \). Time is discrete and varies over \( t = 0, 1, 2, ... \). There is a finite set of payoffs \( \mathcal{A} = \{a_1, a_2, ..., a_N\} \), one of which is realised independently every period \( t = 1, 2, ... \). Payoffs are real-numbered, \( a_n \in \mathbb{R} \), and ordered such that \( a_1 \leq a_2 \leq ... \leq a_N \), and inequality holds at least in some cases.\(^{11}\) The realised payoff in period \( t \) is denoted \( a_t \).

For each \( A^m \), the probability of realising payoff \( a^n \) in any period is represented by \( \theta_{nm} = p(a^n | A^m) \). For each \( m = 1, 2, ..., M \), let \( \theta_m = (\theta_{1m}, \theta_{2m}, ..., \theta_{Nm}) \), where \( \theta_m \in \Delta^{N-1} = \{\theta_m \in \mathbb{R}_+^N : \sum_{n=1}^N \theta_{nm} = 1\} \).\(^{12}\) The possible states in \( \mathcal{A} \) are ordered in the sense of first order stochastic dominance. That is, if \( j > i \), \( p(a \geq a^n | A^j) \geq p(a \geq a^n | A^i) \) for all \( a^n \), and \( p(a \geq a^n | A^j) > p(a \geq a^n | A^i) \) at least for some \( a^n \). Therefore, assuming that the agent has monotonically increasing preferences, \( A^M \) is her most preferred state, and \( A^1 \) is the least preferred. This assumption about preferences is needed because the model is about how better experienced payoffs make the agent more optimistic about the more preferred states.

In period \( t = 0 \), nature draws \( A^m \) for all subsequent periods but the agent does not know which state is drawn. Her prior belief is that \( A^m \) is drawn with probability \( \mu_0(A^m) \) for \( m = 1, 2, ..., M \). She needs to observe realised payoffs to update this belief. In each period, she also chooses an action \( c_t \in \{0, 1\} \). If \( c_t = 0 \), she will not obtain the payoff \( a_{t+1} \) but she can still observe it; and if \( c_t = 1 \), she will obtain the payoff in period \( t + 1 \).

\(^{11}\)The payoffs cannot be all equal, otherwise they are not informative of the underlying state and the agent cannot update belief.

\(^{12}\)If two payoffs were equal (\( a^i = a^j \)), they can be taken as the same outcome and their probabilities can be added together in each state of the world \( \theta_{im} + \theta_{jm} \).
Define $R_t$ as the agent’s payoff experience, or reinforcement value, up to period $t \geq 1$,

$$R_t = \phi R_{t-1} + c_{t-1} a_t,$$

where $\phi \in [0, 1]$ represents the discounting or forgetting of past experienced payoffs. So $R_t$ is just a discounted sum of previous obtained payoffs. When $\phi = 1$, experienced payoffs from all previous periods are weighted equally, and when $\phi = 0$, only the most recent payoff matters. The specification of $R_t$ is related to reinforcement learning in psychology and economics literature (e.g. Thorndike, 1898; Erev and Roth, 1998; Camerer and Ho, 1999). In reinforcement learning, an action is reinforced and more likely chosen subsequently if it has higher reinforcement values; actions associated with lower reinforcement values are more likely avoided. The PBBD model can thus be considered as linking reinforcement to belief updating. Although it would be interesting to estimate $\phi$, in the following experimental tests, the results will not rely on the value of this parameter. So most tests are performed with the assumption of either $\phi = 0$ or $\phi = 1$.

The agent starts with no payoff experience $R_0 = 0$, a prior belief of $\mu_0$, and chooses an action $c_0$. In $t = 1$, a new payoff $a_1$ is realised, and her payoff experience updates to $R_1$. The agent updates her belief to $\mu_1$, and accordingly chooses action $c_1$. Choice is not modelled here, as it depends on the agent’s preferences. But as a simple example of a risk-neutral expected utility maximizer, she will choose $c_t = 1$ if expected payoff is positive given her posterior belief, and $c_t = 0$ otherwise. The agent’s problem can be summarised in Figure 1.

Suppose the payoff in period $t$ is $a_t = a^n$. One can show that a Bayesian agent’s log likelihood ratio of $A^1$ relative to $A^2$ is

---

13 Reinforcement learning (and its variants) is a parsimonious model that can describe choice behaviour relatively well in repeated decisions from experience (e.g. Erev et al., 2010) and even in some situations with decisions from description (e.g. Nax et al., 2016). But this literature does not directly address the impact of payoffs on beliefs.

14 There are other features in reinforcement learning models that could be included here, such as the effects of forgone payoffs, inertia, surprise, regret, and reference point (see e.g. Erev and Roth, 1998; Foster and Young, 2006; Nevo and Erev, 2012). Their relationship with belief biases could be interesting for future research.

15 Therefore, this model is about how people process a new signal under the influence of all previous payoff experience. Admittedly, it would also be interesting to investigate how people process a new signal under the influence of payoffs associated with that signal, as pointed out by an anonymous referee. But this is out of the scope of the current model.
\[ \ln \Lambda_{t}^{12} = \ln \Lambda_{t-1}^{12} + \ln \frac{\theta_{n1}}{\theta_{n2}}, \tag{2} \]

where \( \Lambda_{t}^{12} \) represents the Bayesian log likelihood ratio of \( A^1 \) relative to \( A^2 \) in period \( t \). In each period \( t \), his log likelihood ratio between any two states is updated exactly by the log likelihood ratio of observing \( a_t = a^n \) between the two states. This Bayesian belief updating process is not influenced by experienced payoffs. Specifically, actions \( c_t \) and the previously obtained payoffs \( R_t \) do not matter for belief updating.

By contrast, a PBBD agent’s belief updating in period \( t \) is influenced by \( R_t \). Although this bias can take many forms, here I propose a mechanism in which gains make the agent overweight positive news relative to negative news, and losses do the opposite: this is because the agent misperceives new information as being better or worse depending on her experienced payoffs. When \( R_t > 0 \), with some probability \( q(R_t) \in [0, 1] \) she could misperceive the true payoff \( a_t = a^i \) as a better payoff \( a^j > a^i \). But note that if \( a_t = a^N \), the largest possible payoff, then she does not misperceive. Conversely, when \( R_t < 0 \), with probability \( q(R_t) \) she could misperceive \( a_t = a^j \) as a worse payoff \( a^j < a^i \), and she does not misperceive if \( a_t = a^1 \), the smallest possible payoff. Additionally, \( q(0) = 0 \), that is when \( R_t = 0 \) she does not misperceive.

This is equivalent to underweighting the actual realised payoff \( a_t \) when updating beliefs, but placing a non-zero weight on better payoffs when \( R_t > 0 \), and on worse payoffs when \( R_t < 0 \).

We can re-write Equation (2) as

\[ \ln \Lambda_{t}^{12} = \ln \Lambda_{t-1}^{12} + \sum_{i=1}^{N} I(a^i = a_t) \ln \frac{\theta_{i1}}{\theta_{i2}}, \tag{3} \]

where \( I(a^i = a_t) \) is an indicator function that is equal to 1 if \( a^i = a_t \) and 0 otherwise.
Then the biased agent’s log likelihood ratio of $A^1$ relative to $A^2$ updates from period $t - 1$ to $t$ according to the following

$$
\ln \Lambda_{t,Biased}^{12} = \ln \Lambda_{t-1,Biased}^{12} + \sum_{i=1}^{N} \left[ (1 - q(R_t)) I(a^i = a_t) + q^+(a^i|a_t, R_t) I(a^i \geq a_t) I(R_t > 0) + q^-(a^i|a_t, R_t) I(a^i \leq a_t) I(R_t < 0) \right] \ln \frac{\theta_{i1}}{\theta_{i2}} 
$$

where $I(a^i \geq a_t) = 1$ when $a^i \geq a_t$, and 0 otherwise; similarly, $I(a^i \leq a_t)$, $I(R_t > 0)$ and $I(R_t < 0)$ are equal to 1 respectively when $a^i \leq a_t$, $R_t > 0$ and $R_t < 0$; and $q^+ \in [0, 1]$ and $q^- \in [0, 1]$ are respectively the weights placed on each $a^i$ when $a^i \geq a_t$ and when $a^i \leq a_t$ respectively. Note that $q^+$ and $q^-$ depend on $a_t$ and $R_t$, such that regardless of whether $R_t > 0$ or $R_t < 0$, the total weights always sum up to 1.\(^{16}\)

The agent overestimates the probability of the more preferred states when $R_t > 0$, and underestimates when $R_t < 0$, compared with a Bayesian agent. When $R_t = 0$, Equation (4) becomes the same as Equation (3) and the agent is Bayesian.

**Hypothesis 1 (H1):** With positive reinforcement, a PBBD agent overestimates the likelihood of more preferred states relative to Bayesian; with negative reinforcement, she underestimates.

Additionally, two things are interesting and open for future research. The first one is the properties of $q(R_t)$. It is possible that $q(R_t)$ could be asymmetric between $R_t > 0$ and $R_t < 0$. It is also possible that the magnitude of the belief bias could vary with the size of the experienced payoffs. For instance, more extreme payoff experience may lead to a larger belief bias.\(^{17}\) However, in what follows, a simple working assumption is that $q(R_t)$ is an exogenous constant when $R_t \neq 0$.

Secondly, when $R_t > 0$, $a_t$ is underweighted by $q(R_t)$ and this is redistributed

---

\(^{16}\)For instance, one specific form is $q^+(a^i|a_t, R_t) = \sum_{a^j \geq a_t} q^+(a^j) q(R_t)$. Thanks to the suggestion of an anonymous referee.

\(^{17}\)For instance, $q$ may have the following desirable properties: i. $q(0) = 0$; ii. $q(|R_t|) > 0$; iii. $q(R_t) \to 1$, when $|R_t| \to \infty$. Intuitively, when there is no prior experience, or when positive and negative experiences exactly cancel out, there is no bias; more extreme experienced payoffs lead to larger bias; and this probability is bounded between 0 and 1. A potential functional form of $q(R_t)$ could take the logistic form $q(R_t) = \frac{2}{1+e^{-\nu|R_t|}} - 1$, where $\nu \geq 0$ is a measure of the sensitivity of $q$ to $R_t$. When $\nu = 0$, the agent is unbiased.
to payoffs larger than \( a_t \); similarly when \( R_t < 0 \), \( q(R_t) \) is redistributed to payoffs smaller than \( a_t \). But how \( q(R_t) \) is redistributed is open for discussion. For instance, in a simple case, \( q(R_t) \) could be uniformly distributed among payoffs larger or smaller than \( a_t \). In what follows, this complication is avoided by assuming binary payoffs.

2.2 A Simple Example

This section offers a simple example of the model that is more readily applicable in the following experimental test. In this example and the experimental tests, I retain the most essential features of the model, making some simplifications. First, payoffs are obtained before new signals are observed and beliefs are updated. This is because if payoffs are obtained based on choices, then confounds, such as self-justification or desirability bias, can also drive beliefs.\(^\text{18}\) Second, the relationship between the magnitude of \( R_t \) and belief bias is interesting to test, but the main results of this paper will focus on the sign of \( R_t \). This is because a key novel prediction of the model is that gains and losses will lead to different belief biases. I will, however, also provide suggestive evidence that more extreme realised payoffs are likely associated with larger belief biases.

Suppose there is a risky stock, which is either good or bad, \( \mathcal{A} = \{G, B\} \). In each period it generates a payoff of either \( U \in \mathbb{R}_+ \) if price goes up, or \( D \in \mathbb{R}_- \) if price goes down, \( \mathcal{A} = \{U, D\} \). For a good stock the probability of price going up is \( \theta \in (\frac{1}{2}, 1) \), and for a bad stock the probability of price going up is \( 1 - \theta \). Price cannot stay unchanged, so the probability of price going down is one minus the probability of price going up.

In period \( t = 0 \), nature draws either \( G \) or \( B \). Then in each subsequent period \( t \), the agent observes payoff \( a_t \) in that period, updates her belief to \( \mu_t \) and accordingly makes an investment decision \( c_t \in \{0, 1\} \). In this process, her payoff experience is updated according to Equation (1).

Suppose the true sequence of signals are \( S = U, U, U, D, D \). A Bayesian agent has the following posterior likelihood ratio of good relative to bad stock, regardless of experienced payoffs:

\(^\text{18} \)Supplementary Materials Section 5 reports an additional experiment which allowed for the simultaneous updating of payoffs and beliefs, with subjects’ endogenous choice of action, and also found that beliefs were biased by experienced payoffs. Thanks to the suggestion of two anonymous referees.
\[ \Lambda_{\text{Bayesian}} = \frac{P(G|S)}{P(B|S)} = \frac{\theta}{1 - \theta}. \]  

(5)

For a PBBD agent, in particular, after gains she could misperceive \( D \) signals as \( U \) (but not misperceiving \( U \)), and over-predict the probability of \( G \); and the converse is true after losses. For simplicity, assume (a) that the agent obtained payoff experience \( R \) before observing sequence \( S \) and no new payoff is obtained while observing \( S \), and (b) that \( q \), the probability of misperceiving signals, is an exogenous constant.

Therefore, the biased agent can be considered as perceiving signals from a biased source, where signals are drawn according to the following probabilities:

\[
\sigma = P(U|R > 0, G) = P(D|R < 0, B) = \theta + q(1 - \theta); \\
\sigma' = P(U|R > 0, B) = P(D|R < 0, G) = (1 - \theta) + q\theta,
\]

(6) \( \sigma' \)

(7)

where \( \sigma \) is the probability of perceiving \( U \) (\( D \)) given \( R > 0 \) (\( R < 0 \)) and a good (bad) stock, and \( \sigma' \) is the probability of perceiving \( U \) (\( D \)) given that \( R > 0 \) (\( R < 0 \)) and a bad (good) stock. This is because, for instance, when the stock is good, the probability of observing \( U \) is \( \theta \), and the probability of observing \( D \) but misperceiving it as \( U \) is \( q(1 - \theta) \) for an agent with \( R > 0 \).

The mistake of a PBBD agent is in that (a) she misperceives signals contradicting with her prior payoff experiences, and (b) she is unaware of her bias and still uses the original outcome distribution (i.e. \( P(U|G) = \theta \) and \( P(U|B) = 1 - \theta \)) to update beliefs. Suppose, a PBBD agent with \( R > 0 \) misperceives the last \( D \) as \( U \), so her perceived sequence is \( \tilde{S} = U, U, U, D, U \). Her posterior likelihood ratio would be

\[
\Lambda_{R>0}^{\text{Biased}} = \left( \frac{\theta}{1 - \theta} \right)^3,
\]

(8)

with \( \Lambda_{R>0}^{\text{Biased}} > \Lambda_{\text{Bayesian}} \). However, if the agent is aware of her bias, she could have done better. With the perceived sequence \( \tilde{S} = U, U, U, D, U \), if she knew how the bias works and the probability \( q \) of misperceiving signals, her posterior likelihood ratio would be
\[ \Lambda_{R>0}^* = \frac{(\theta + q(1-\theta))^4(1-\theta)}{((1-\theta) + q\theta)^4\theta}, \]  

(9)
as any \( U \) is potentially the result of misperception, and \( \Lambda^* \) can be considered as the best she could have done given her perceived signals. It can be shown that \( \Lambda_{R>0}^* < \Lambda_{R>0}^{\text{Biased}} \). Therefore, with positive reinforcement, the biased agent overestimates the likelihood of \( G \).

Similarly, for a biased agent who has \( R < 0 \), her perceived sequence could be \( \tilde{S}' = U, U, D, D, D \), i.e. she misperceived the third \( U \) as \( D \), and her posterior likelihood ratio is

\[ \Lambda_{R<0}^{\text{Biased}} = \left( \frac{\theta}{1-\theta} \right)^{-1}, \]  

(10)
with \( \Lambda_{R<0}^{\text{Biased}} < \Lambda_{R<0}^\text{Bayesian} \). If she took into account of her bias, her posterior likelihood ratio would be

\[ \Lambda_{R<0}^* = \frac{\theta^2(\theta q + (1-\theta))^3}{(1-\theta)^2(q(1-\theta) + \theta)^3}, \]  

(11)
with \( \Lambda_{R<0}^{\text{Biased}} < \Lambda_{R<0}^* \). Therefore, with negative reinforcement, the biased agent underestimates the likelihood of \( G \). These will be tested in the experiment.

It is also noteworthy that after observing sufficiently large number of signals, a Bayesian agent should know the truth, but the biased agent’s belief does not converge to Bayesian, unless she is aware of her bias and the value of \( q \).\(^{19}\) In the above simple example, the order of signals does not matter. The order signals could matter when the agent could obtain payoffs while observing the sequence, or when there are other path-dependent belief biases (such as Rabin, 2002).

Some comparative statics can be derived. First, for instance, if more extreme experienced payoffs lead to higher probability of misperception, the agent will have more signals misperceived in expectation, and hence a larger bias, i.e. \( \frac{\partial}{\partial R}(\frac{\Lambda_{R>0}^{\text{Biased}}}{\Lambda^*}) > 0 \).\(^{20}\)

\(^{19}\)The proof of this is in Supplementary Materials Section 1.1.

\(^{20}\)The proof of this is in Supplementary Materials Section 1.2.
Additionally, denote the numbers of true $U$ and $D$ signals by $n_U$ and $n_D$, and the numbers of perceived $U$ and $D$ signals by $\tilde{n}_U$ and $\tilde{n}_D$, respectively. Given a fixed length of signal stream, perceiving more signals consistent with one’s experienced payoffs (larger $\tilde{n}_U$ when $R > 0$ or larger $\tilde{n}_D$ when $R < 0$) generates a larger bias because there are more opportunities to misperceive, and thus $\frac{\partial}{\partial n_U}(\frac{\Lambda_{\text{Biased}}}{\Lambda^*}) > 0.$\textsuperscript{21}

2.3 Signal Weighting

Experienced payoffs lead to misperception of signals, and therefore the agent places wrong weights on signals when updating beliefs relative to a Bayesian agent. This section derives the implication of the model for biased signal weighting, which will then be directly testable in the following experiments.

Without knowing a biased agent’s perceived signals ($\tilde{n}_U$ and $\tilde{n}_D$), an objective observer can form an expectation about a biased agent’s belief distortion, given $n_U$, $n_D$, and the sign of $R$. The observer can calculate the posterior likelihood ratios assuming the biased agent misperceives a proportion $q$ of the signals contradicting with her payoff experience, and obtain the following

$$\Lambda_{\text{Observer}}^{R>0} = \left(\frac{\theta}{1-\theta}\right)^{n_U-n_D+2qn_D}; \quad (12)$$

$$\Lambda_{\text{Observer}}^{R<0} = \left(\frac{\theta}{1-\theta}\right)^{n_U-n_D-2qn_U}. \quad (13)$$

It is easy to see that $\Lambda_{\text{Observer}}^{R<0} < \Lambda_{\text{Bayesian}} < \Lambda_{\text{Observer}}^{R>0}$, i.e. the observer believes the biased agent overestimates $G$ when $R > 0$, and underestimates when $R < 0$ in expectation.

From Equations (12) and (13):

$$\ln \Lambda_{\text{Observer}}^{R>0} = \ln \left(\frac{\theta}{1-\theta}\right) n_U - \ln \left(\frac{\theta}{1-\theta}\right) (1-2q)n_D; \quad (14)$$

$$\ln \Lambda_{\text{Observer}}^{R<0} = \ln \left(\frac{\theta}{1-\theta}\right) (1-2q)n_U - \ln \left(\frac{\theta}{1-\theta}\right) n_D. \quad (15)$$

\textsuperscript{21}The proof of this is in Supplementary Materials Section 1.3.
These can be compared with a Bayesian agent’s posterior log likelihood ratio,

$$\ln \Lambda^{Bayesian} = \ln \left( \frac{\theta}{1-\theta} \right) n_U - \ln \left( \frac{\theta}{1-\theta} \right) n_D. \quad (16)$$

Therefore, a Bayesian agent places equal weights on each new $U$ or $D$ signal observed. A PBBD agent, on the other hand, places relatively smaller weight on $D$ signals than on $U$ signals when $R > 0$, and smaller weight on $U$ than on $D$ when $R < 0$.\(^{22}\) The bias in signal weighting is determined by $q$, the probability of misperceiving signals. Hence the following hypothesis, which will be directly tested in the experiment.

**Hypothesis 2 (H2):** With positive reinforcement, a PBBD agent overweights good signals relative to bad signals; with negative reinforcement, the opposite is true.

## 3 Experiments

This section reports the results from two experiments that check the validity of the model. For most parts of the experimental test, I focus on a parsimonious model with a constant $q \in (0, 1)$, that is a constant probability of misperceiving signals inconsistent with one’s payoff experience.

### 3.1 Experiment 1: Stock Price Prediction

#### 3.1.1 Design

This experiment tests H1 in a simple and natural setup, where subjects observe stock prices adapted from real-world data, obtain payoffs exogenously manipulated between treatments, and make predictions about future prices of the stocks.\(^{23}\)

Participants were told that they would observe price sequences adapted from real stocks traded on NYSE. All prices were scaled to within 0 to 200 ECU (Experimental Cash Unit) and participants did not have any other information about the stocks.\(^{24}\) They viewed 24 unique price sequences, one at a time, each containing 101 periods

\(^{22}\)If more extreme $R$ leads to an increase in $q$, then it should also lead to larger relative underweighting of signals contradicting with one’s experienced payoff.

\(^{23}\)Full instructions of both experiments can be found in Experimental Information.

\(^{24}\)All ECU earnings in this experiment were finally converted to cash at a rate of 150 ECU = $1.
(Period 0 to 100). Two sequences were repeated to check consistency. All sequences were shown in cross-sectionally different random orders. Before viewing each stock sequence, subjects received 800 ECU cash endowment plus some additional cash or share endowments depending on treatments. They first observed prices up to Period 20 when endowed shares, if any, were sold. The Period 20 announcement contained the ECUs gained/lost from the stock, and their updated cash balance. Subjects were also told that price changes after Period 20 would not influence payoffs. Then they predicted Period 100 price, viewed prices up to Period 80, and predicted Period 100 price again. Subjects then moved on to the next stock without feedback on accuracy, which was to avoid learning effect. But they were told that they could view the true Period 100 prices after the experiment.

Here is how gains/losses were determined. Subjects were randomly assigned to 4 treatments (GAIN10, GAIN20, LOSS20 and NO) with different payoff experiences in the first 20 periods. Payoffs were exogenously manipulated according to whether the share endowment was a long position (GAIN), a short position (LOSS) or no position (NO); 20 shares or 10 shares. That is, GAIN treatment subjects gained from price increases and lost from price decreases in the first 20 periods. LOSS means the opposite. The instructions did not mention long or short position at all to avoid confusion. Subjects were only told that they receive some share endowments of the stock in Period 0, that shares were sold between Period 0 and Period 20, and how much they gained/lost in Period 20.

Here are three more important design choices. (1) Subjects had no stakes in the stock when reporting beliefs. This is to avoid any endowment effect, or a belief bias due to personal stake, such as desirability bias or wishful thinking (e.g. Brunnermeier and Parker, 2005; Mayraz, 2011). For example, if subjects had owned shares, they could have been more likely to over-predict price increases because these are more desirable. (2) Subjects did not make any buying or selling decisions, which avoids self-justification concerns. Otherwise, subjects could later distort beliefs about the

25 It was guaranteed that the two repetitions were far apart so that participants couldn’t recognise.

26 Instead of using long/short positions, another way to let some subjects gain while others lose from the same price chart is to assign roles of consumers and producers in the same market, like that done in Mayraz (2011), but such framing naturally induces the desirability bias, which is a confound. Another method is to let one group observe mirror image of the price charts of the other group during the first 20 periods. However, subjects in different treatments won’t see the same price chart, which could give them a reason to form different beliefs rather than just the experienced gain/loss.
stock so as to maintain a positive self-image, not purely because of gain/loss from the stock. (3) Wealth effect was controlled for, which otherwise could also explain the different risk attitudes or beliefs as a result of different payoffs. Although treatment NO subjects had no gain/loss from stocks, their cash endowment in Period 0 of each stock matched the amount of cash remaining for the same stock in treatment GAIN20.

In this design, subjects in all treatments had exactly the same information (price charts). Personal gains/losses were not additionally informative about future prices. Conventional theories of belief updating that are payoff-independent should predict no belief difference between treatments. Therefore, the design is a robust test of the effect of experienced payoffs on beliefs.

I elicited subjects’ median belief about Period 100 price using the Exchangeability Method (e.g., Baillon, 2008; Abdellaoui et al., 2011). This method has several advantages over its alternatives. In each elicitation, subjects divided the state space (0 to 200) into two equally-likely subspaces: they chose a number between 0 and 200 so that they were indifferent between winning a prize if the true price is above that number and winning the same price if the true price is below.

Formally, subjects were asked to choose a value \( r^* \) on a slider bar representing \( r \in [0, 200] \), so that they were indifferent between two lotteries: Lottery A, win 200 ECUs if Period 100 price is within \([r, 200]\) and 0 otherwise, denoted \( 200_{[r, 200]}0 \); and Lottery B, \( 200_{[0,r]}0 \). Clearly, for small enough \( r \) (say 0), Lottery A is preferable; for large enough \( r \) (say 200), Lottery B is preferable. A choice of \( r^* \) indicates one prefers Lottery A for all \( r < r^* \), and Lottery B for all \( r > r^* \). To determine the prediction reward, a number \( Z \) was drawn from a uniform distribution on \([0, 200]\); if \( Z < r^* \), Lottery A \( (200_{[Z,200]}0) \) was played for real; if \( Z > r^* \), Lottery B \( (200_{[0,Z]}0) \) was played; if \( Z = r^* \), both lotteries were equally likely played. The same question was asked in Periods 20 and 80 for each stock. Supplementary Materials Section 2 demonstrates that this method is incentive compatible in the current experiment.

\footnote{For example it avoids issues of changing risk attitudes, nonlinear probability weighting, complicated math, and source dependence. For a review of belief elicitation methods, see e.g., Schlag et al. (2015).}

\footnote{As they moved the slider, the value of \( r \) in Lotteries A and B changed correspondingly. To avoid the hard concept of indifference, I asked them to choose so that for values smaller than \( r^* \) they prefer Lottery A, and for values larger than \( r^* \) they prefer Lottery B, and \( r^* \) is the smallest value at which they just switch from choosing Lottery A to choosing Lottery B.}

\footnote{Subjects were given examples and practice questions, were allowed to ask questions, and were also told that the reward-maximizing strategy was to always choose according to one’s true beliefs. They could not start the experiment until they correctly answered all comprehension questions.}
In the end, rewards from the experiment consisted of two parts, the cash balance and the prediction reward, each from a randomly selected stock.

3.1.2 Procedure

The experiment was conducted over 8 sessions at Nuffield College Centre for Experimental Social Sciences (CESS), University of Oxford. Experimental sessions were advertised within the subject pool of CESS, and participants were recruited through ORSEE (Greiner, 2015). Totally 122 participants (62 male, 74 full-time students, average age 28) volunteered for the experiment. Non-student participants were residents in Oxford city. There were 28 participants in treatment GAIN10, 29 in GAIN20, 32 in LOSS20 and 33 in NO. Experiments 1 and 2 were conducted together, as two separate stages. In addition to the earnings from the experimental tasks, there was a flat show-up fee of $4. Each session took 1.5 to 2 hours and the average payment was $20. Finally, participants completed a questionnaire about their demographic information, educational backgrounds, math and economics classes taken, etc.

The experimental procedure was ethically approved by the Central University Research Ethics Committee at the University of Oxford. Participants were fully consented prior to the experiment, and then completed the experiment via computer interface. They were not allowed to use calculators or any computer programme, but were offered paper and pens. The experimental task started after subjects correctly answered a set of comprehension questions.

3.1.3 Results

Do exogenously obtained payoffs induce a systematic belief bias? This experiment addresses this question by comparing elicited beliefs between treatments that varied payoff experiences. Summary statistics of these elicited beliefs are in Table A3 of Supplementary Materials. To check belief consistency, I compared reported beliefs on the two repeated stocks. Between each pair of these stocks, the Wilcoxon test did not reveal any significant belief difference ($p > 0.10$). The pairwise correlations between the two elicitations were significantly positive ($p < 0.001$). Therefore subjects reported consistent beliefs.\(^{30}\)

H1 states that more optimism should be expected after gains than after losses.

\(^{30}\)Details of the consistency check results are in Table A4 of Supplementary Materials.
More specifically, for stocks that increased up to Period 20, GAIN subjects gained, while LOSS subjects lost; so compared with condition NO, GAIN subjects should overestimate Period 100 prices and LOSS subjects should underestimate. The opposite should be true for stocks that decreased in the first 20 periods.

This effect is demonstrated graphically in Figure 2. In order to compare between treatments, I generated a measure that is easy to aggregate over all stocks. This measure is the deviation of elicited beliefs from a benchmark. Since there is no objectively rational benchmark for belief about Period 100 prices in this experiment, I use the average belief on each stock across all treatments and subjects as a benchmark. The cumulative distributions of these belief deviations are plotted in Figure 2, for beliefs in Period 20 and Period 80 separately, and for stocks that increased and decreased in the first 20 periods separately. The top panels show that for stocks that increased by Period 20, GAIN10 and GAIN20 subjects gained, whereas LOSS20 subjects lost, so GAIN10 and especially GAIN20 subjects had more optimistic beliefs than NO, which had more optimistic beliefs than LOSS20 subjects. By contrast, the middle panels show that the reverse is true for stocks that decreased by Period 20, because GAIN10 and GAIN20 subjects lost, whereas LOSS20 subjects gained. These observations are in line with the model’s predictions. (all are significant with \( p < 0.01 \) in Wilcoxon tests of each treatment relative to NO, except in Panels D and E, GAIN10 is not significantly different from NO with \( p > 0.10 \), although the difference is in the predicted direction). Additionally, the two bottom panels categorise observations only according to whether subjects gained or lost, regardless of treatments. This confirms the model’s prediction that subjects who gained reported higher expected future prices than those who lost (\( p < 0.001 \) in Wilcoxon tests).

The effects are also demonstrated in the OLS regression results of Table 1: the dependent variable was the belief deviation from the mean belief on each stock across all subjects. The main independent variables are dummy variables GAIN and LOSS, which respectively equal to 1 when the subject gained or lost on a stock in the first 20 periods, and 0 otherwise. For subjects in the NO treatment, both GAIN and LOSS are equal to 0. Period 20 beliefs are tested in regressions (1) to (3), and Period 80 beliefs are tested in regressions (4) to (6). Regressions (2) (3) (5) and (6) included control variables. Variable Uncertainty is equal to 1 when the standard deviation of all beliefs on a stock exceeds the median standard deviation across all stocks. Uncertainty surrounding a stock is expected to have an influence on the magnitude of
Figure 2: Cumulative Distribution

Note: This figure shows the cumulative distribution of belief deviations from mean belief on each stock in Experiment 1. The curves were drawn for the four treatments respectively (Panels A B D E), or beliefs following gains and losses respectively (Panels C F).

belief deviation: for instance, in the extreme case with no uncertainty, there should be no role for belief biases; whereas larger uncertainty can probably associate with
larger belief bias. $\Delta P(0 \rightarrow 20)$ and $\Delta P(20 \rightarrow 80)$ are price change between Periods 0 and 20, and between Periods 20 and 80. The price patterns may also influence beliefs. In regressions (3) and (6) stock fixed effects and subject fixed effects were added. The results are very persistent: gains in the previous period lead to more optimistic belief deviations, and losses lead to more pessimistic belief deviations. A gain increased the belief deviation by 14.16 percentage points (or 38% of a standard deviation) in Period 20, and 12.60 percentage points (or 31% of a standard deviation) in Period 80 ($p < 0.01$) according to regressions (3) and (6).

Table 1: Beliefs in Experiment 1

This table reports OLS regression results for beliefs in Experiment 1. The dependent variable is the belief deviation from mean belief; the main independent variables are Gain (Loss), i.e. a dummy variable when the subject gained (lost) on the stock in the first 20 periods; GainLoss is a continuous variable of the actual gain or loss. Control variables: Uncertainty measures uncertainty about a stock, which is a dummy variable that is 1 when the belief standard deviation on a stock is higher than the median among all stocks; $\Delta P(0 \rightarrow 20)$ and $\Delta P(20 \rightarrow 80)$ are price changes from Periods 0 to 20 and Periods 20 to 80. Regressions (1)-(4) are for Period 20 beliefs, Regressions (5)-(8) for Period 80 beliefs.

|                | Period 20 Beliefs | Period 80 Beliefs |
|----------------|-------------------|-------------------|
|                | (1)              | (2)              | (3)              | (4)              | (5)              | (6)              | (7)              | (8)              |
| Gain           | 13.49***         | 14.09***         | 14.16***         | 11.95***         | 12.53***         | 12.60***         |                   |                   |
|                | (8.79)           | (9.22)           | (3.86)           | (7.92)           | (8.34)           | (3.77)           |                   |                   |
| Loss           | -8.49***         | -9.08***         | -9.16**          | -9.28***         | -9.86***         | -9.93***         |                   |                   |
|                | (-5.58)          | (-6.05)          | (-2.44)          | (-6.37)          | (-6.80)          | (-3.10)          |                   |                   |
| GainLoss       |                  |                  |                  | 0.02***          |                  |                  | 0.02***          |                  |
|                |                  |                  |                  | (9.31)           |                  |                  | (10.73)          |                  |
| Uncertainty    | 0.26             | 1.42             | -0.00            | 0.23             | 0.60             | 0.00             |                   |                   |
|                | (0.21)           | (0.38)           | (-0.00)          | (0.17)           | (0.18)           | (0.00)           |                   |                   |
| $\Delta P(0 \rightarrow 20)$ | -0.09*** | -0.10           | -0.04            | -0.08**          | -0.10            | -0.04            |                   |                   |
|                | (-3.54)          | (-1.17)          | (-0.42)          | (-2.47)          | (-1.32)          | (-0.44)          |                   |                   |
| $\Delta P(20 \rightarrow 80)$ | 0.00             | 0.02             | -0.00            | 0.00             | 0.02             | -0.00            |                   |                   |
|                |                  |                  |                  |                  |                  |                  |                  |                  |
| Stock FE       | No               | No               | Yes              | Yes              | No               | No               | Yes              | Yes              |
| Subject FE     | No               | No               | Yes              | Yes              | No               | No               | Yes              | Yes              |
| $N$            | 3172             | 3172             | 3172             | 3172             | 3172             | 3172             | 3172             | 3172             |
| Adjusted $R^2$ | 0.07             | 0.07             | 0.06             | 0.04             | 0.06             | 0.07             | 0.06             | 0.04             |

$t$ statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
This experiment also provides suggestive evidence for the effect of the magnitude of experienced payoffs on subsequent beliefs. Figure 2 suggests that after larger gains (GAIN20) beliefs were more biased than after smaller gains (GAIN10). Additionally, columns (4) and (8) in Table 1 use the magnitude of experienced payoffs (GainLoss) as independent variable and find that belief deviations were significantly positively correlated with the magnitude of experienced payoffs. However, admittedly these are not sufficient to fully capture the shape of $q(R)$.

3.2 Experiment 2: Drawing Balls from Urns

3.2.1 Design

The main purpose of this experiment is to test H2, an important behavioural mechanism proposed in the model. That is, experienced payoffs distort processing of new information, rather than only inducing a one-time belief bias. To test this, this experiment relies on a design that can compare signal weighting in belief updating relative to the objective Bayesian benchmark. In the meantime, gains and losses are again exogenously manipulated and information-free.

The experiment consisted of a series of prediction tasks with computerised ball drawing from urns (see e.g. Camerer, 1987). There were totally 26 rounds. In each round, subjects observed ball drawings from one of two urns (A or B) with replacement. In the beginning of each round, a die roll determined which urn was used for that round (Urn A if 1-4; Urn B if 5-6). Urn A contained 6 type-P balls and 4 type-Q balls; Urn B the opposite. The two types of balls had different but similar colours, and were clearly contrasted in the instructions and in the beginning of each round.31 Subjects had all this information, observed ball drawings and guessed future draws.

In the beginning of each round, subjects received 500 ECUs cash endowment plus some additional cash or share endowments. There were 4 treatments differing in payoff structures from the first 5 ball drawings in each round: treatment P50 subjects gained (lost) 50 ECUs from each P-ball (Q-ball) drawn; treatment P100 subjects gained (lost) 100 ECUs from each P-ball (Q-ball); treatment Q100 subjects gained (lost) 100 ECUs from each Q-ball (P-ball); treatment NO subjects did not gain/lose

31 The sequences of outcomes were predetermined by using a real bingo cage with two types of balls. Subjects were allowed to inspect the bingo cage and balls after the experiment if they wanted to. The reason for calling them P balls and Q balls in the experiment was that the colours were not fixed between rounds.
from ball drawings, but received additional cash that matched with the earnings of P100 subjects in the corresponding rounds. After the first 5 draws, subjects observed 8 more draws without earning anything from the draws. Both after the 5th and the 13th draws, they predicted the 14th draw. Similar to Experiment 1, during belief elicitation subjects had no further stake in the ball drawings in order to avoid desirability bias.

Subjects were asked to guess the probability of the 14th draw being a P-ball: they needed to select a number $r^*$ on a slider representing $r \in [0, 1]$, so that they were indifferent between two lotteries: Lottery A, win 200 ECUs if the 14th draw is a P-ball, and 0 otherwise, denoted $200_P0$; Lottery B, win 200 ECUs with probability $r$, and 0 otherwise, denoted $200_r0$. Karni (2009) offers theoretical proof for the lottery method here. The instructions about belief elicitation here are similar to Experiment 1. Finally, the ending cash balance of one randomly selected round was paid, and additionally one prediction was randomly selected for the Prediction Reward. Prediction Reward was determined in a similar way as in Experiment 1: a number $Z$ was randomly drawn from a uniform distribution on $[0, 1]$; if $Z < r^*$, Lottery A was played for real; if $Z > r^*$, Lottery B ($200_Z0$) was played; if $Z = r^*$, both lotteries were equally likely played.

3.2.2 Results

The summary statistics of beliefs in this experiment are in Table A3 of Supplementary Materials. I test the treatment effects by comparing reported beliefs with the Bayesian benchmarks, and use the differences between the two elicited beliefs on each sequence to measure biased signal weighting in belief updating. Figure A2 of Supplementary Materials summarises beliefs in this experiment across treatments and compare them with the Bayesian benchmark.

I first show that Period 5 beliefs were biased by payoffs experienced in the first five draws. According to H1, gains should make subjects more optimistic than losses. Subjects in treatments P50 and P100 gained after observing sequences with more P balls, so they should be more optimistic about future gains (P ball), while after observing sequences with more Q balls, they lost and should be more pessimistic about

\[^{32}\text{Again the concept of indifference was avoided. Because the underlying proportions of P balls were 60\% and 40\%, so subjects were told that only values within these boundaries were possible. Their responses reflected that they paid attention to the instructions.}\]
future gains. By contrast, subjects in treatment Q100 gained from Q balls, so their bias should be in the opposite direction when compared with P50 and P100 subjects. Note that desirability bias should not be present here as subjects had no further gain/loss from the balls when beliefs were elicited. Moreover, simple extrapolation predicts that beliefs should be correlated with observed outcomes, but there should be no systematic difference between treatments when they observed exactly the same sequences. Therefore, neither of these alternative mechanisms generates a belief bias dependent on experienced payoffs.

To visually show this, in Figure 3 I plotted deviations of Period 5 beliefs from the Bayesian benchmark by treatments, respectively for sequences with more P balls and more Q balls. A positive deviation suggests overoptimism about P balls relative to Bayesian.

![Figure 3: Experiment 2 First Elicitation: Deviations from Bayesian](https://ssrn.com/abstract=2964289)

*Note:* This bar chart plots belief deviations from Bayesian in the first belief elicitation of Experiment 2. The vertical axis represents the deviation of reported beliefs about the 14th draw being a P ball minus Bayesian belief. Observations were grouped according to treatments, and according to whether the sequence contained more P or Q balls in the first 5 draws.

In this figure, treatment NO provides a baseline where subjects had no gain/loss from the ball drawings. Beliefs of treatment NO subjects exhibited no significant bias when there were more Q balls, but significant downward biases when there were more P balls (mean $-0.51$; one-sample two-tailed t-test compared with 0: $p < 0.001$; df:
This could reflect their base rate neglect (e.g. Kahneman and Tversky, 1973): although the proportions of P balls and Q balls were symmetric between the two urns, the chance of urn A (with more P balls than Q balls) being chosen was 2/3. Subjects did not sufficiently account for the unbalanced prior probabilities of the two urns.

The between-treatment differences reflect the effects of experienced payoffs. Consistent with the PBBD predictions, after observing the same sequence with more P balls, P100 and P50 subjects gained, while Q100 subjects lost, so P100 and P50 subjects were more optimistic about future gains (P ball) than NO subjects (P50−NO= 0.83, p < 0.001, df: 866; P100−NO= 1.33, p < 0.001, df: 852), who in turn were more optimistic about P balls than Q100 subjects (NO−Q100= 0.26, p = 0.06, df: 908). Meanwhile, after observing the same sequence with more Q balls, P100 and P50 subjects lost, while Q100 subjects gained, so Q100 subjects were more optimistic about future gains (Q ball) than NO subjects (Q100−NO= −1.70, p < 0.001, df: 778), who in turn were more optimistic than and P100 and P50 subjects though not significantly (NO−P50= −0.17, p = 0.30, df: 730; NO−P100= −0.23, p = 0.15, df: 742).

Admittedly, some observations were not completely consistent with the PBBD prediction when compared with the Bayesian benchmark. Specifically, if PBBD is the only belief bias at work, after seeing a sequence with more P balls, Q100 subjects lost, so they should be more pessimistic about future gains, or more optimistic about future P balls than Bayesian. And similarly, after seeing a sequence with more Q balls, P100 and P50 subjects should be more pessimistic about future P balls than Bayesian. The non-presence of such patterns could be attributable to some residual desirability bias, which made subjects who benefited from P (or Q) balls more optimistic about P (or Q) balls in general. Given that PBBD and desirability bias work in opposite directions in these instances, it was possible that PBBD cancelled out some of the desirability bias but did not completely drive beliefs to the other direction. P50 and P100 subjects significantly overestimated P (Q) balls compared with treatment NO by 0.48, and Q100 subjects’ mean deviation was −1.10 (significantly different at p < 0.001; df: 2312).

Note that there are at least two key differences between desirability bias and PBBD. First, desirability bias has no prediction when the decision maker no longer has

---

33The inconsistency with the PBBD prediction and the potential residual desirability bias can also be seen from the summary in Figure A2 of Supplementary Materials.
a stake, whereas PBBD only requires prior payoff experience. Second, gains are always more desirable than losses and should be overestimated according to desirability bias, whereas PBBD predicts that both gains and losses could be overestimated depending on prior experience. It is hard to fully avoid the residual desirability bias, or to disentangle it from PBBD when their predictions overlap. Hence in what follows, I present results that are not contaminated by the potential desirability bias. PBBD predicts that after gains, subjects should be more optimistic about future gains than after losses. Within treatment, P100 subjects were significantly biased to overestimate P balls after they gained than after they lost ($p < 0.001; df: 726$); Q100 subjects were significantly biased to overestimate Q balls after they gained than after they lost ($p < 0.001; df: 830$).

Another interesting observation not contaminated by the potential desirability bias is related to the magnitude (rather than sign) of experienced payoffs. If the magnitude of payoffs is positively correlated with the size of belief bias, then P100 subjects should overestimate P balls more than P50 subjects after seeing sequences with more P balls. This is because the only difference between these treatments is that P100 subjects gained more on such sequences than P50 subjects. I find that this is indeed the case. For sequences with more P balls, P100 subjects significantly overestimated future P balls relative to P50 subjects (Difference= 0.50, $p < 0.01$, df: 796). This is suggestive evidence for a positive correlation between the magnitude of experienced payoffs and the size of belief bias. However, this is not sufficient to fully establish the relationship, because the experiment did not have enough variation in payoffs between subjects who observe the same ball drawing sequence. Future research on this is needed.

Therefore, gains made subjects more optimistic about the good state. Interestingly, P50 and P100 beliefs were not significantly different when they saw more Q balls. This suggests a potential asymmetry: payoff-based belief distortion might be stronger after gains than after losses. This could reflect an optimism bias, or the ostrich effect (e.g. Galai and Sade, 2006; Karlsson et al., 2009), as subjects were

---

34 But note that this comparison is not holding information constant. Subjects in the same treatment who gained and those who lost saw different sequences of ball drawings.

35 I also calculated the average deviation from Bayesian at the individual level: no subject was consistently Bayesian, in that all subjects’ deviations from Bayesian were significantly different from 0 at the 5% level. And there was considerable heterogeneity in the magnitude of their biases. See distribution of beliefs at individual level in Figure A3.
probably more likely to update beliefs after gains when they had good prior beliefs.

Table 2 reports OLS regression results that confirm the treatment effect, using belief deviations from Bayesian as the dependent variable. Independent variables include treatment dummies ($P_{50}$, $P_{100}$, $Q_{100}$) and controls for the observed ball $P$ balls in the first 5 and 13 periods respectively. $P_{100}$ subjects overestimated $P$-balls and $Q_{100}$ subjects underestimated. The coefficient on $P_{100}$ was significantly larger than that on $P_{50}$ when they gained (Regression (1) Wald test: $p = 0.03$), but not when they lost (Regression (3) Wald test: $p = 0.37$). Table 2 also shows that in Period 13 both $P_{100}$ and $Q_{100}$ subjects overestimated $P$ ball for sequences with more $P$ balls in the first five draws, and underestimated $Q$ ball for sequences with more $Q$ balls, but were respectively more biased towards the ball from which they gained: in Regression (2) $P_{50}$ and $P_{100}$ subjects were more optimistic about $P$ balls than $Q_{100}$ subjects (Wald test: $p < 0.01$); Regression (4) $Q_{100}$ subjects were more optimistic about $Q$ balls than $P_{50}$ and $P_{100}$ subjects (Wald test: $p < 0.01$). The regressions included as independent variables the number of $P$ balls in the first 5 and 13 periods respectively, controlling for belief dependence on observed outcomes: There was belief in mean reversion: more $P$ balls in sequence led to lower belief of $P$ balls in the 14th draw. But the treatment effects were significant even when these controls were present.

The results so far could suggest that there was an asymmetry in beliefs following gains and losses. Experienced payoffs might just cause a one-time shift in beliefs. However, the following results show that this was not the case, in that experienced payoffs also led to biases in processing new information. This supports the behavioural mechanism proposed in the model. To test H2 regarding belief updating, I investigate how beliefs were updated using the 8 new ball drawings between Period 5 and Period 13. With these new signals, a Bayesian agent should update to increase belief about $P$ ball if there were more $P$ balls in the new signals, and the updating magnitude should be positively correlated with the strength of new evidence, i.e. the number of $P$ balls relative to that of $Q$ balls. Since the proportion of balls in each urn was fixed given $\theta$, it can be shown that the change in Bayesian log likelihood ratio from Period 5 to Period 13 ($\Delta \ln \Lambda^{Bayesian}$) should be directly proportional to the strength of new evidence supporting $P$ balls ($\Delta(p - q)$, i.e. the difference between the number of $P$ balls and that of $Q$ balls in the new signals): $\Delta \ln \Lambda^{Bayesian} = 0.4055 \Delta(p - q)$.$^{36}$

$^{36}$See Section 3 of Supplementary Materials for a short proof.
Table 2: Treatment Effects in Experiment 2

This table reports OLS regression results for a test of treatment effects in Experiment 2. The dependent variable is the deviation of elicited beliefs from Bayesian. All regressions controlled for sequence fixed effects and subject fixed effects. $P_{50}$, $P_{100}$, $Q_{100}$ are dummy variables for the respective treatments. $P(1 - 5)$ and $P(1 - 13)$ are the number of P balls in the first 5 periods and 13 periods, respectively. Regressions (1) (2) are for sequences with more P balls by the 5th draw; (3) (4) for sequences with more Q balls by the 5th draw. (1) (3) for Period 5 beliefs; (2) (4) for Period 13 beliefs.

|                | More P Balls | More Q Balls |
|----------------|--------------|--------------|
|                | Period 5     | Period 13    | Period 5     | Period 13    |
| $P_{50}$       |             |              |              |              |
|                | 0.83***      | 3.18***      | 0.17         | -2.55***     |
|                | (2.98)       | (5.58)       | (0.90)       | (-5.14)      |
| $P_{100}$      |             |              |              |              |
|                | 1.33***      | 3.88***      | 0.23*        | -2.75***     |
|                | (4.68)       | (7.17)       | (1.73)       | (-5.10)      |
| $Q_{100}$      |             |              |              |              |
|                | -0.26        | 1.19**       | -1.70***     | -4.76***     |
|                | (-1.03)      | (-2.10)      | (-12.76)     | (9.28)       |
| $P(1 - 5)$     |             |              |              |              |
|                | -2.62***     |             | -3.91***     |              |
|                | (-16.16)     |              | (-39.66)     |              |
| $P(1 - 13)$    |             | -2.38***     | -2.81***     |              |
|                | (-6.56)      |              | (-3.44)      |              |
| Constant       | 8.48***      | 14.14***     | 6.32***      | 15.57***     |
|                | (15.12)      | (6.17)       | (31.40)      | (2.99)       |
| Sequence FE    | Yes          | Yes          | Yes          | Yes          |
| Subject FE     | Yes          | Yes          | Yes          | Yes          |
| $N$            | 1708         | 1708         | 1464         | 1464         |
| $R^2$          | 0.46         | 0.27         | 0.80         | 0.24         |

$t$ statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Equivalently, $\Delta \ln \Lambda_{Bayesian} = 0.4055\Delta p - 0.4055\Delta q$, where $\Delta p$ ($\Delta q$) represents the number of P (Q) balls in the new signals; that is, a Bayesian agent should place equal weight on new P and Q balls drawn, and the weight is exactly 0.4055.

By contrast, according to Equations 14 and 15, a PBBD agent underweights signals inconsistent with experienced payoffs (relative to those that are consistent). I test this using the regression specification in Equation 17,

$$\Delta \ln \Lambda_i^k = \alpha_1 \Delta p^k + \alpha_2 \Delta q^k + \alpha_3 |G L_i^k| \times \Delta p^k + \alpha_4 |G L_i^k| \times \Delta q^k + \varepsilon_i^k, \quad (17)$$
where $\Delta \ln \Lambda_i^k$ is the change in subject $i$’s log likelihood ratio on sequence $k$, calculated using the elicited beliefs; $\Delta p^k$ ($\Delta q^k$) is the number of P (Q) balls in draws between Period 5 and Period 13 in sequence $k$; $|GL_i^k|$ is the absolute value of gain or loss of subject $i$ from sequence $k$. The regression contains no intercept so that I can obtain estimates of the weights placed on P and Q signals directly without suffering from the collinearity problem. The null hypothesis is that $\alpha_1 = -\alpha_2 = 0.4055$ and $\alpha_3 = \alpha_4 = 0$ according to Bayesian updating. If experienced payoffs lead to biased weighting of signals, then $\alpha_3$ and $\alpha_4$ should be significantly different from zero, with their signs depending on treatment and gain/loss. For example, in P50 or P100 treatments, after a gain from P balls, subjects should overweight new P balls relative to Q balls, thus $\alpha_3 \geq 0$ and $\alpha_4 \geq 0$; after a loss, Q balls should be overweighted, thus $\alpha_3 \leq 0$ and $\alpha_4 \leq 0$. The opposite should be true for Q100.

Table 3 presents regression results from Equation 17, with standard errors clustered at the subjects level and sequence-fixed effects. Using data of all treatments, Regression (1) shows that belief updating was not Bayesian: we can confidently reject $\alpha_1 = 0.4055$ and $\alpha_2 = -0.4055$ ($p < 0.001$); instead, the weights were significantly less than 0.4055, suggesting an overall conservatism bias, and slightly larger weights on Q than P balls (Wald test $p < 0.001$). This again might reflect the base rate neglect. Regression (2) use data from the baseline group NO: subjects placed larger weights on both P and Q balls, but still did not update sufficiently compared with Bayesian (Wald test for $\alpha_1 = 0.4055$: $p = 0.09$; for $\alpha_2 = -0.4055$: $p < 0.001$). Regressions (3) to (8) reveal that in general, if there had been no gain/loss ($|GL| = 0$), subjects would have underweighted new signals, and that P50 and P100 subjects were inclined to overweight P balls (Wald test for $\alpha_1 = \alpha_2$, $p < 0.01$, except in Regression (3)), while Q100 subjects were inclined to overweight Q balls (Wald test for $\alpha_1 = \alpha_2$, $p = 0.026$ in regression (7) and $p < 0.01$ in regression (8)). For P50/P100 treatments, after gains $\alpha_4$ was significantly positive, and after losses $\alpha_3$ was significantly negative. These together suggest that larger gains (losses) made P50 and P100 subjects attach relatively larger weight on P (Q) balls. The opposite was true for Q100 subjects: $\alpha_3$ was significantly negative after they gained, and both $\alpha_3$ and $\alpha_4$ were significantly positive after they lost. More interestingly, in regressions (9) and (10) I test whether P100 subjects weighted each additional P (Q) ball differently than P50 subjects. This was done by interacting the number of P or Q balls with

\footnote{Note that $|GL| = 0$ was only hypothetical because only subjects in NO treatment had 0 gain/loss.}
a dummy variable for P100 treatment. The result suggests that, after gains P100 subjects tended to more severely underweight each Q ball than P50 subjects, but the pattern did not emerge after losses. This is consistent with the asymmetry observed in Figure 3, i.e. the payoff-based belief distortion was stronger after gains than after losses.

In order to visualise the effects of experienced payoffs on signal weighting, Figure 4 illustrates the results from Table 3. Panels A and B show weights placed on each new P ball and Q ball respectively. The marginal effect of each additional P ball on the log likelihood was calculated as \( \alpha_1 + \alpha_3|GL| \), and that of each additional Q ball was \( \alpha_2 + \alpha_4|GL| \). The Bayesian benchmark (0.4055 in Panel A and -0.4055 in Panel B) was clearly marked by a horizontal line in each panel. The NO baseline result indicates a tendency of insufficient updating, weighting each ball (P or Q) less than 0.4055. PBBD drove the weights towards Bayesian on balls that were consistent with prior gains, i.e. P (Q) balls for P50/P100 (Q100) subjects; and it drove the weights further away from Bayesian on balls that were inconsistent with prior gains, i.e. Q (P) balls for P50/P100 (Q100) subjects. In the former case, although the payoff-induced bias drove conservative subjects towards Bayesian, a closer examination reveals that they over-adjusted, placing a weight even larger than 0.4055 on these signals.
Table 3: Belief Updating in Experiment 2

This table reports OLS regressions (no intercept) that test biases in belief updating in Experiment 2. The dependent variable is the change of log likelihood ratio between Period 5 and Period 13 beliefs. Independent variables: $\Delta p$ and $\Delta q$ are the number of P and Q balls from the 6th to the 13th draws; $|GL|$ is the absolute value of gain or loss from each sequence. Regression (1) is for all treatments, (2) for NO; (3) (4) for P50, (5) (6) for P100, (7) (8) for Q100; (3) (5) (7) for beliefs following gains; (4) (6) (8) for beliefs following losses. Regressions (9) (10) use both P50 and P100 treatments, with $\Delta p$ and $\Delta q$ interacting with a dummy variable for P100. All regressions include sequence fixed effects and subject fixed effects.

|                | All  | NO   | P50  | P100 | Q100  | P50&P100 |
|----------------|------|------|------|------|-------|----------|
|                | (1)  | (2)  | (3)  | (4)  | (5)   | (6)      |
| $\Delta p$ ($\alpha_1$) | 0.17*** | 0.31*** | 0.32*  | 0.22*** | 0.40*** | 0.36***  | 0.31*** | -0.00 | 0.43*** | 0.00 |
|                | (7.00) | (5.39) | (1.94) | (6.10) | (2.84) | (4.25)  | (4.42) | (-0.01) | (6.56) | (0.23) |
| $\Delta q$ ($\alpha_2$) | -0.19*** | -0.29*** | -0.32*** | -0.18*** | -0.33*** | -0.26** | -0.34*** | -0.52*** | -0.18*** | -0.22*** |
|                | (-10.06) | (-5.65) | (-6.22) | (-3.26) | (-4.99) | (-2.38) | (-4.20) | (-7.91) | (-7.14) | (-4.33) |
| $|GL| \times \Delta p$ ($\alpha_3$) | 0.00 | -0.0015*** | 0.00 | -0.0015*** | -0.0011** | 0.0019*** |
|                | (0.25) | (-3.35) | (-0.55) | (-3.39) | (-2.36) | (3.11)  |
| $|GL| \times \Delta q$ ($\alpha_4$) | 0.0021*** | -0.00 | 0.0016*** | -0.00 | -0.00 | 0.0019*** |
|                | (3.65) | (-1.45) | (3.46) | (-0.89) | (-0.86) | (6.42)  |

$t$ statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Figure 4: Experiment 2 Updating

Note: This figure plots signal weighting in Experiment 2: Panel A shows weights on P balls, and Panel B shows weights on Q balls from the 6th to the 13th draws. The weights and their 95% confidence intervals were calculated from the regression results in Table 3. ‘Sm’ and ‘Lg’ on the horizontal axis represent small and large gain/loss, which in P100 and Q100 treatments mean 150 and 300 ECUs, and in P50 treatment mean 50 and 100 ECUs; ‘0’ represents no gain/loss, which was calculated from the regression results hypothetically.
4 Robustness Checks and Alternative Explanations

4.1 Robustness Checks

4.1.1 Experiment 1

In order to check the robustness of the results in Experiment 1, I use several alternative specifications in Table A5 of Supplementary Materials. All regressions in this table use the same dependent variable as in Table 1, and contain stock fixed effect, subject fixed effect and control variables. Three different alternative specifications were tested. First, due to the repetitions in this experiment, the results could be influenced by players’ experience or learning throughout the session. In regressions (1) to (4), I separately run regressions for the first 10 periods and the last 10 periods in each session. Second, it is possible that the results are driven by specific stocks. In the experiment, subjects saw stocks in random orders, and there were 24 unique stocks. In regressions (5) to (8), I split the stocks into two 12-stock subsamples and run the same regressions. Third, I consider alternative specifications to account for experienced gain or loss. In regressions (9) to (12), instead of using dummy variables for Gain and Loss, I use a continuous variable, namely variable $GainLoss$, which measures the actual gain/loss from a stock in the first 20 periods. Regressions (9) and (10) use a linear specification of $GainLoss$, whereas regressions (11) and (12) include the quadratic form of $GainLoss$. All regressions produce results that are consistent with findings in Table 1. The effect of previous payoffs on subsequent beliefs was robust in Experiment 1.\textsuperscript{38}

4.1.2 Experiment 2

Similar to the above robustness checks of Experiment 1, here I also use alternative specifications to test robustness of results in Table 2 and Table 3, respectively. I tried two different ways to split the sample. First, I split the periods into the first half and the last half of the experiment, to rule out learning effect. Second, I split the sequences into two equally-sized subsamples. These results are reported in Table A6 and Table A7 of Supplementary Materials. These changed the coefficients slightly but did not change the main results. The findings were robust.

\textsuperscript{38}In the quadratic form, within meaningful ranges of the the value of $GainLoss$, optimism was increasing in gains.
4.2 Alternative Explanations

4.2.1 Change of risk attitudes

It is possible that gains and losses influenced subjects’ risk attitudes, and the experimental results on beliefs were driven by differences in risk attitudes. The effect of previous payoffs on subsequent risk preferences was discussed in Thaler and Johnson (1990) in their house money effect, and then met with conflicting evidence. Recently Imas (2016) finds that realised losses increase risk aversion, but paper losses decrease it. In my experiment, all gains and losses were realised before beliefs were elicited, which avoided the complication of realised versus paper gain/loss. However, I still use several ways to address the concern about risk attitudes.

First, I elicited subjects’ risk attitudes in two ways at the end of the experiment before announcing earnings, such that these risk attitude measures are not affected by learning and experiencing gains/losses between rounds. One elicitation was a survey measure: subjects were asked to rate their general risk-taking on a scale of 1-10 (variable RiskSurvey), with larger numbers meaning more risk-taking. The other elicitation was using the multiple price list method (Holt and Laury, 2002) (variable RiskChoice), with larger numbers meaning more risk-taking, and numbers represent the number of risky options chosen after switching from safe to risky option on the list. If gains and losses influenced risk attitudes, these measures should be correlated with earnings from the experiment. As final earnings (from a randomly chosen round) were not announced when the risk attitude was measured, I use the total earnings here. Total earnings had a Spearman rank correlation of $-0.07$ ($p = 0.47$) with the survey measure of risk attitude, and $-0.02$ ($p = 0.83$) with the choice list measure. Therefore, earnings in the experiment were not significantly correlated with risk attitudes.

Furthermore, I check whether subjects who were more risk-taking reported more optimistic beliefs, which if true would undermine my results. From both experiments, I calculated the average belief deviation for each subject, and its correlation with elicited risk attitudes. The results in Table 4 do not show significant correlation. Additionally, Table A8 of Supplementary Materials reports regression results that control for risk attitudes. Only the survey risk attitude measure had some significant effect. But even after controlling for subjects’ risk attitudes, the effects of gains and losses on subsequent beliefs were still robust, and quantitatively similar. Therefore, this also alleviates the concern that changing risk attitudes may contaminate the
Table 4: Correlation between Risk Attitudes and Beliefs

This table shows the Spearman correlation between risk attitudes and belief deviations in Experiment 1 and 2, with p values in parentheses. Belief deviations in Experiment 1 are deviations from mean belief on each stock, and in Experiment 2 are deviations from the Bayesian benchmark.

|                | Experiment 1 | Experiment 2 |
|----------------|--------------|--------------|
| RiskSurvey     | 0.14         | 0.05         |
|                | (0.15)       | (0.58)       |
| RiskChoice     | -0.15        | 0.04         |
|                | (0.10)       | (0.70)       |

4.2.2 Hot-hand and gambler’s fallacies

PBBD is related to the hot-hand and gambler’s fallacies. The hot-hand fallacy is the belief that a recently observed streak of like outcomes will continue, whereas the gambler’s fallacy is the opposite. By contrast, PBBD is based on payoffs, whereas not all past outcomes are associated with payoffs. Therefore, PBBD is proposing that outcomes may receive different weights in belief formation depending on whether they generated positive or negative payoffs. For instance, an investor may observe the full history of a stock’s price, but only held the stock and obtained payoffs during some periods. In such cases, PBBD could generate different predictions than observation-based belief biases, for instance, when payoffs are obtained from an unrepresentative subsample of historical outcomes. A stock might have generated positive overall historical return, but an unlucky investor might happen to own the stock during some bad periods.

In order to check whether PBBD is different from observation-based hot hand/gambler’s fallacies, I included variables that account for observed outcomes in regression analyses in both experiments. In Table 1 for Experiment 1, variables $DP(0 - 20)$ and $DP(20 - 80)$ are the price changes between Periods 0 and 20, and between Periods

---

39 Also for the following reasons, the results documented in this paper are not likely contaminated by risk attitudes: the stakes were small; finally only one randomly selected round was paid; and the belief elicitation method was robust to different risk attitudes.
20 and 80. The results indicate that there was not so much hot hand or gambler’s fallacy in Period 20 beliefs, and there was some hot hand fallacy in Period 80 beliefs: subjects believed stocks with larger historical price increase would have higher future price. However, adding these control variables did not change the effects of payoffs on beliefs. In Table 2 for Experiment 2, I added the number of P balls in the first 5 draws, and in the first 13 draws. The negative sign on these coefficients suggests the gambler’s fallacy: if subjects observed more P balls in the sequence, they tended to predict lower probability for future P balls. Again, adding these variables did not change the effects of payoffs on beliefs.

Section 4 of Supplementary Materials contains an additional check of whether PBBD generates different results than the observation-based fallacies. All of these results indicate that PBBD was robust even when hot hand or gambler’s fallacy was present.

4.2.3 Confirmation bias

Confirmation biased agents are more likely to believe in or to actively seek for information consistent with their prior belief. Rabin and Schrag (1999) model this as a misperception of signals inconsistent with one’s prior belief, thus overweighting confirming evidence. However, confirmation bias does not address where the prior belief comes from. By contrast, PBBD proposes that prior beliefs are actually shaped by experienced payoffs, and subsequently there is a confirmation bias with respect to these payoffs. PBBD can generate different predictions than simple confirmation bias. For instance, it’s possible that one may have experienced gains from a stock but hold a pessimistic view about its future performance. If in my experiment, positive prior beliefs had always coincided with positive payoff experiences, then it would have been hard to disentangle the effect of payoffs from that of simple confirmation bias. Fortunately, this was not the case.

In order to check whether confirmation bias was a confound, I first show how belief updating depended on prior beliefs or experienced payoffs in Figure A5. In Experiment 2, I elicited beliefs in both Periods 5 and 13. Treating beliefs in Period 5 as the prior belief, I calculated the log likelihood ratio \( \ln \Lambda_{i,5}^k \) for each subject \( i \) on sequence

\[ \text{40} \]
The hot hand fallacy in Experiment 1 and gambler’s fallacy in Experiment 2 are not contradictory, but actually consistent with the literature (e.g. Rabin, 2002): when people know the underlying process, they are more likely to exhibit gambler’s fallacy; but when they do not know, they are more likely to exhibit the hot hand fallacy.
$k$. When $\ln \Lambda_{i,5}^k > 0$, the subject’s Period 5 prior belief favors urn A, and thus each new P ball is confirming with this prior belief; when $\ln \Lambda_{i,5}^k < 0$, the subject’s prior belief favors urn B; when both are 0, the subject is indifferent. In Panel A, I summarise $\Delta \ln \Lambda_i^k$, subject $i$’s update of log likelihood ratio on sequence $k$, for $\ln \Lambda_{i,5}^k < 0$ (Q ball is confirming with prior) and $\ln \Lambda_{i,5}^k > 0$ (P ball is confirming with prior) respectively, across different treatments. A positive value of $\Delta \ln \Lambda_i^k$ means the subject updated more towards urn A. Treatment NO exhibited no significant confirmation bias (and the sign was even contradictory with confirmation bias), and all other treatments exhibited some confirmation bias in the sense that subjects updated more towards urn A if their prior belief favored urn A, although this was completely driven by the bias on the $\ln \Lambda_{i,5}^k < 0$ side. Panels B to D show each treatment separately, for $\ln \Lambda_{i,5}^k < 0$ and $\ln \Lambda_{i,5}^k > 0$, gains and losses respectively. When P50 and P100 subjects gained, they were more likely to update towards urn A, and when they lost, they were more likely to update towards urn B. The opposite was true for Treatment Q100, although only significant for gains. However, there was no significant confirmation bias when we compare between different prior beliefs given experienced payoffs.41

I further checked the effects of confirmation bias and payoff experiences in belief updating by rerunning the regressions in Table 3 with variables that control for confirmation bias. These variables are dummy variables respectively for $\ln \Lambda_{i,5}^k < 0$ and $\ln \Lambda_{i,5}^k > 0$, and I interacted them with variables $\Delta p^k$ and $\Delta q^k$, such that the coefficients of the interaction terms indicate the effects of observing signals confirming with one’s prior belief. The results reported in Table A8 show that controlling for confirmation bias did not render the effect of payoffs insignificant.

5 Discussion

This paper proposes a model for the effects of experienced payoffs on subsequent beliefs in repeated decision-making under uncertainty. The model combines a reinforcement learning component with belief updating. Simply repeating previously successful actions is an intuitive and backward-looking heuristic rule, so reinforcement learning could correspond to System 1 processing (e.g. Kahneman, 2003). However, contemplating future probabilities and updating beliefs could entail the use of Sys-

---

41Note that the missing observations in Figure A5 was due to the fact that no subject’s prior belief favored urn B after observing more P balls.
tem 2. The mechanism proposed here suggests that the output of System 1 (i.e. reinforcements) may exert some externality on System 2 (i.e. belief updating).

The findings in this paper have important implications for financial decision-making, job search process, adoption of new technology, etc. The following discussion will focus on financial applications. If investors fall under the spell of PBBD, this could give rise to unjustified optimism (or pessimism) on specific assets, asset classes or trading strategies, correlated with personal payoff experiences. For instance, this could explain why young fund managers who invested in tech stocks during the tech bubble were more optimistic and less likely to believe there was a bubble than those (mostly old fund managers) who did not have stakes involved (e.g. Greenwood and Nagel, 2009; Bernile et al., 2017). Having different personal payoff experiences could be an important way to generate heterogeneous beliefs, and thus trades, without asymmetric information. What traders fail to realise is that their personal payoffs may carry little to no useful information for the market’s future performance. Other non-payoff-based theories cannot directly generate such a prediction.

PBBD can also be a driving force during the formation of bubbles. This is because investors gain from the bubble asset during the bubble formation period, which makes them overoptimistic, and even underweight signals that contradict their positive experience. Therefore, PBBD agents would be overly optimistic during bubble formation and over-invest in the bubble asset; when the bubble bursts, they suffer a lot, become overly pessimistic and under-invest. Barberis (2013) argues that extrapolation is an important belief-based mechanism that contributes to asset price bubbles. And this should be generalised to extrapolation of payoffs, rather than just observations.

PBBD interestingly interacts with other established behavioural biases. Benartzi and Thaler (1995) use myopic loss aversion to explain the equity premium puzzle. Short-sighted portfolio evaluation makes investors feel their gains/losses more frequently, which could lead to larger belief biases. Moreover, investors also suffer from the ‘ostrich effect’, looking at their portfolios more often when they have good prior beliefs (Karlsson et al., 2009). If so, they observe and experience gains more often during bullish markets, biasing their beliefs more upwards.

PBBD also has important policy implications. Firstly, it informs household finance, in terms of the presentation of information to individual investors.42 PBBD

42Kaufmann et al. (2013) and Bradbury et al. (2014) also suggest how experiential information
agents care too much about experienced payoffs and overweight information consistent with own payoffs. To reduce this bias, information presented to investors (e.g. when they access their brokerage accounts) should put less emphasis on payoffs, but direct more attention to stock or market news, especially highlighting information that contradicts experienced payoffs. Compared with other belief biases, the fact that PBBD relies on observable payoffs makes its implied interventions easier to implement. Secondly, PBBD implies that continued gains/losses are dangerous for fund managers or individuals in management positions in general. Therefore, it might be beneficial to form a team of mixed-experience managers, or to rotate the person in top management role.

6 Conclusion

Experienced payoffs are important in sequential decision-making. This is usually studied by psychologists and economists in reinforcement learning models. However, little is known about how experienced payoffs shape beliefs. This paper models a belief distortion based on experienced payoffs: experienced gains make an agent underweight bad news relative to good new; losses do the opposite. I conducted two experiments and found supportive evidence for both the predicted belief bias and the behavioural mechanism. These experiments provided a robust test of the effects of experienced payoffs, as experienced payoffs were information-free from a Bayesian perspective.

Experimental subjects were more optimistic (pessimistic) after previous gains (losses), relative to those who had no payoff but merely observed the same information. They also overweighted new signals consistent with prior payoff experiences when updating beliefs. The model has significant implications for finance and other areas, and for practical policy interventions.

A natural direction of future research is to investigate the potential effects of PBBD on asset pricing, and the dynamic interactions between PBBD agents and rational agents through market booms and busts. Another potential avenue of future research would be to characterise the function $q(R)$, and to see how it depends on the values of reinforcement, strength of prior beliefs, and domain of gain or loss. It would also be interesting to investigate how the magnitude of PBBD varies with the level of subjective uncertainty, stake sizes, etc.
Moreover, some of the experimental results were not completely consistent with PBBD. This suggests the coexistence of multiple belief patterns, including PBBD, desirability bias and some potential hot-hand and gambler’s fallacies. This calls for a parsimonious model that can generate all of these empirical patterns, which will be tremendously beneficial for the field.

*Department of Finance, School of Business and Economics, Maastricht University & Nuffield College, University of Oxford.*
References

Abdellaoui, M., Baillon, A., Placido, L. and Wakker, P.P. (2011). ‘The rich domain of uncertainty: Source functions and their experimental implementation’, American Economic Review, vol. 101(2), pp. 695–723.

Anagol, S., Balasubramaniam, V. and Ramadorai, T. (2019). ‘Learning from noise: Evidence from india’s ipo lotteries’, https://ssrn.com/abstract=2568748 (last accessed: 12 July 2019).

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Vega, C. (2007). ‘Real-time price discovery in global stock, bond and foreign exchange markets’, Journal of International Economics, vol. 73(2), pp. 251–277.

Baillon, A. (2008). ‘Eliciting subjective probabilities through exchangeable events: An advantage and a limitation’, Decision Analysis, vol. 5(2), pp. 76–87.

Barberis, N.C. (2013). ‘Psychology and the financial crisis of 2007–2008’, in (M. Haliassos, ed.), Financial Innovation: Too Much Or Too Little?, pp. 15–28, MIT Press: Cambridge.

Benartzi, S. and Thaler, R.H. (1995). ‘Myopic loss aversion and the equity premium puzzle’, Quarterly Journal of Economics, vol. 110(1), pp. 73–92.

Bernile, G., Bhagwat, V. and Rau, P.R. (2017). ‘What doesn’t kill you will only make you more risk-loving: Early-life disasters and ceo behavior’, Journal of Finance, vol. 72(1), pp. 167–206.

Bradbury, M.A., Hens, T. and Zeisberger, S. (2014). ‘Improving investment decisions with simulated experience’, Review of Finance, vol. 19(3), pp. 1019–1052.

Breiter, H.C., Aharon, I., Kahneman, D., Dale, A. and Shizgal, P. (2001). ‘Functional imaging of neural responses to expectancy and experience of monetary gains and losses’, Neuron, vol. 30(2), pp. 619–639.

Brunnermeier, M.K. and Parker, J.A. (2005). ‘Optimal expectations’, American Economic Review, vol. 95(4), pp. 1092–1118.
Bucher-Koenen, T. and Ziegelmeyer, M. (2013). ‘Once burned, twice shy? financial literacy and wealth losses during the financial crisis’, Review of Finance, vol. 18(6), pp. 1–32.

Camerer, C. and Ho, T. (1999). ‘Experience-weighted attraction learning in normal form games’, Econometrica, vol. 67(4), pp. 827–874.

Camerer, C.F. (1987). ‘Do biases in probability judgment matter in markets? experimental evidence’, American Economic Review, vol. 77(5), pp. 981–997.

Chiang, Y.M., Hirshleifer, D., Qian, Y. and Sherman, A.E. (2011). ‘Do investors learn from experience? evidence from frequent ipo investors’, Review of Financial Studies, vol. 24(5), pp. 1560–1589.

Choi, J.J., Laibson, D., Madrian, B.C. and Metrick, A. (2009). ‘Reinforcement learning and savings behavior’, Journal of Finance, vol. 64(6), pp. 2515–2534.

Chua, P., Krams, M., Toni, I., Passingham, R. and Dolan, R. (1999). ‘A functional anatomy of anticipatory anxiety’, Neuroimage, vol. 9(6), pp. 563–571.

Costa-Gomes, M.A., Huck, S. and Weizsäcker, G. (2014). ‘Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect’, Games and Economic Behavior, vol. 88, pp. 298–309.

Erev, I., Ert, E., Roth, A.E., Haruvy, E., Herzog, S.M., Hau, R., Hertwig, R., Stewart, T., West, R. and Lebiere, C. (2010). ‘A choice prediction competition: Choices from experience and from description’, Journal of Behavioral Decision Making, vol. 23(1), pp. 15–47.

Erev, I. and Roth, A.E. (1998). ‘Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria’, American Economic Review, vol. 88(4), pp. 848–881.

Foster, D.P. and Young, H.P. (2006). ‘Regret testing: Learning to play nash equilibrium without knowing you have an opponent’, Theoretical Economics, vol. 1(3), pp. 341–367.

Galai, D. and Sade, O. (2006). ‘The “ostrich effect” and the relationship between the liquidity and the yields of financial assets’, Journal of Business, vol. 79(5), pp. 2741–2759.
Gallagher, J. (2014). ‘Learning about an infrequent event: evidence from flood insurance take-up in the United States’, *American Economic Journal: Applied Economics*, vol. 6(3), pp. 206–233.

Gervais, S. and Odean, T. (2001). ‘Learning to be overconfident’, *Review of Financial Studies*, vol. 14(1), pp. 1–27.

Giannetti, M. and Wang, T.Y. (2016). ‘Corporate scandals and household stock market participation’, *Journal of Finance*, vol. 71(6), pp. 2591–2636.

Gilovich, T., Vallone, R. and Tversky, A. (1985). ‘The hot hand in basketball: On the misperception of random sequences’, *Cognitive Psychology*, vol. 17(3), pp. 295–314.

Greenwood, R. and Nagel, S. (2009). ‘Inexperienced investors and bubbles’, *Journal of Financial Economics*, vol. 93(2), pp. 239–258.

Greiner, B. (2015). ‘Subject pool recruitment procedures: organizing experiments with orsee’, *Journal of the Economic Science Association*, vol. 1(1), pp. 114–125.

Guiso, L., Sapienza, P. and Zingales, L. (2013). ‘Time varying risk aversion’, National Bureau of Economic Research.

Hanaoka, C., Shigeoka, H. and Watanabe, Y. (2018). ‘Do risk preferences change? evidence from the great east Japan earthquake’, *American Economic Journal: Applied Economics*, vol. 10(2), pp. 298–330.

Helweg-Larsen, M. (1999). ‘(the lack of) optimistic biases in response to the 1994 northridge earthquake: The role of personal experience’, *Basic and Applied Social Psychology*, vol. 21(2), pp. 119–129.

Hoffmann, A.O. and Post, T. (2015). ‘How return and risk experiences shape investor beliefs and preferences’, *Accounting and Finance*.

Holt, C.A. and Laury, S.K. (2002). ‘Risk aversion and incentive effects’, *American Economic Review*, vol. 92(5), pp. 1644–1655.

Imas, A. (2016). ‘The realization effect: Risk-taking after realized versus paper losses’, *American Economic Review*, vol. 106(8), pp. 2086–2109.
Kahneman, D. (2003). ‘Maps of bounded rationality: Psychology for behavioral economics’, American Economic Review, vol. 93(5), pp. 1449–1475.

Kahneman, D. and Tversky, A. (1973). ‘On the psychology of prediction’, Psychological Review, vol. 80(4), pp. 237–251.

Karlsson, N., Loewenstein, G. and Seppi, D. (2009). ‘The ostrich effect: Selective attention to information’, Journal of Risk and Uncertainty, vol. 38(2), pp. 95–115.

Karni, E. (2009). ‘A mechanism for eliciting probabilities’, Econometrica, vol. 77(2), pp. 603–606.

Kaufmann, C., Weber, M. and Haisley, E. (2013). ‘The role of experience sampling and graphical displays on one’s investment risk appetite’, Management Science, vol. 59(2), pp. 323–340.

Kaustia, M. and Knüpfer, S. (2008). ‘Do investors overweight personal experience? evidence from ipo subscriptions’, Journal of Finance, vol. 63(6), pp. 2679–2702.

Knüpfer, S., Rantapuska, E.H. and Sarvimäki, M. (2017). ‘Formative experience and portfolio choice: Evidence from the finnish great depression’, Journal of Finance, vol. 72(1), pp. 133–166.

Knutson, B. and Bossaerts, P. (2007). ‘Neural antecedents of financial decisions’, Journal of Neuroscience, vol. 27(31), pp. 8174–8177.

Koudijs, P. and Voth, H.J. (2016). ‘Leverage and beliefs: personal experience and risk-taking in margin lending’, American Economic Review, vol. 106(11), pp. 3367–3400.

Kuhnen, C.M. (2015). ‘Asymmetric learning from financial information’, Journal of Finance, vol. 70(5), pp. 2029–2062.

Kuhnen, C.M. and Knutson, B. (2011). ‘The influence of affect on beliefs, preferences, and financial decisions’, Journal of Financial and Quantitative Analysis, vol. 46(03), pp. 605–626.

Malmendier, U. and Nagel, S. (2011). ‘Depression babies: Do macroeconomic experiences affect risk-taking?’, Quarterly Journal of Economics, vol. 126(1), pp. 373–416.
Malmendier, U. and Nagel, S. (2016). ‘Learning from inflation experiences’, *Quarterly Journal of Economics*, vol. 131(1), pp. 53–87.

Malmendier, U., Tate, G. and Yan, J. (2011). ‘Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies’, *Journal of Finance*, vol. 66(5), pp. 1687–1733.

Mayraz, G. (2011). ‘Wishful thinking’, https://ssrn.com/abstract=1955644 (last accessed 12 July 2019).

Nax, H.H., Burton-Chellew, M.N., West, S.A. and Young, H.P. (2016). ‘Learning in a black box’, *Journal of Economic Behavior and Organization*, vol. 127, pp. 1–15.

Nevo, I. and Erev, I. (2012). ‘On surprise, change, and the effect of recent outcomes’, *Frontiers in Psychology*, vol. 3, pp. 1–9.

Rabin, M. (2002). ‘Inference by believers in the law of small numbers’, *Quarterly Journal of Economics*, vol. 117(3), pp. 775–816.

Rabin, M. and Schrag, J.L. (1999). ‘First impressions matter: A model of confirmatory bias’, *Quarterly Journal of Economics*, vol. 114(1), pp. 37–82.

Schlag, K.H., Tremewan, J. and Van der Weele, J.J. (2015). ‘A penny for your thoughts: A survey of methods for eliciting beliefs’, *Experimental Economics*, vol. 18(3), pp. 457–490.

Seru, A., Shumway, T. and Stoffman, N. (2010). ‘Learning by trading’, *Review of Financial studies*, vol. 23(2), pp. 705–739.

Thaler, R.H. and Johnson, E.J. (1990). ‘Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice’, *Management Science*, vol. 36(6), pp. 643–660.

Thorndike, E.L. (1898). ‘Animal intelligence: An experimental study of the associative processes in animals’, *Psychological Monographs: General and Applied*, vol. 2(4), pp. i–109.

Tversky, A. and Kahneman, D. (1973). ‘Availability: A heuristic for judging frequency and probability’, *Cognitive Psychology*, vol. 5(2), pp. 207–232.