Testing Models of Belief Bias: An Experiment

Alexander Coutts[*†]
Nova School of Business and Economics

November 16, 2018

Abstract

Optimistic beliefs affect important areas of economic decision making, yet direct knowledge on how belief biases operate remains limited. To better understand these biases I introduce a theoretical framework that trades off anticipatory benefits against two potential costs of forming biased beliefs: (1) material costs which result from poor decisions, of Brunnermeier and Parker (2005), and (2) direct psychological costs of distorting reality, of Bracha and Brown (2012). The experiment exploits the potential of the increasingly popular BDM elicitation procedure adopted to lotteries to distort beliefs in different directions, depending on which costs are most important. Relative to an elicitation procedure without distortionary incentives, beliefs are biased in the optimistic direction. Increasing payments for accuracy further increases belief reports, in many cases away from the truth, consistent with psychological costs of belief distortion. Yet the overall results suggest that such theories of optimism fail to explain how beliefs respond to financial incentives.

JEL classification: C91, D03, D80, D81, D83, D84.

Keywords: Beliefs · Optimism · Pessimism · Overconfidence · Anticipation · Affective expected utility

[*]Nova School of Business and Economics, Universidade NOVA de Lisboa, Rua da Holanda 1, Campus de Carcavelos, 2775-405 Carcavelos, Portugal; alexander.coutts@novasbe.edu

[†]Acknowledgements: This research has been generously supported by grant #98-14-06 from the Russell Sage Foundation. I am heavily indebted to David Cesarini for numerous discussions and comments. I am grateful for helpful comments from Hunt Allcott, Andrew Demers, David Dillenberger, David Eil, Kfir Eliaz, Guillaume Fréchette, Xavier Gabaix, Nicole Hildebrandt, John Leahy, Elliot Lipnowski, David Low, Amnon Maltz, Joseph Mullins, Giorgia Romagnoli, Andrew Schotter, Emilia Soldani, Tobias Salz, Séverine Toussaert, Pedro Vicente, Natanya Wilson, Christopher Woolnough, Sevgi Yuksel, and numerous participants at ESA International and North American meetings in 2014. Finally this paper benefitted immensely from very helpful comments and suggestions from the associate editor and three anonymous referees.

1 Introduction

Optimistic beliefs play an important role in decision making, yet a lack of direct evidence hinders the ability of researchers to model them. Accurately modeling optimism is critical for developing theory and informing policy. There is evidence that optimistic beliefs affect decisions such as medical testing, saving for retirement, starting a new business, or investing in the stock market.[1] This has motivated theorists to challenge the benchmark rational model of decision making under uncertainty. While substantial progress has been made, rigorous tests of existing theory and direct evidence about optimism are scarce. This paper develops a theoretical framework and conducts a corresponding lab experiment that provides this direct evidence.

In the framework of optimistic belief formation, individuals derive pleasure from the anticipation of future outcomes. For example, the belief that one will be financially successful or will not have an incurable illness in the future may increase utility today. Early thinkers such as Bentham posited an important role for anticipatory utility, which motivates belief distortion in modern theories of optimism. The framework nests two such prominent models, Brunnermeier and Parker (2005) (henceforth BP) and Bracha and Brown (2012) (henceforth BB), and assumes biased beliefs emerge endogenously from a subconscious optimization problem where self-delusion is welfare enhancing. For example, one may be happier believing the probability they have HIV is lower than it is.

The experiment tests these theories, by comparing belief estimates about an identical event in the lab, where some subjects have the opportunity to earn financial prizes of $80 if the event occurs, while others do not. Those with the $80 stake have a greater incentive to be hopeful or optimistic that the event will occur, due to benefits from anticipation. Comparing these two groups allows for a first test of whether anticipation of these prizes alters belief reports. However, the primary innovation of the framework and experiment is that it allows for a test of how optimistic beliefs are constrained, by leveraging the properties of an incentive compatible elicitation procedure in order to distinguish two plausible cost mechanisms.

First, there are indirect material costs associated with holding incorrect beliefs, as individuals will subsequently make suboptimal choices. For example, believing that one has a lower probability of having HIV may lead an individual to not get tested, which has negative consequences for later life outcomes. These costs are emphasized in the optimal expectations model of BP, where the key tradeoff is between the anticipatory benefits from holding optimistic beliefs and the costs of these beliefs due to worse decision-making.

Second, there may be direct psychological costs of distorting reality, i.e. constraints on the ability to manipulate beliefs that increase as one moves further away from reality. For example, an individual engaging in high HIV risk behaviors may struggle to rationalize

[1] Evidence that optimistic beliefs play a role in these situations can be found in Oster et al. (2013) and Grulich (2000), Puri and Robinson (2007), Landier and Thesmar (2009), and Easterwood and Nutt (1999) respectively.

a belief that they are at low risk, given available statistics and information campaigns. These costs are emphasized in the affective decision making model of BB, involving a static simultaneous moves game between a rational and an emotional process. The rational process makes choices, acting as a standard expected utility maximizer given beliefs. The emotional process chooses beliefs, values only anticipation, and incurs a direct cost of distorting reality.

The framework and lab experiment distinguish the relative importance of these two costs by varying the payment for accuracy under a specific elicitation procedure, a variant of the Becker-DeGroot-Marschak (BDM) method applied to lotteries, henceforth referred to as the lottery method. This method is increasingly utilized by experimental economists as it is incentive compatible, i.e. agents maximize their chances of earning the accuracy payment by reporting the true probability, for any risk-attitudes, see Karni (2009).[2] The lottery method has an additional key property that is exploited in this experiment: expected payments for accuracy are increasing in the likelihood of the event being predicted.

As a result, increasing the accuracy payment has two effects. Increasing rewards for accuracy motivates agents to be more hopeful that the event will occur, as with the case of financial prizes. This follows since the more likely they believe the event is, the more they anticipate receiving the reward. At the same time, the elicitation procedure is incentive compatible, meaning that individuals suffer greater material losses from reporting incorrect beliefs. These expected losses are increasing in the size of the accuracy payment.

In the BP model, this second effect dominates. The overall effect of increasing accuracy payments is to raise the material costs of distorted beliefs, and hence beliefs become less optimistic. In the BB model, it is primarily mental costs of belief distortion that constrain optimistic beliefs. Mental costs are unaffected by the accuracy reward levels, hence in the framework the BB model makes the prediction that increasing the accuracy payment leads to greater belief distortion.

Distinguishing theories of optimism bias is critical for developing policy to reduce harmful instances of such bias. Oster et al. (2013) suggest that individuals are optimistically biased about their risk of having Huntington Disease (HD), a degenerative neurological disorder, and are subsequently unlikely to get tested. As Oster et al. (2013) note, not knowing one's status precludes one from taking important life decisions. Grulich (2000) observes that optimistic individuals are more likely to engage in behaviors that increase the risk of HIV.

While both approaches to modeling the costs of optimism are consistent with observed behavior in these examples, the policy prescriptions for de-biasing individuals differ. Take the example of policies to increase rates of testing for HIV, if one believes that individuals' beliefs about having HIV are distorted downwards. To de-bias this belief, if future costs affect belief formation, campaigns could emphasize improved treatment options for those

[2] Some recent empirical studies using the lottery method include: Holt and Smith (2009), Mobius et al. (2014), Benoît et al. (2015), Ertac et al. (2017), Ambuehl and Li (2018), and Buser et al. (2018).

with HIV (as in the CDC's "HIV Treatment Works" campaign). On the other hand, if mental costs of belief distortion are important, a different approach could be to emphasize a salient statistic or image that forces individuals to confront their risk status, but doesn't necessarily contain novel information. For example, the AIDS Service Foundation of Greater Kansas City ran a campaign with powerful imagery of individuals with quotes such as "more than one million people in the USA have HIV" written on their faces.[3]

The elicitation procedure is central to distinguishing these two approaches, as reported beliefs will either increase or decrease depending on which cost mechanism dominates. As a robustness treatment I further examine belief reports under a different elicitation procedure that is completely void of incentives for belief distortion.

The overall results suggest the following. Beliefs are distorted in important ways, resulting in a rejection of the standard rational expectations framework. However, the optimism framework is unable to explain all of the observed patterns. First, individuals on average make non-motivated mistakes, as beliefs differ significantly from true probabilities, but not unanimously in the optimistic direction. Controlling for these mistakes by taking the robustness treatment as the relevant counterfactual, beliefs are biased upwards by 13% when distortionary incentives are present, in line with optimism models. The framework additionally predicts that when subjects are given financial prize stakes of \$80, their beliefs will become more optimistically biased. This is observed, but only in sessions where accuracy payments are low; when accuracy payments are moderate or high, or in the overall sample, significant effects are not observed.

In what concerns distinguishing the cost mechanisms in models of optimism bias, as accuracy payments are increased using the incentive compatible elicitation procedure, the average belief report increases significantly by approximately 9%, in many cases a movement *away* from the truth. While significant in the overall sample, this finding is driven by events where individuals do not have the \$80 prize state. Similar patterns are not observed in robustness treatments. Adhering to the framework, this result implies that mental costs, as in BB, are required to rationalize the observed patterns.

These main results cast some doubt on the viability of models of optimism to explain behavior pertaining to financial stakes in the lab. Yet they highlight important takeaways for theoretical and empirical work. Financial stakes are shown to alter belief reports, in a way that a number of models of belief formation and choice are unable to explain. Despite the failure of theories of optimism to capture how beliefs respond to financial incentives, these theories continue to add explanatory power beyond standard models unable to account for these patterns.

Beyond these insights, the premise of the theoretical framework involves using features of an increasingly common incentive compatible elicitation procedure to manipulate beliefs through changes in financial stakes. The empirical results confirm that altering payments

[3]Examples of both materials can be found in Online Appendix E. The existence of mental costs of distortion do not directly imply that powerful imagery per-se will alter these costs, this would critically depend on the true form of such a mental cost function.

under the lottery method alters belief reports, and that these patterns may go against initial intuitions, i.e. larger accuracy payments can lead to more, rather than less biased beliefs. To my knowledge this provides the first empirical evidence documenting such distortions through changes in financial stakes, and suggests additional caution in using these methods to elicit beliefs.

2 Related Literature

The notion that individuals gain utility from the anticipation of future consumption, follows early economic thinkers such as Bentham and more recent theoretical work such as Loewenstein (1987), Caplin and Leahy (2001), and Koszegi (2010).[4] In contrast to the current paper and related models by BP and BB on direct belief distortion, these models assume rational expectations and focus on the implications of beliefs entering the utility function for broader choice behavior. In particular, they do not allow for individual's to hold incorrect beliefs about states of the world. Akerlof and Dickens (1982) present one of the earliest models of subconscious belief distortion, in the context of workers who wish to minimize discomfort from fear of working in a dangerous industry. Their model is a variant of BP, where agents trade off the benefits of belief distortion due to fear reduction against the costs of worse decisions, e.g. not investing in safety equipment.

An emerging literature discusses the possibility that biased beliefs arise through other channels besides direct belief distortion, typically through information manipulation. For example, individuals may deceive themselves through biased recall (Benabou and Tirole (2002), Benabou and Tirole (2006)), selective sampling (Eliaz and Spiegler (2006), Carrillo and Mariotti (2000)), or biased information processing (Landier (2000), Yariv (2005), Mayraz (2013)). In this paper the focus is on static belief formation, with no role for information gathering or recall, which precludes such channels.

Few lab experiments have been conducted in economics to test for the presence of optimistic beliefs.[5] Most related is Mayraz (2013) who focuses on finding evidence of optimistic beliefs in a novel experiment. Sessions were divided into "farmers" and "bakers", with income depending differentially on hypothetical prices of wheat, by construction these were non-predictable. Relative to bakers, farmers predicted higher prices, suggestive of optimistic belief formation. Further, Mayraz (2013) shows how changing costs for accuracy did not alter beliefs, and interprets this as consistent with a model of optimistic information

[4]Existing work has considered other channels through which individuals may benefit from holding optimistic beliefs, such as social signalling or self-discipline as in Benabou and Tirole (2002) and Carrillo and Mariotti (2000).

[5]Within psychology, there exist some direct experimental tests of optimism, however participants in these experiments are typically not provided financial incentives for accurate responses, and these studies are often not designed with the aim of distinguishing or testing theory, e.g. Vosgerau (2010). Related are experiments on testing models of overconfidence which assume individuals benefit from holding positive views about their self-image or ability, e.g. Eil and Rao (2011) and Mobius et al. (2014).

processing in which there are no costs to belief distortion, unlike the models considered in this paper. Beyond these differences, I use an incentive compatible elicitation procedure, and include a control where there are no motives to distort beliefs.

3 Experiment Design

3.1 Background

The lab experiment was conducted at New York University at the Center for Experimental and Social Science (CESS). 462 subjects participated in 43 sessions, which were conducted initially in spring 2014 (31 primary sessions) with 8 followup and 4 robustness sessions conducted in spring 2016. The experiment was programmed and conducted in z-Tree, Fischbacher (2007), with some components done by hand in order to increase salience. Average subject payments were \$25 for approximately 75 minutes in the primary sessions, and \$15 for 50 minutes in the followup and robustness sessions. Instructions for the experiment can be found in Online Appendix F.[6]

The primary sessions differed as they involved an additional incentivized component designed to study updating in response to noisy signals, the subject of a related paper, Coutts (2018).[7] The analysis in this paper focuses only on initial belief formation, whereas the other study looks exclusively at updating patterns. One concern is that knowledge of future participation in this updating task may have altered initial belief formation. For the theoretical models of interest in the framework of this paper, belief formation is independent of future opportunities to update. Further, as the updating task was completed by all subjects, the comparative static analysis remains valid.

3.2 Events

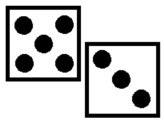
The experiment involved subjects estimating the probability of 4 (primary sessions) or 8 (followup/robustness sessions) different binary events, E, in the lab. Figure 1 summarizes the events: six events are "objective", following the definition of Gilboa and Schmeidler (2001), involving dice, coins, or playing cards. The outcome of these events was determined by chance, programmed on z-Tree, and individuals could not affect the outcome. For three of the dice events, the experiment also varied whether individuals were given control over selecting their own numbers, in order to test the "illusion of control", Langer (1975).

[6] The motivation and design with testable hypotheses was outlined in Russell Sage Foundation grant proposal #98-14-06, prior to the experimental data collection.

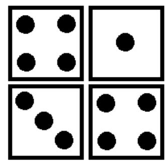
[7] While there is overlap between the two papers, a primary aim of studying updating of beliefs was to test whether updating about personal qualities such as intelligence resembled updating about objective events, following studies on asymmetric updating for example Mobius et al. (2014). Of interest is the possibility that subjects may have pursued different dynamic strategies, i.e. forming biased priors then updating differently depending on the financial stakes. However, as detailed in Coutts (2018), there is no evidence that subjects pursue different updating strategies when financial stakes are different.

Figure 1: Description of Events

All Sessions

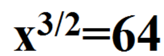


(a) (E_1) Easy Dice: The computer rolls two dice. Event occurs when two different specified numbers were the only numbers to come up (e.g. 5-3, or 3-5, 3-3, 5-5). In the control treatment individuals select the two numbers. The probability of this is $\frac{4}{36}$ or approximately 11.11%.



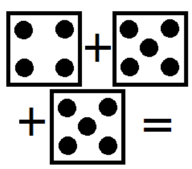
(b) (E_2) Hard Dice: The computer rolls four dice. Event occurs when exactly two out of those four dice was a specified number (e.g. 4). In the control treatment individuals select this number. The probability of this is $\binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 = \frac{150}{1296}$ or approximately 11.57%.

(c) (E_3) Weather: Event occurs if the individual correctly estimated the average temperature on a specified random day in NYC in the previous year (2013/2015), $\pm 5°$F. In the sample, 26.59% of subjects were in the correct range.



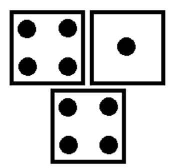
(d) (E_4/E_5) Quiz: Event occurs if the individual scored in the top 15% on a skill-testing multiple choice quiz, relative to students in pilot sessions (E_4: self). For a subset of participants the event pertained to a random partner's performance instead of their own (E_5: other).

Followup and Robustness Sessions Only

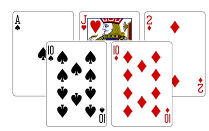


(e) (E_6) Sum Dice: The computer rolls three dice. Event occurs when sum of three numbers is less than or equal to 14. There are 20 out of 216 combinations where this is not true. The probability of this is thus $\frac{196}{216}$ or approximately 90.74%.

(f) (E_7) Coins: The computer flips eight coins. Event occurs when at least three coins are tails. The probability of this is $1 - \left(\binom{8}{0}+\binom{8}{1}+\binom{8}{2}\right)\left(\frac{1}{2}\right)^8 = \frac{219}{256}$ or approximately 85.55%.



(g) (E_8) Three Dice: The computer rolls three dice. Event occurs when out of two different specified numbers, at least one appears at least once. In the control treatment individuals select the two numbers. The probability of this is $1-\frac{4}{6}^3$ or approximately 70.37%.



(h) (E_9) Cards: The computer selects 5 cards from a standard 52-card deck. Event occurs if there is at least one pair (includes three/four of a kind). Probability of this is approximately: 49.29%. See http://www.math.hawaii.edu/~ramsey/probability/pokerhands.html.

Two events involved performance on a quiz and on a question about historical weather. The quiz (self) event involved whether a subject scored in the top 15% on a five minute skill testing quiz consisting of math and verbal questions taken by all subjects before the experiment. Subjects were incentivized by being truthfully informed that achieving a high score on the quiz would result in an increased probability of earning income in the experiment.[8] To determine whether subjects were in the top 15% they were compared to a reference group taking the same quiz during pilot sessions. A random subset of subjects (30%) were selected instead to participate in the quiz (other) event, regarding the performance of a random anonymous partner in the room, instead of their own performance. Finally, the weather event involved correctly estimating within 5°F the average temperature on a given, random day in the previous calendar year (2013 or 2015) in New York City.

Events were presented in random order, with one being selected at random to determine payment upon completion of the experiment.[9] While the primary sessions included only events that were relatively unlikely ($\pi < 0.5$), the followup and robustness sessions included others with $\pi \geq 0.5$.

3.3 Belief Elicitation

To incentivize truthful belief reports, individuals faced the elicitation procedure of Karni (2009), which I refer to as the lottery method.[10] The method is implemented as follows. Given the individual's belief report b, a random number r is drawn from any distribution with full support on $[0, 1]$, here I use the uniform distribution. If the individual's report $b \geq r$ she is paid an amount $a > 0$ if E occurs, and 0 if E does not occur. If $b < r$, she plays a lottery that pays out a with probability r, and 0 otherwise.

The incentive compatibility of this method follows from a dominance argument, which requires only that individuals exhibit probabilistic sophistication over lotteries, see Machina and Schmeidler (1995). It does not require assumptions about risk preferences. Accuracy payments a were randomized at the session level: low ($3), moderate ($10 - primary sessions only), or high ($20).[11] The lottery method has the key property that one is more likely to earn the accuracy payment a when the event is more likely, given any reported probability. This fact does not affect the incentive compatibility of the procedure, but will be critical for the theoretical predictions described later.

[8] Paying subjects directly for their performance would introduce another financial stake, which would preclude incentive compatibility of the elicitation procedure, see Karni and Safra (1995). If subjects gain utility from their ego (as in Mobius et al. (2014)) this would pose a similar threat to incentive compatibility. The results to be presented are robust to excluding the quiz event.

[9] During the primary sessions one event (Easy Dice - see event descriptions) was fixed as the last event. In the followup and robustness the order of all events was randomized.

[10] See Karni (2009) for a more detailed description of the lottery method, though the method itself has been described in a number of earlier papers. See Schlag et al. (2015) for details about earlier descriptions of this mechanism. The mechanism is also referred to in various papers as the "crossover method", "matching probabilities", and "reservation probabilities".

[11] Common to other elicitation procedures such as the quadratic scoring rule, the lottery method involves very flat playoffs around the true beliefs. For an event with $\pi = 0.5$, a distorted report of $\hat{\pi} = 0.6$ would result in only a $\$0.005 \cdot a$ reduction in payoffs. Although small monetary consequences do not alter the predictions of the framework, this does raise a concern that the magnitude of some belief changes may be difficult to detect in the experiment.

3.4 Financial Stakes

A primary test of the optimism framework involves varying the size of financial stakes, introduced in the form of potential prize payments, $P \geq 0$, randomized at the subject-event level. Subjects endowed with a financial stake P in an event j could earn that amount if the event occurred, independent of their decisions in the experiment. These stakes were only used in primary sessions, i.e. E_j for $1 \leq j \leq 5$. Half of subjects had the chance to earn an extra \$80 if a given event occurred, while the other half would earn nothing extra. Assignment of the prize stake, $P \in \{\$0, \$80\}$ was done by hand, with each subject drawing a poker chip with the prize amount listed.

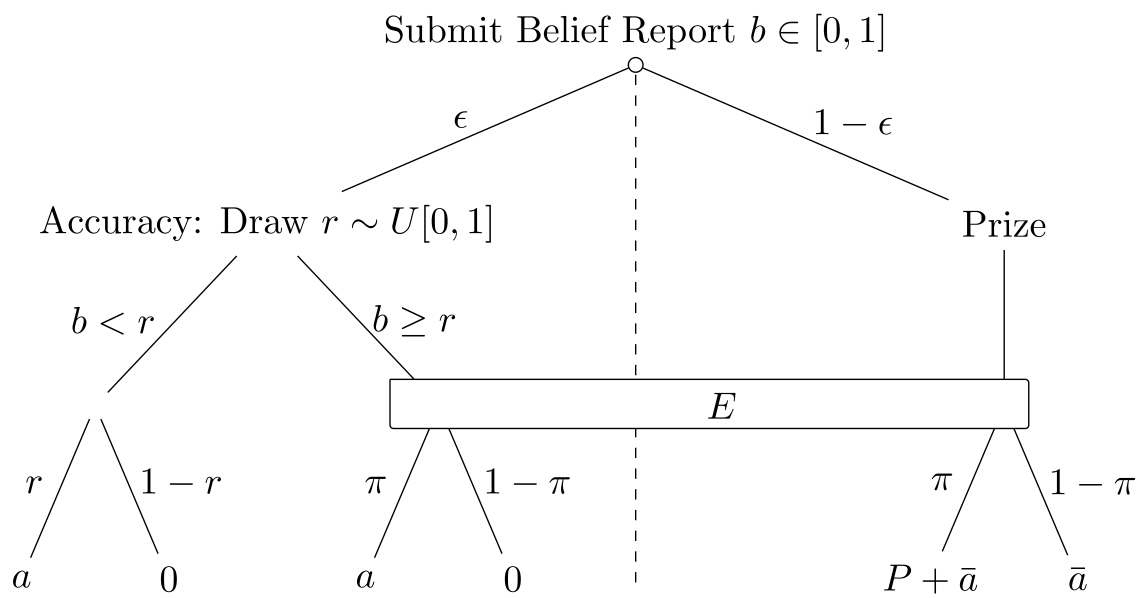
3.5 Partition: Accuracy State and Prize State

It is theoretically impossible to truthfully elicit beliefs about events that non risk neutral individuals have a financial stake in, see Karni and Safra (1995), due to hedging between accuracy and prize payments. To restore incentive compatibility of the elicitation procedure, subjects were only eligible to receive the accuracy *or* the prize payment, determined at random at the end of the experiment, shown in Figure 2.

First the individual submits a belief report b, regarding the probability E occurs. Next, the world is partitioned into two states, an accuracy state with probability $\epsilon \in (0, 1]$ or a prize state with probability $1-\epsilon$. In the accuracy state, the individual has the opportunity to earn $a > 0$ following the elicitation procedure. In contrast, in the prize state the individual's belief report is not relevant. Instead, the individual receives a guaranteed payment $\bar{a} \geq a$, and has the opportunity to earn $P \geq 0$ only if the event occurs.[12] This partition is similar to one utilized by Blanco et al. (2010). In the primary sessions, there was a 50% probability that an individual would end up in the prize or accuracy state ($\epsilon = 0.5$). In the followup and robustness sessions the prize state was removed entirely ($\epsilon = 1$).

[12]The payment of $\bar{a}$ was to ensure that the prize state was preferred to the accuracy state. This was motivated by earlier theoretical analysis intended to rule out counterintuitive patterns in belief distortion. It is not required for the comparative statics of interest. I thank an anonymous referee for pointing this out.

Figure 2: Structure of Stylized Model and Elicitation Procedure



Individual submits belief report b regarding probability that binary event E occurs. In accuracy state, payment depends on the elicitation procedure drawing r from the uniform distribution on $[0,1]$. In prize state, payment depends only on outcome of the event, not the belief report. π is the true probability of event E.

The first half of the experiment was dedicated to instructions and hands-on practice with the elicitation procedure in z-Tree. After the detailed practice, subjects took the five minute skill testing quiz, followed by answering the weather question. Before introducing the subsequent event, the prize amount was drawn by hand (primary sessions only), next the event was revealed, and then the elicitation occurred. This was repeated for all events. To conclude the experiment, one event was randomly selected for payment, and for the primary sessions it was determined, by drawing poker chips, whether the subject ended up in the prize or accuracy state. The experiment concluded with a short questionnaire.

3.6 Robustness Treatments

In 4 sessions ($N = 54$), subjects were allocated to a robustness treatment with no incentives to hold optimistic beliefs. Rather than eliciting beliefs about a single event occurrence, subjects were asked to estimate the frequency that the event had occurred for all subjects that previously participated.[13]

Subjects received the accuracy payment of $a \in \$\{3, 20\}$ if they were within 5 percentage points of the true frequency. The outcome of 318 binary events with probability π of "success" follows a binomial distribution. The optimal choice is to report the modal frequency of successes, $\frac{\lfloor (318+1)\pi \rfloor}{318} \approx \pi$.[14] Thus it follows that this elicitation procedure is incentive compatible in the current framework, as the mean and mode frequencies are equivalent given the granularity of the elicitation procedure. This procedure is robust to optimistic belief distortion as expected payments are constant with respect to the frequency of the event having occurred. Since individuals have no other financial or personal stakes in these events, they are indifferent to the frequencies of occurrence, their only concern is with being in the correct interval, see Online Appendix B.

To give subjects familiarity with the quiz and weather events so that they would be in a better position to evaluate others' performance, they were paid \$0.50 for each point on the quiz. For the weather question they earned \$0.50 for a correct answer. Table 1 presents a summary of the different types of sessions in the experiment.

Table 1: Summary of Experimental Sessions

	Regular		Robustness
	Primary	Followup	
Participants	326	84	52
Events	1-5	1-9	1-3, 5-9*
Prize stake: P	$\{0, 80\}$	–	–
Accuracy payment: a	$\{3, 10, 20\}$	$\{3, 20\}$	$\{3, 20\}$
Elicitation Procedure	Lottery Method	Lottery Method	Past Frequencies

See Figure 1 for description of events. The prize state was removed in followup and robustness sessions.
*Robustness sessions asked about past frequencies of events, which excludes E_4: own performance on test.

[13] Since data was only available for four events in the primary sessions, the other four events were simulated 318 times (326 subjects participated, however one session crashed, leading to data for 318). Subjects were informed but did not know which of the events were simulated. Instead of being told, "The computer has rolled four dice. What is the probability that exactly two of the four come up 6?", subjects are instructed, "318 students previously faced an event where the computer rolled four dice, and their payment depended on how often exactly two of the four came up 6. In what percent of these dice rolls did 6 come up exactly twice?".

[14] $\lfloor \cdot \rfloor$ is the floor function. Note that $\pi - \frac{1-\pi}{318} = \frac{(318+1)\pi - 1}{318} \leq \frac{\lfloor (318+1)\pi \rfloor}{318} \leq \frac{(318+1)\pi}{318} = \pi + \frac{1}{318}$. Thus $\left| \frac{\lfloor (318+1)\pi \rfloor}{318} - \pi \right| \leq \frac{1}{318}$, which is smaller than the 1% granularity of the elicitation procedure. No probabilities were near the extremes of < 0.05 or > 0.95. When 319π is an integer there exists a second mode of $319\pi - 1$.

4 Theory

I now introduce a flexible framework of belief bias that admits BP, BB, as well as a benchmark Rational Expectations (RE) agent as special cases. Individuals form beliefs at the subconscious level, and subsequently take actions given these beliefs in accordance with maximizing subjective expected utility. Once actions have been pinned down, one solves for optimal beliefs by working backwards. The key feature of this framework is that the action corresponds directly to reporting a belief to the lottery method.

I apply the framework to the context of the regular sessions in the experiment which utilize the lottery method. In robustness sessions, subjects have no motive to distort beliefs about the frequencies of interest.

4.1 Preliminaries

Individual utility is given by $u(\cdot)$, a strictly increasing, differentiable function that is independent of time and across states of the world. There are two states, determined by the outcome of whether a binary event E occurs, according to objective probability $\pi \in (0,1)$.

Subjective beliefs about the probability that E occurs are given by $\hat{\pi} \in [0,1]$. Beliefs will be selected optimally, based on the tradeoff between benefits of belief distortion from anticipation and the costs of worse decision making or mental costs of distortion. While I allow $\hat{\pi}$ to lie on the boundary, following BP, I do not allow subjective beliefs to assign positive probability to non-existent states.

Given beliefs, $\hat{\pi}$, individuals take actions, unaware that their beliefs may be biased. An action will correspond to a belief report, denoted by $b(\hat{\pi}) \in [0,1]$. Denote the individual's subjective expected utility from consumption given their belief $\hat{\pi}$ by $U\big(P, a, b(\hat{\pi}); \hat{\pi}\big)$, presented explicitly in Appendix A. Note that it is increasing in both P and a.

As individuals in the framework are unaware they are biased, incentive compatibility of the elicitation procedure remains in place, with individuals reporting the biased belief they believe to be true, $b(\hat{\pi}) = \hat{\pi}$. With this action fixed, one works backwards to determine how optimal beliefs are formed.[15]

4.2 Optimal Beliefs

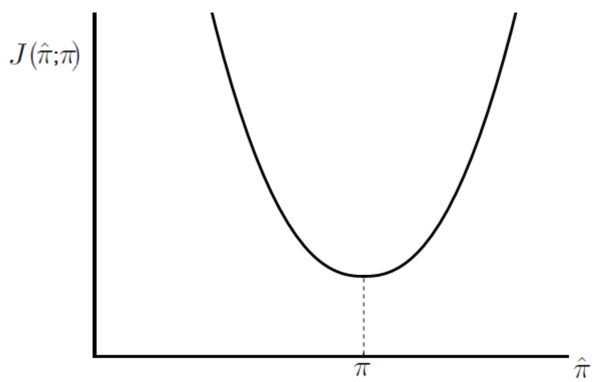
Optimal beliefs trade off the costs and benefits to holding optimistically biased beliefs. Following models of belief distortion such as BP and BB, I assume benefits to anticipation

[15] More broadly, this framework is a sequential moves game between a subconscious process that forms beliefs and a conscious, rational process that subsequently makes choices. Given that there is a dominant strategy for the rational process to truthfully report beliefs, the equilibrium outcome is identical if this were a simultaneous moves game, as in BB. In the lab it is possible to create such an environment through appropriate choice of elicitation procedure. However, even if one had observational data on choices and beliefs, because these are equilibrium quantities, it would be difficult to generate testable predictions with this framework more generally.

are proportional to subjective expected utility, $U\big(P, a, b\,(\hat{\pi})\,; \hat{\pi}\big)$. The first costs, emphasized in BP, are the material costs from worse decision making. These are captured by the true expected utility, i.e. $U\big(P, a, b\,(\hat{\pi})\,; \pi\big)$. True expected utility is increasing in π and decreasing in $|\hat{\pi} - \pi|$ due to the lottery method.

Following BB, the second costs are direct psychological costs to distorting beliefs, $J(\hat{\pi}; \pi)$. The intuition for such a cost function is based on evidence from psychology, that people use mental strategies such as biased search to justify their beliefs. As beliefs are further away from the truth, search costs to support these beliefs become greater. I follow BB in assuming $J(\hat{\pi}; \pi)$ is a non-negative, strictly convex, essentially smooth function on $(0, 1)$. The function reaches a minimum at $\hat{\pi} = \pi$, and is such that in the limit as $\hat{\pi}$ goes to either 0 or 1 for any $\pi \in (0, 1)$, $J(\hat{\pi}; \pi)$ approaches infinity at a higher rate than the utility function, guaranteeing that holding extreme beliefs is never optimal. An example of this mental cost function can be seen in Figure 3.[16]

Figure 3: Mental Cost Function



Example mental cost function, $J(\hat{\pi}; \pi)$, of holding distorted beliefs, $\hat{\pi}$, in BB model for two states. π is the true probability of one of the states.

Optimal beliefs $\hat{\pi}$ are chosen according to the maximization of a flexible weighting of true expected utility (first term, weighted by $\alpha \in \{0, 1\}$), subjective expected utility from anticipation (second term, weighted by $\gamma \geq 0$), and the mental cost of belief distortion (third term, weighted by $\beta \in \{0, 1\}$):

$$\max_{\hat{\pi} \in [0,1]} \;\; \alpha U\big(P, a, b(\hat{\pi}); \pi\big) + \gamma U\big(P, a, b(\hat{\pi}); \hat{\pi}\big) - \beta J\big(\hat{\pi}; \pi\big)\;, \;\; \alpha \in \{0, 1\},\; \gamma \geq 0,\; \beta \in \{0, 1\}. \tag{1}$$

[16]It further satisfies the following properties: $\lim_{\hat{\pi} \to 0} |J'(\hat{\pi}; \pi)| = \lim_{\hat{\pi} \to 1} |J'(\hat{\pi}; \pi)| = +\infty$ and $\lim_{\hat{\pi} \to 0} J(\hat{\pi}; \pi) = \lim_{\hat{\pi} \to 1} J(\hat{\pi}; \pi) = +\infty$, where $J'(\hat{\pi}; \pi)$ is the first derivative of $J(\hat{\pi}; \pi)$.

When there are no benefits to anticipation, i.e. $\gamma = 0$, the result is a standard RE agent, holding optimal beliefs $\hat{\pi} = \pi$. Conversely, whenever $\gamma > 0$, optimal beliefs will be such that $\hat{\pi} > \pi$, see Appendix B. While BP assume that $\gamma = 1$, to avoid degenerate beliefs, $\hat{\pi} = 1$, I follow Oster et al. (2013), Spiegler (2008), and Bridet and Schwardmann (2014), and do not impose this restriction. That α and β are binary is for expositional purposes only. The analysis is identical for any non-negative parameters, as the properties of the mental cost function are unchanged by multiplication by any positive constant.

As will be seen, removing the first term ($\alpha = 0$), material costs, in Equation 1 results in the BB model, while removing the third term ($\beta = 0$), mental costs, results in the BP model. Appendix B outlines the first and second order conditions, showing that a solution to the optimization problem is guaranteed to exist, and the solution is generically unique. When mental costs are absent, the solution may be at a corner ($\hat{\pi} = 1$) if the weight on anticipation γ is too large relative to material costs, see Appendix B. An interior solution is required for the comparative static analysis. This is assumed, and can be verified from the experimental results.

4.2.1 Comparative Statics

An increase in either P or a increases both true expected utility and subjective expected utility from anticipation, while leaving mental costs unchanged. The probability of receiving P is independent of the belief report $b(\hat{\pi})$, see Figure 2. The implication is that the marginal costs of belief distortion are unchanged, while the marginal benefits increase. Intuitively, optimal beliefs $\hat{\pi}$ will increase as P increases. Equation 2 shows this formally. For brevity, $\Delta u_a = u(a) - u(0) > 0$.

$$\frac{\partial \hat{\pi}}{\partial P} = \frac{(1-\epsilon)u'(P+\bar{a})}{(\alpha-\gamma)\epsilon\Delta u_a + \beta J''(\hat{\pi};\pi)} > 0 \tag{2}$$

The denominator of this equation is the negative of the second order condition, see Equation B2, hence for any interior solution this term is positive. As the prize P increases, belief bias increases unambiguously in the framework.

In contrast, the probability of receiving a is decreasing as subject beliefs, $\hat{\pi}$, move further away from the truth, π. Hence there are two effects of increasing a. The first is that it increases the marginal benefits of belief distortion, as with P. But the second is that it increases the marginal costs. Which effect dominates depends on the relative weights in the framework, as can be seen from Equation 3:

$$\frac{\partial \hat{\pi}}{\partial a} = \frac{\epsilon u'(a)\big(\gamma\hat{\pi} - \alpha(\hat{\pi}-\pi)\big)}{(\alpha-\gamma)\,\epsilon\Delta u_a + \beta J''(\hat{\pi};\pi)}. \tag{3}$$

When both mental and material costs are present, $\alpha = \beta = 1$, the sign of the comparative static $\frac{\partial \hat{\pi}}{\partial a}$ is ambiguous. The denominator again is positive, while the numerator is positive when the weight on anticipation is large relative to material costs and the extent of bias $\hat{\pi} - \pi$. The models of BP and BB follow from simple reductions of the framework. They make opposing predictions for this comparative static, due to their emphasis on the different cost mechanisms.

4.2.2 Brunnermeier and Parker (2005) (BP)

In the optimal expectations framework of BP, there are no mental costs of distorting beliefs, $\beta = 0$. Optimal expectations are formed based on the tradeoff between anticipatory utility and the material costs of worse decision making. The current setup can be found in the portfolio choice application of the model presented by BP Section II, with $\alpha = \gamma = 1$.[17]

Appendix C shows the closed form solution to optimal beliefs in the BP model. Substituting in the closed form solution for $\hat{\pi}^{BP}$, the comparative static for $\frac{\partial \hat{\pi}^{BP}}{\partial a}$, where $\Delta u_P = u(P + \bar{a}) - u(\bar{a}) \geq 0$, is:

$$\frac{\partial \hat{\pi}^{BP}}{\partial a} = \frac{-u'(a)\left(\frac{(1-\epsilon)\gamma}{\epsilon}\frac{\Delta u_P}{\Delta u_a}\right)}{(1-\gamma)\Delta u_a} \leq 0. \tag{4}$$

For interior solutions, $\frac{\partial \hat{\pi}^{BP}}{\partial a} \leq 0$, with equality only when $\Delta u_P = 0$. As the material costs of holding optimistically biased beliefs increase, beliefs become less biased. Increasing a increases both benefits and costs, however because utility from anticipation is less than actual utility, material costs dominate.

When there is no prize stake, i.e. $P = 0$ ($\Delta u_P = 0$), beliefs are biased upwards by a constant proportion $\hat{\pi}^{BP} = \frac{\pi}{1-\gamma}$, invariant to accuracy payments a, see Appendix C. The experimental results thus can provide an estimate for γ that does not require assumptions about the functional form of utility, an exercise conducted in Section 5.1.3.

Note that if $\beta = 0$, the restriction that $\alpha = 1$ is without loss of generality, resulting in the following corollary finding. When mental costs of belief distortion are absent ($\beta = 0$), the comparative static $\frac{\partial \hat{\pi}}{\partial a}$ is necessarily non-positive. In other words, a positive comparative static for a in the framework can only be rationalized by the existence of mental costs of belief distortion.

4.2.3 Bracha and Brown (2012) (BB)

In the affective decision making model of BB, choice and beliefs are determined simultaneously by the Nash equilibrium outcome of an intrapersonal game between two cognitive processes, a rational process which chooses actions, and an emotional process that chooses beliefs. More generally, material costs of belief distortion may play a role as the rational process in BB corresponds to a standard RE agent given beliefs. However, the reduction of

[17]BP allow for utility from past consumption. In the current framework there is no role for memory.

the action space solely to a belief report in the framework precludes any further material costs. Thus, in BB, $\alpha = 0$ and $\gamma = \beta = 1$. For consistency I continue to retain the flexibility of $\gamma \geq 0$. Appendix C presents the implicit solution for optimal beliefs in the BB model. Since $\beta > 0$, an interior solution is guaranteed. The comparative static for a in the BB model is:

$$\frac{\partial \hat{\pi}^{BB}}{\partial a} = \frac{\gamma \epsilon u'(a) \hat{\pi}^{BB}}{J''(\hat{\pi}^{BB}; \pi) - \gamma \epsilon \Delta u_a} > 0. \tag{5}$$

Again the denominator is positive, deriving from the second order condition. Because the accuracy payment a only affects the benefits from holding optimistic beliefs, and not the costs, an increase in a leads individuals to form more optimistic beliefs. This can be seen again looking at Equation 1, and noting that when $\alpha = 0$, the true probability π only enters this equation through mental costs, which are independent of a.

4.3 Testable Predictions

A primary prediction of the framework is that when facing the lottery method with $\gamma > 0$, individuals will hold biased beliefs $\hat{\pi}^* > \pi$, in both the BP and BB models. In the BP model, one can estimate γ, as when $P = 0$, $\hat{\pi}^* = \frac{\pi}{1-\gamma}$. In robustness sessions, beliefs will be unbiased, and invariant to a.

Regarding the comparative statics, both models predict that increasing P in regular sessions will lead to more optimistically biased beliefs, as this increases benefits of anticipation without altering costs. However, as a result of the emphasis on different ways to model the costs of holding biased beliefs, the BP and BB models give opposing comparative static predictions for the effect of changing the accuracy payment a. Moreover, a positive comparative static for a is only possible with the existence of mental costs, $J(\hat{\pi}; \pi)$.[18]

Prediction 1:

$$\frac{\partial \hat{\pi}^*}{\partial P} > 0 \text{ for BP and BB agents,}$$
$$\frac{\partial \hat{\pi}^*}{\partial P} = 0 \text{ for a RE agent.}$$

[18] Of mention is also whether either model, BP or BB makes testable predictions regarding the cross-partial derivative $\frac{\partial^2 \hat{\pi}}{\partial a \partial P}$. From Equation 4, it is possible to note that in the BP model, $\frac{\partial^2 \hat{\pi}}{\partial a \partial P} \leq 0$, while in the BB model, as in the general framework, the resulting sign is ambiguous.

Prediction 2:

$$\frac{\partial \hat{\pi}^*}{\partial a} < 0 \text{ for a BP agent when } \Delta u_P > 0,$$
$$\frac{\partial \hat{\pi}^*}{\partial a} = 0 \text{ for a RE agent and an BP agent when } \Delta u_P = 0, \text{ and}$$
$$\frac{\partial \hat{\pi}^*}{\partial a} > 0 \text{ for a BB agent.}$$

5 Experimental Results

Before examining the two main comparative statics, I present an overview of individuals' belief reports. The theoretical framework predicts that beliefs will be optimistically (upward) biased in regular sessions which utilize the lottery method, but that beliefs will be unbiased in the robustness sessions. Table 2 presents summary statistics for individual belief reports for each of the events. The left part of the table examines beliefs in regular sessions, where there are incentives to distort beliefs when $\gamma > 0$. The right part examines beliefs in robustness sessions, where such incentives are absent. The final column presents the difference, with statistical significance indicated by an unpaired t-test.

Examining the regular sessions, one can reject that average beliefs are correct for each event at the 1% level. While the optimism framework predicts that regular sessions should exhibit unanimous upward bias relative to true probabilities, this is not evident from the data. For events E_1 to E_5 with true probabilities less than 49%, beliefs are indeed biased upwards, for events E_6 to E_9, with probabilities greater than 49%, average belief reports are biased downwards, i.e. a pessimistic bias.

The robustness sessions are free from any incentives to distort beliefs, yet one can analogously reject that average beliefs equal the true probabilities for every event at the 1% level, except for the Easy Dice event (which can be rejected at the 5% level), an indication that individuals on aggregate make non-motivated mistakes.

Consequently, I relax the presumption that counterfactual beliefs will be on average unbiased, and to take the robustness sessions as the relevant benchmark.[19] Average belief reports in the regular sessions are higher than reports in robustness sessions, for every event in the study. To conduct an aggregate test of the difference in reported beliefs that accounts for correlated errors and unbalanced observations, the final rows present the coefficient on a dummy for robustness sessions in a regression of the belief report on event level dummy variables, see Online Appendix D.2.

The average belief report is 6.35 percentage points greater in regular sessions, controlling for the type of event. Overconfidence may affect belief reports for the weather event, E_3, involving perceptions about self-estimates of the weather. Dropping this event

[19] Note that this requires a re-interpretation of how π was defined in the framework, as rather than the true objective probability, it references counterfactual beliefs observed without incentives to distort beliefs.

Table 2: Summary Statistics About Reported Beliefs

Event	π	$\mu_{\hat{\pi}}$	$\sigma_{\hat{\pi}}$	N	$\mu_{\hat{\pi}}$	$\sigma_{\hat{\pi}}$	N	Diff
		Regular	Sessions		Robustness	Sessions		
(E_1) Easy Dice	11.11	17.69	15.48	402	15.92	15.63	52	1.77
(E_2) Hard Dice	11.57	20.15	17.39	402	17.67	13.54	52	2.48
(E_3) Weather	26.59	62.21	22.14	410	44.17	25.02	52	18.04***
(E_4) Quiz Self	15.00	48.10	27.32	281	–	–	–	–
(E_5) Quiz Other	15.00	27.19	18.52	121	26.71	21.22	52	0.48
(E_6) Sum Dice	90.74	59.98	24.25	84	55.88	26.21	52	4.09
(E_7) Coins	85.55	53.12	25.54	84	45.40	27.26	52	7.72 *
(E_8) Three Dice	70.37	47.65	27.13	84	38.15	24.07	52	9.50 **
(E_9) Cards	49.29	26.55	19.95	84	20.52	19.46	52	6.03 *
All Excluding E_4							2087	6.35***
All Excl. E_3 & E_4							1625	4.25***

†π refers to the calculated underlying probability of the event. For the Weather event this probability is the empirical frequency in the data. $\mu_{\hat{\pi}}$ is the mean of the belief report $\hat{\pi}_i$ and $\sigma_{\hat{\pi}}$ is the standard deviation. Final two rows are from regression of belief report on event level indicators and dummy for robustness session.

reveals overall belief reports are 4.25 percentage points greater when incentives to distort are present. Thus beliefs are biased upwards by 13% in regular sessions relative to robustness, consistent with the framework. This finding also is evidence that the lottery method, which has attractive features due to incentive compatibility independent of risk attitudes, may directly distort beliefs, with clear implications for its use for belief elicitation.

5.1 Testing the Two Comparative Static Predictions

The dependent variable $b_{ij} \in [0, 100]$ is the reported belief of individual i regarding the percent chance of the event j occurring. $1 \leq j \leq 9$, an integer, indexes the events corresponding to the order in Table 2. Online Appendix D contains additional investigations of the comparative static tests looking at event level interactions, and additional tests regarding the illusion of control.[20]

[20] While there are some patterns of interest which support differences for those with control, these are not robust over the whole sample. One intriguing finding is that the result for $a = \$3$ is primarily driven by subjective events (where overconfidence may play a role) and those were subjects have control over selecting dice numbers. This potentially points to an interaction between overconfidence/control and optimism, which has found recent empirical support in Heger and Papageorge (2018).

5.1.1 Prediction 1: Does optimism bias increase with the prize stake P_{ij}?

The first comparative static of interest involves examining the coefficient β_1 on an indicator of whether or not the subject had an \$80 prize stake in the event, $1\{P_{ij} > 0\}$. E_j is an event specific fixed effect, α_k is a session level fixed effect, and ϵ_{ij} is an idiosyncratic error for each probability report.

$$b_{ij} = \beta_1 \cdot 1\{P_{ij} > 0\} + \sum_{1 \leq j \leq 9} \gamma_j \cdot E_j + \alpha_k + \epsilon_{ij} \tag{6}$$

The general framework predicts that individuals given a positive financial stake in an event will believe the event is more likely, i.e. $\beta_1 > 0$, in contrast to the standard RE model where $\beta_1 = 0$. Table 3 examines whether there are patterns in the data consistent with Prediction 1, using all regular sessions. Note that $P_{ij} = 0$ for events $j > 5$, as followup and robustness sessions did not involve prize stakes.[21] The table is split into each of the three accuracy treatments, with the aggregate data presented in the final column.

The effect of having a prize stake is only significant for the low accuracy payment ($a = \$3$) sessions, where β_1 is significant and positive at the 2% level. This is sizeable, and corresponds to a 5.65 percentage point increase in the prior probability reported by an individual, or 16%. For the other accuracy payment sessions, and the aggregate data, the effect is not significant. Finally, note that the direction of the bias is such that individuals are becoming *less accurate* as the prize stake increases, as probabilities are overestimated for all of the events $j \leq 5$ where subjects had the potential to earn a prize stake.

[21] One concern with pooling primary and followup sessions is that removal of the prize state in followup sessions also removed the 50% possibility of earning the fixed payment $\bar{a}$ rather than be paid by the elicitation procedure, which could alter belief reports. In Online Appendix C I show that belief reports do not differ significantly between primary and followup sessions. Moreover, Table 3 includes session fixed effects, as such the results are unchanged by pooling.

Table 3: Impact of Financial Prize on Beliefs (Testing Hypothesis 1)

All Events. Dependent Variable: Belief Report				
Regressor	Acc = \$3	Acc = \$10	Acc = \$20	All
$\{P = 80\}(\beta_1)$	5.650**	−0.909	−1.901	0.938
	(2.244)	(1.773)	(2.179)	(1.210)
Easy Dice (γ_1)	17.200***	10.186***	10.187***	9.554***
	(2.930)	(3.661)	(3.324)	(3.508)
Hard Dice (γ_2)	16.871***	12.390***	15.776***	12.000***
	(2.973)	(3.811)	(3.444)	(3.543)
Weather (γ_3)	58.494***	55.803***	56.867***	53.968***
	(3.393)	(4.461)	(3.656)	(3.774)
Quiz Self (γ_4)	41.466***	42.032***	45.365***	39.822***
	(3.857)	(5.034)	(4.199)	(3.891)
Quiz Other (γ_5)	23.687***	15.888***	26.718***	19.325***
	(3.806)	(4.077)	(4.591)	(3.791)
Sum Dice (γ_6)	58.034***		58.454***	54.048***
	(4.828)		(5.090)	(4.478)
Coins (γ_7)	52.323***		50.274***	47.190***
	(5.563)		(5.306)	(4.765)
Three Dice (γ_8)	48.190***		43.274***	41.726***
	(5.192)		(5.645)	(4.724)
Cards (γ_9)	25.390***		24.120***	20.619***
	(4.401)		(4.993)	(4.326)
Session Fixed Effects	YES	YES	YES	YES
R^2	0.41	0.54	0.45	0.45
Observations	784	436	732	1952

Analysis uses OLS regression. Difference significant from zero at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

5.1.2 Prediction 2: Does optimism bias change with accuracy payments?

The empirical estimation of Prediction 2 is similar to Equation 6. As the comparative static predictions are in general non-linear, I utilize dummies for $a = \$10$ and $a = \$20$, where $a = \$3$ is the omitted category. β_2 and β_3 are the coefficients of interest, demonstrating the relationship between accuracy payments a and beliefs b_{ij}. As a is randomized at the session level, I do not include session fixed effects.

$$b_{ij} = \beta_2 \cdot 1\{a = 10\} + \beta_3 \cdot 1\{a = 20\} + \sum_{1 \leq j \leq 9} \gamma_j \cdot E_j + \epsilon_{ij} \tag{7}$$

Table 4 presents the results from the specification in Equation 7, separately for regular and robustness sessions. Regular sessions are further split into the two prize stake levels. Increasing accuracy payments is associated with significantly greater belief reports, a prediction made by BB, and only possible in the framework given the existence of the mental cost function ($\beta > 0$). The prediction of the BP model is a negative or zero effect when there is no stake or a positive stake respectively, which does not appear to be supported in the data.

The observed effect is driven by those who are not provided the \$80 financial stake. For the entire sample, moving the accuracy payment from \$3 to either \$10 or \$20 has the effect of increasing beliefs by approximately 9% overall. The similar magnitude of the coefficients also points to strong non-linearities, and suggests that higher payments may only lead to small changes in beliefs. Of note is that the positive comparative static finding is consistent with the work of Mayraz (2013), who also found that higher accuracy payments led to a larger, though not statistically significant, degree of bias.

The final column of Table 4 presents the results for the robustness sessions, where individuals had no incentive to distort beliefs. The coefficient is not statistically different from zero, and small in magnitude. One concern is that a lack of power works against finding statistical significance, as only 52 individuals participated. Indeed, in Appendix Table D1 I conduct the analysis separately for the followup sessions that did not contain the prize state, a total of 84 individuals. These sessions find evidence of a positive comparative static, though it is not significant at conventional levels. Looking at the magnitude, the coefficient is six times as large as the coefficient in robustness sessions, which is only suggestive evidence that similarly positive effects may be present.

The overall results do not unanimously support the framework. However some patterns can be interpreted through its lens. First, mental costs must be present to generate the observed results. Next, the evidence is strongest for optimistic belief distortion when increasing prize stakes when accuracy payments are low, or increasing accuracy payments when prize stakes are low. This would suggest sharp diminishing returns to belief distortion, potentially indicative of the shape of the mental cost function. Given these effects, it is also not surprising that when interacting accuracy and prize payments, all three terms are significant, with the interaction being negative, shown in Online Appendix D.4.[22]

[22] Recall that BP predicts a negative comparative static $\frac{\partial^2 \hat{\pi}}{\partial a \partial P}$, though this is not inconsistent with BB or the general framework.

Table 4: Impact of Accuracy Payment on Beliefs (Testing Hypothesis 2)

All Events. Dependent Variable: Belief Report				
Regressor	No Stake	Stake = \$80	All	Robustness Check
$\{a = 10\}(\beta_2)$	5.014***	−0.734	2.995**	
	(1.848)	(2.397)	(1.492)	
$\{a = 20\}(\beta_3)$	4.880***	0.729	3.738***	0.456
	(1.642)	(2.504)	(1.376)	(2.996)
Easy Dice (γ_1)	14.412***	18.240***	15.555***	15.625***
	(1.217)	(2.392)	(1.156)	(2.420)
Hard Dice (γ_2)	16.424***	21.390***	18.016***	17.375***
	(1.366)	(2.339)	(1.212)	(2.617)
Weather (γ_3)	58.183***	63.783***	60.062***	43.875***
	(1.721)	(2.187)	(1.366)	(4.080)
Quiz Self (γ_4)	44.229***	49.418***	45.917***	
	(2.219)	(3.149)	(1.820)	
Quiz Other (γ_5)	24.602***	27.021***	25.162***	26.413***
	(2.448)	(2.956)	(1.852)	(3.992)
Sum Dice (γ_6)	57.710***		58.241***	55.586***
	(2.807)		(2.754)	(4.362)
Coins (γ_7)	50.853***		51.384***	45.106***
	(2.922)		(2.876)	(4.631)
Three Dice (γ_8)	45.389***		45.919***	37.856***
	(3.106)		(3.057)	(4.014)
Cards (γ_9)	24.282***		24.812***	20.221***
	(2.383)		(2.319)	(3.697)
Session Fixed Effects	NO	NO	NO	NO
R^2	0.42	0.46	0.43	0.29
Observations	1318	634	1952	416

Analysis uses OLS regression. Difference significant from zero at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

5.1.3 Estimating γ in the BP model

It is possible to estimate the anticipatory parameter γ, by examining the difference in beliefs between robustness and regular sessions when $P = 0$. The estimate, $\hat{\gamma}$, ranges from 0.02 for the Quiz (other) event to 0.23 for the Cards event, with an average estimate of $\hat{\gamma} = 0.12$, see Online Appendix D.6. Thus, the assumption that $\gamma = 1$ is rejected, and the results would suggest a much smaller role for anticipatory utility, representing 12% of the utility gained from actual consumption utility.

6 Conclusion

I study a theoretical framework of optimistic belief distortion when there are benefits from the anticipation of receiving financial payoffs. The framework nests two models, the optimal expectations model of Brunnermeier and Parker (2005) (BP), and the affective decision making model of Bracha and Brown (2012) (BB). Both models differ in how optimistic beliefs are constrained, highlighting intuitive mechanisms. BP states that biased beliefs lead to poor outcomes through suboptimal decisions. On the other hand, BB emphasize the importance of psychological costs of belief distortion.

In an experiment designed to test the framework and distinguish these models I find only weak evidence for optimistic belief formation. In line with the framework is that reported beliefs are significantly greater when incentives to distort are present. Further, greater payments for accuracy result in more optimistic belief reports, a result that is only possible in the framework given the existence of mental costs of belief distortion. This lends some support to models such as BB, while the results are incompatible with the predictions of BP. Inconsistent with the framework were that some predictions, such as increasing belief bias with financial prizes, were not borne out in all treatments.

Despite this weak evidence, the results are inconsistent with standard rational expectations models of belief formation and behavior, and remain incompatible with a number of plausible alternative theories, discussed further in Online Appendix A. There are clear takeaways for empirical studies of beliefs. The lottery method is a popular procedure to elicit beliefs due to its invariance to risk attitudes. Yet the results, and indeed the premise of this paper, have shown that this method itself may directly distort beliefs, which raises concerns with its use as an incentive compatible elicitation procedure.

While the analysis of this paper has shed some light on the mechanisms that underly belief distortion, much remains unknown. Patterns of belief formation observed are not well explained by existing models, and it is an open question the extent to which belief distortion may be relevant outside of the lab for more important life events such as health status. It will be important for future work to study these relevant domains.

Appendix

A Properties of Expected Utility

The following presents subjective expected utility $U\big(P, a, b(\hat{\pi}); \hat{\pi}\big)$ given an individual with belief $\hat{\pi}$, who takes an action (belief report) $b(\hat{\pi}) \in [0, 1]$ using the lottery method with the uniform distribution.

$$U\big(P,a,b(\hat{\pi});\hat{\pi}\big) = \epsilon\bigg[b(\hat{\pi})\Big(\hat{\pi}u\big(a\big)+(1-\hat{\pi})u\big(0\big)\Big)+\big(1-b(\hat{\pi})\big)\bigg(\frac{\int_{b(\hat{\pi})}^{1} rdr}{1-b(\hat{\pi})}u\big(a\big)+ \qquad \text{(A1)}$$

$$\left(1-\frac{\int_{b(\hat{\pi})}^{1} rdr}{1-b(\hat{\pi})}\right)u(0)\bigg)\bigg]+(1-\epsilon)\Big[\hat{\pi}u\big(P+\bar{a}\big)+(1-\hat{\pi})u\big(\bar{a}\big)\Big]$$

The probability ϵ is the probability of ending up in the accuracy state. In this state, a random number is drawn $r \sim U[0,1]$. When $r \leq b(\hat{\pi})$, which occurs with probability $b(\hat{\pi})$, the individual receives a when the event occurs (which occurs with probability $\hat{\pi}$ in the individual's mind). Recall that utility is given by $u(\cdot)$. When $r > b(\hat{\pi})$, which occurs with probability $1-b(\hat{\pi})$, the individual earns a with expectation $\frac{\int_{b(\hat{\pi})}^{1} rdr}{1-b(\hat{\pi})}$, given $b(\hat{\pi})$. Finally with probability $1-\epsilon$ the individual ends up in the prize state, in which case they receive $P+\bar{a}$ if the event occurs, and $\bar{a}$ if it does not (again they believe that the event occurs with probability $\hat{\pi}$).

True expected utility, i.e. from the perspective of an unbiased observer, is $U\big(P,a,b(\hat{\pi});\pi\big)$, since the event in fact occurs with true probability π. Note that the reported belief still depends on $\hat{\pi}$, the individual's subjective belief.

Regarding incentive compatibility of the elicitation procedure for any agent in the framework, it suffices to examine the choice of action, given held beliefs $\hat{\pi}$.

Proposition 1. *An agent in the framework truthfully reports her belief.*

Proof. The objective function is found in Equation A1. Since $\epsilon > 0$ the report $b(\hat{\pi})$ is relevant. The maximization problem is $\max_{b(\hat{\pi})\in[0,1]}\ U\big(P,a,b(\hat{\pi});\hat{\pi}\big)$. Setting the resulting first order condition to zero yields the optimal report $b^*(\hat{\pi}) = \hat{\pi}$. □

An important note is that this result relies on an implicit assumption, maintained throughout this paper, that individuals only distort the primary probability of interest π. If individuals distorted probabilities of other elements of the model, such as the probability of the prize state, ϵ or the distribution used for the lottery method, this could undermine the dominance argument for truthful reporting or alter the comparative static analysis.

In the lab experiment, the randomization of all other elements of the experiment such as how the accuracy and prize state were determined and the lottery method was highly transparent. In followup sessions, the prize state was removed entirely. Further, if subjects are able to transform probabilities being explicitly told to them, there are no guarantees that subjects are not distorting any other components of the experimental design.[23]

[23] In a previous version of this paper I took steps to show how the results would be affected if individuals were permitted to distort these other probabilities. Under a further set of relatively strong assumptions I was able to show that the results would continue to hold.

B Existence and Uniqueness of Optimal Beliefs

Given truthful reporting of held beliefs, I substitute $b(\hat{\pi}) = \hat{\pi}$ into Equation 1, below.

$$\max_{\hat{\pi}\in[0,1]} \alpha U\big(P, a, \hat{\pi}; \pi\big) + \gamma U\big(P, a, \hat{\pi}; \hat{\pi}\big) - \beta J\big(\hat{\pi}; \pi\big)\ , \ \ \alpha \in \{0,1\},\ \gamma \geq 0,\ \beta \in \{0,1\}.$$

Both the first and final terms, true expected utility and the mental cost of belief distortion, are decreasing in $|\hat{\pi} - \pi|$, and strictly concave in $\hat{\pi}$. The second term, anticipation, is an increasing, convex function in the belief $\hat{\pi}$. To prove existence of a solution (maximum) I use the extreme value theorem, taking account for the discontinuity of $J(\cdot)$ at $\{0, 1\}$.

Proposition 2. *A solution always exists to the maximization problem of Equation 1.*

Proof. Let $f(\hat{\pi})$ be the objective function in Equation 1, which is continuous for $\hat{\pi} \in (0, 1)$. The issue is that it is not continuous at $\hat{\pi} \in \{0, 1\}$ as $\lim_{\hat{\pi}\in\{0,1\}} J(\hat{\pi}; \pi) \to \infty$, and not defined at these values. Instead I define a new trimmed compact set $[\xi, 1 - \xi]$ and argue that the maximum of $f(\hat{\pi})$ obtained in this set is the same as that which would obtain in the set $[0, 1]$.

Note that $\gamma U(\hat{\pi}, \hat{\pi}) + \alpha U(\pi, \hat{\pi})$ is continuous and finite for all $\hat{\pi} \in [0, 1]$. Denote $M = \max_{\hat{\pi}\in[0,1]} \gamma U(\hat{\pi}, \hat{\pi}) + \alpha U(\pi, \hat{\pi})$. Now find $\xi \in (0, 1)$ such that $\min\{J(\xi; \pi), J(1 - \xi; \pi)\} > M + J(\pi; \pi)$. Note that for any $\hat{\pi}' \in [0, \xi]$ or $[1 - \xi, 1]$, $f(\pi) > f(\hat{\pi}')$, i.e. $f(\hat{\pi}')$ cannot be the maximum.

Thus I apply the extreme value theorem to $f(\hat{\pi})$ on the compact set $\hat{\pi} \in [\xi, 1 - \xi]$ to prove the existence of a maximum. □

The first order condition of Equation 1 is given by Equation B1, where $J'(\hat{\pi}; \pi)$ denotes the first derivative of the mental cost function, and $u\big(a\big) - u\big(0\big) = \Delta u_a > 0$ and $u\big(P + \bar{a}\big) - u\big(\bar{a}\big) = \Delta u_P \geq 0$, optimal beliefs satisfy,

$$\epsilon \Delta u_a \left(\gamma \hat{\pi} + \alpha(\pi - \hat{\pi})\right) + \gamma \left(1 - \epsilon\right) \Delta u_P - \beta J'(\hat{\pi}; \pi) \geq 0, \tag{B1}$$

with equality for an interior solution. The second order condition, where $J''(\hat{\pi}; \pi) > 0$ by strict convexity of $J(\cdot)$ is given by,

$$\epsilon \Delta u_a \left(\gamma - \alpha\right) - \beta J''(\hat{\pi}; \pi). \tag{B2}$$

Because $\gamma \geq 0$, the belief $\hat{\pi} < \pi$ will never be optimal. A sufficient condition for an interior solution to exist is $\beta = 1$, by continuity and by the properties of $J'(\pi; \pi)$. This solution is generically unique.[24] Note that whenever $\gamma > 0$, and as long as one of Δu_a or Δu_P is

[24] The solution will not be unique in the degenerate case where marginal mental costs are exactly tangent to the linear marginal benefits of belief distortion on some interval. This case is ruled out by standard regularity conditions, see Bracha and Brown (2012). Uniqueness is not required for the comparative static analysis.

positive: $\hat{\pi} > \pi$. To see this note that the first order condition would not be satisfied for an interior solution when $\hat{\pi} = \pi$.

When $\beta = 0$, as in the BP model, the solution may be at the corner $\hat{\pi} = 1$. A sufficient condition for an interior solution when $\beta = 0$ is that $\gamma < \frac{\alpha(1-\pi)\epsilon\Delta u_a}{(1-\epsilon)\Delta u_P + \epsilon\Delta u_a} \leq \alpha$. Note that the benefits from anticipatory utility must be strictly less than those from actual expected consumption utility. When this condition is not satisfied, the optimization problem is strictly increasing in $\hat{\pi}$ for $\hat{\pi} < 1$, hence the corner solution is also unique. Uniqueness of the interior solution when $\beta = 0$ is guaranteed as the objective function is strictly concave.

For an interior solution the second order condition in Equation B2 will be strictly negative. If not, one could increase $\hat{\pi}$ slightly, and the marginal benefits would increase more than the marginal costs, contradicting the fact that the solution was an optimum.

C Optimal Beliefs in BP and BB

From the first order condition in Equation B1, one can find a closed form solution for optimal beliefs in the BP model, setting $\alpha = 1$, and $\beta = 0$:

$$\hat{\pi}^{BP} = \min\left\{\frac{\pi}{1-\gamma} + \frac{(1-\epsilon)\gamma}{\epsilon(1-\gamma)}\frac{\Delta u_P}{\Delta u_a}, 1\right\}. \tag{C1}$$

From Equation C1 it is clear that the optimal belief may be at the corner, $\hat{\pi}^{BP} = 1$. The condition for an interior solution was provided in the previous section.

The first order condition which determines optimal beliefs in BB can be directly read from Equation B1, imposing $\alpha = 0$, and $\beta = 1$:

$$\gamma\epsilon\Delta u_a\hat{\pi}^{BB} + \gamma(1-\epsilon)\Delta u_P - J'(\hat{\pi}^{BB};\pi) = 0. \tag{C2}$$

Note that a linear cost function, $J''(\hat{\pi};\pi) = 0$, as in Benabou and Tirole (2006), is not sufficient to constrain beliefs. Such a function necessitates following the BP framework to guarantee an interior solution, and would thus generate similar comparative static predictions to BP.

D Supplemental Tables

D.1 Analysis in Followup Sessions

Table D1 presents the analogue of Table 4 for the followup sessions only, where there was no longer any prize stake ($\epsilon = 1$). The coefficient on β_3 is positive, though smaller, and only marginally significant at the 10% level in Column 1; the standard error increases even further with individual clustering. One question is whether the coefficient of 2.9 is driven by events with low probabilities of occurring. In fact, examining the events $1 \leq j \leq 5$ (probabilities < 0.5) and $6 \leq j \leq 9$ (probabilities ≥ 0.5) separately the coefficient β_3 is 3.3 and 2.6 respectively, both non-significant, and not statistically different from one another.

Thus it appears the result is not driven by differences in the probabilities of the events occurring.

Using data from the primary sessions, I examine how common it would be to observe a non-significant result when there are significantly fewer subjects. I randomly sample from the previous data without replacement 10,000 times, selecting 84 subjects. In 43% of the regressions the coefficient is significant at conventional ($\leq 10\%$) levels with clustered standard errors, indicating the difficulties of making inferences with relatively small samples, given the nature of these specific data. It is worth pointing out that the robustness sessions are similarly low powered, although the coefficient in the robustness session is additionally much smaller in absolute magnitude, at 0.456.

Table D1: Impact of Accuracy Payment on Beliefs (Testing Hypothesis 2) - Replication Sessions

All Events. Dependent Variable: Belief Report†		
$\{a = 20\}(\beta_3)$	2.907*	2.907
	(1.757)	(2.372)
Easy Dice (γ_1)	18.626***	18.626***
	(1.872)	(1.931)
Hard Dice (γ_2)	16.448***	16.448***
	(1.724)	(2.002)
Weather (γ_3)	56.924***	56.924***
	(3.042)	(3.166)
Quiz Self (γ_4)	38.600***	38.600***
	(3.425)	(3.580)
Quiz Other (γ_5)	27.994***	27.994***
	(4.182)	(4.287)
Sum Dice (γ_6)	58.626***	58.626***
	(2.803)	(2.986)
Coins (γ_7)	51.769***	51.769***
	(2.960)	(3.078)
Three Dice (γ_8)	46.305***	46.305***
	(3.103)	(3.244)
Cards (γ_9)	25.198***	25.198***
	(2.329)	(2.575)
Session Fixed Effects	NO	NO
R^2	0.34	0.34
Observations	672	672

Analysis uses OLS regression. Difference significant from zero at * 0.1; ** 0.05; *** 0.01. † Robust standard errors, clustered at individual level only in the last column. R^2 corrected for no-constant.

D.2 Prediction 2: Examining Upward and Downward Biased Priors

Table D2 examines Prediction 2 splitting the data into events that are overestimated and events that are underestimated. This addresses the question of whether the optimistic bias observed in Table 4 is partly accounted for by subjects becoming more accurate. In fact, it can be seen in Column 1 that the pattern of positive effects of larger accuracy payments is in fact most significant when events are already overestimated, indicating that subjects are on average becoming less accurate.

Table D2 also presents the same analysis for the robustness sessions where there are no incentives to distort beliefs. While Table 4 demonstrated that the average effect of the accuracy payment was near zero for these sessions, in fact, one can see that there is some evidence that individuals do respond to a, as beliefs appear to become more accurate, with a having a negative effect on overestimated events, but a positive effect for underestimated events.

One puzzle is why the coefficient in column 1 for $\{a = 20\}\times$ *Underestimated* is smaller than the corresponding coefficient for $\{a = 20\}\times$ *Overestimated*. If there exists both a desire for optimism and a desire for accuracy, these forces work together for downward biased subjects, but against each other for upward biased subjects, which would suggest an opposite pattern than what is observed. While I do not have an explanation for this, it is important to note that the coefficients are not significantly different from one another.

Table D2: Impact of Accuracy Payment on Beliefs (Over or Under)

All Events. Dependent Variable: Belief Report		
Regressor	All	Robustness Check
$\{a = 10\}\times$ Overestimated	3.159^{**}	
	(1.505)	
$\{a = 20\}\times$ Overestimated	4.074^{***}	-0.963
	(1.401)	(2.892)
$\{a = 20\}\times$ Underestimated	2.553	1.876
	(3.408)	(4.457)
Easy Dice (γ_1)	15.391^{***}	16.553^{***}
	(1.197)	(2.061)
Hard Dice (γ_2)	17.852^{***}	18.303^{***}
	(1.209)	(2.339)
Weather (γ_3)	59.898^{***}	44.803^{***}
	(1.383)	(4.017)
Quiz Self (γ_4)	45.749^{***}	0.000^{***}
	(1.821)	(0.000)
Quiz Other (γ_5)	25.008^{***}	27.341^{***}
	(1.855)	(3.782)
Sum Dice (γ_6)	58.791^{***}	54.658^{***}
	(3.193)	(4.914)
Coins (γ_7)	51.934^{***}	44.177^{***}
	(3.398)	(5.384)
Three Dice (γ_8)	46.469^{***}	36.927^{***}
	(3.404)	(4.651)
Session Fixed Effects	NO	NO
R^2	0.43	0.29
Observations	1952	416

Analysis uses OLS regression. Difference significant from zero at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

References

Akerlof, George A and William T Dickens (1982) "The Economic Consequences of Cognitive Dissonance," *American Economic Review*, Vol. 72, No. 3, pp. 307–319.

Ambuehl, Sandro and Shengwu Li (2018) "Belief updating and the demand for information," *Games and Economic Behavior*, Vol. 109, pp. 21–39.

Benabou, R. and J. Tirole (2002) "Self-Confidence and Personal Motivation," *The Quarterly Journal of Economics*, Vol. 117, No. 3, pp. 871–915.

Benabou, Roland and Jean Tirole (2006) "Belief in a just world and redistributive politics," *Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 699 – 746.

Benoît, Jean Pierre, Juan Dubra, and Don A. Moore (2015) "Does the better-than-average effect show that people are overconfident?: Two experiments," *Journal of the European Economic Association*, Vol. 13, No. 2, pp. 293–329.

Blanco, Mariana, Dirk Engelmann, Alexander K. Koch, and Hans-Theo Normann (2010) "Belief elicitation in experiments: is there a hedging problem?," *Experimental Economics*, Vol. 13, No. 4, pp. 412–438.

Bracha, Anat and Donald J. Brown (2012) "Affective decision making: A theory of optimism bias," *Games and Economic Behavior*, Vol. 75, No. 1, pp. 67–80.

Bridet, Luc and Peter Schwardmann (2014) "Selling dreams : Endogenous overoptimism and collateral use in financial contracts," *mimeo*, pp. 1–62.

Brunnermeier, Markus K and Jonathan A Parker (2005) "Optimal Expectations," *American Economic Review*, Vol. 95, No. 4, pp. 1092–1118.

Buser, Thomas, Leonie Gerhards, and Joël van der Weele (2018) "Responsiveness to feedback as a personal trait," *Journal of Risk and Uncertainty*, Vol. 56, No. 2, pp. 165–192.

Caplin, Andrew and John Leahy (2001) "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, Vol. 116, No. 1, pp. 55–79.

Carrillo, Juan D and Thomas Mariotti (2000) "Strategic Ignorance as a Self-Disciplining Device," *The Review of Economic Studies*, Vol. 67, No. 3, pp. 529–544.

Coutts, Alexander (2018) "Good news and bad news are still news: experimental evidence on belief updating," *Experimental Economics*, pp. 1–27.

Easterwood, John C and Stacey R Nutt (1999) "Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism?," *The Journal of Finance*, Vol. 54, No. 5, pp. 1777–1797.

Eil, David and Justin M Rao (2011) "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, Vol. 3, No. 2, pp. 114–138.

Eliaz, Kfir and Ran Spiegler (2006) "Can anticipatory feelings explain anomalous choices of information sources?," *Games and Economic Behavior*, Vol. 56, No. 1, pp. 87–104.

Ertac, Seda, Mert Gümren, and Levent Koçkesen (2017) "Strategic Feedback in Teams: Theory and Experimental Evidence," No. 111.

Fischbacher, Urs (2007) "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, Vol. 10, No. 2, pp. 171–178.

Gilboa, Itzhak and David Schmeidler (2001) *A Theory of Case-Based Decisions*, Cambridge: Cambridge University Press.

Grulich, Andrew (2000) "HIV risk behaviour in gay men: on the rise? : Monitoring risk behaviour and incidence of infection is essential.," *British Medical Journal*, Vol. 320(7248), pp. 1487–1488.

Heger, Stephanie A and Nicholas W Papageorge (2018) "We should totally open a restaurant: How optimism and overconfidence affect beliefs," *Journal of Economic Psychology*, Vol. 67, pp. 177–190.

Holt, Charles and Angela M. Smith (2009) "An update on Bayesian updating," *Journal of Economic Behavior & Organization*, Vol. 69, No. 2, pp. 125–134.

Karni, Edi (2009) "A Mechanism for Eliciting Probabilities," *Econometrica*, Vol. 77, No. 2, pp. 603–606.

Karni, Edi and Zvi Safra (1995) "The impossibility of experimental elicitation of subjective probabilities," *Theory and Decision*, Vol. 38, No. 3, pp. 313–320.

Koszegi, Botond (2010) "Utility from anticipation and personal equilibrium," *Economic Theory*, Vol. 44, No. 3, pp. 415–444.

Landier, Augustin (2000) "Wishful thinking a model of optimal reality denial," *mimeo*, pp. 1–25.

Landier, Augustin and David Thesmar (2009) "Financial Contracting with Optimistic Entrepreneurs," *Review of Financial Studies*, Vol. 22, No. 1, pp. 117–150.

Langer, Ellen J (1975) "The illusion of control.," *Journal of Personality and Social Psychology*, Vol. 32, No. 2, pp. 311–328.

Loewenstein, George (1987) "Anticipation and the valuation of delayed consumption," *The Economic Journal*, Vol. 97, No. 387, pp. 666–684.

Machina, Mark J. and David Schmeidler (1995) "Bayes without Bernoulli: Simple Conditions for Probabilistically Sophisticated Choice," *Journal of Economic Theory*, Vol. 67, No. 1, pp. 106–128.

Mayraz, Guy (2013) "Wishful Thinking in Predictions of Asset Prices," *mimeo*, pp. 1–35.

Mobius, Markus, Muriel Niederle, Tanya Rosenblat, and Paul Niehaus (2014) "Managing Self-Confidence : Theory and Experimental Evidence," *mimeo*, pp. 1–43.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey (2013) "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, Vol. 103, No. 2, pp. 804–830.

Puri, M and D Robinson (2007) "Optimism and economic choice," *Journal of Financial Economics*, Vol. 86, No. 1, pp. 71–99.

Schlag, Karl H., James Tremewan, and Joël J. van der Weele (2015) "A penny for your thoughts: a survey of methods for eliciting beliefs," *Experimental Economics*, Vol. 18, No. 3, pp. 457–490.

Spiegler, Ran (2008) "On Two Points of View Regarding Revealed Preferences and Behavioral Economics," in *The foundations of positive and normative economics: A handbook*: Oxford Univ. Pr. pp. 95–115.

Vosgerau, Joachim (2010) "How prevalent is wishful thinking? Misattribution of arousal causes optimism and pessimism in subjective probabilities.," *Journal of experimental psychology. General*, Vol. 139, No. 1, pp. 32–48.

Yariv, Leeat (2005) "I'll See It When I Believe It - A Simple Model of Cognitive Consistency," *mimeo*, pp. 1–43.