We Should *Totally* Open a Restaurant: Performance Uncertainty and Optimistic Beliefs*

Stephanie A. Heger†
Department of Economics, Washington University in St. Louis

Nicholas W. Papageorge‡
Department of Economics, Johns Hopkins University

November 6, 2013

ABSTRACT: To understand why people make poor decisions, such as opening up a restaurant that is bound to fail, economists have increasingly examined the possibility that individuals over-estimate their abilities or performance. Typically, this research occurs in settings where individuals have preferences over outcomes (indeed, the restauranteur likely wants his restaurant to succeed), but the role of preferences in shaping beliefs is ignored and the following distinction is overlooked. An individual can engage in *performance over-estimation*, where, independent of payoffs, he believes his performance is better than it actually is. Alternatively, an individual may exhibit *optimism*, where he simply over-weights the probability of outcomes he prefers. We design an experiment to isolate these phenomena and to study how they interact. We show that both optimism and performance over-estimation explain why people tend to over-estimate outcomes with high payoffs. We also show that the same people who over-estimate their abilities are also more likely to be optimistic. Finally, we find a gender difference in the source of beliefs leading to bad decisions. For men, over-estimation of high payoff outcomes is driven by performance over-estimation. For women, the driving factor is optimism.

KEYWORDS: Experiments, subjective beliefs, overconfidence, optimism, entrepreneurship.
JEL CLASSIFICATION: C91 D03 D84 L26.

---

*This project benefitted from many helpful comments and conversations. We particularly thank Barton Hamilton and David Levine for their continual support. Additionally, we thank: Sascha Baghestanian, Jorge Balat, Robert Barbera, Hülya Ereslan, Glenn Harrison, Dan Levin, Filippo Massari, Glenn MacDonald, Erkut Ozbay, Robert Pollak, Maher Said, Robert Slonim, Matthew Wiswall, and participants at the 2012 International ESA Conference at NYU, the 2013 SABE Conference in Atlanta, and the Sydney Experimental Brown Bag Seminar for their helpful comments. We gratefully acknowledge generous financial support from the Center for Research and Strategy (CRES) at Olin Business School, Washington University in St. Louis and the National Science Foundation via grant SES-0851315.
†[Corresponding author] Email: sheger@wustl.edu.
‡Email: papageorge@jhu.edu.

1 Introduction

Economists have increasingly considered the possibility that individuals make poor decisions because they believe their performance is better than it actually is. This phenomenon is often called *overconfidence* and has been linked to sub-optimal investment strategies (Malmendier and Tate, 2005), an over-willingness to compete (Niederle and Vesterlund, 2007) and unreasonably high rates of entrepreneurship (Camerer and Lovallo, 1999; Elston et al., 2006; Koellinger et al., 2007).[1] Research on overconfidence suggests that if individuals held more accurate beliefs about their own performance, then they would make better decisions. Indeed, when an individual is making a decision under uncertainty, it is his perception of the distribution over outcomes that drives choices, rather than an objective measure (Savage, 1954; Manski, 1993). The policy implication is clear: interventions designed to inform individuals could curtail costly behavior. However, this type of intervention could be misguided or ineffective absent a sharp understanding of what sort of information to provide—and to whom.[2]

A difficulty arises since previous research on overconfidence tends to attribute sub-optimal decisions solely to performance mis-estimation, but it does so in scenarios where individuals have preferences over outcomes. The role of preferences in shaping beliefs is largely ignored and the following distinction is therefore overlooked. An individual can engage in *performance over-estimation*, where, uncertain of his abilities and independent of his preferences over outcomes, he believes his performance is better than it actually is. Alternatively, given preferences over outcomes, an individual may exhibit *optimism*, where, irrespective of his abilities or performance, he simply overweights the probability of exogenously determined events that he prefers.[3] Ignoring this distinction could lead to ineffective information interventions. To fix ideas, consider a future failed bio-tech entrepreneur or restauranteur. Neither would gain much from an intervention informing them that their performance will suffer from a lack of technical expertise, training or talent if what drives their choices are optimistic beliefs about factors independent of their performance, like market demand for their product.

In this study, we conduct a laboratory experiment designed to achieve two goals. The

[1]Hamilton (2000) documents high failure rates in entrepreneurship and, among those who remain self-employed for ten years, low returns relative to paid employment. Moskowitz and Vissing-Jorgensen (2002) note that the self-employed, on average, bear higher risk without higher expected returns.

[2] Card et al. (2012) study how information has heterogenous effects in the context of relative income.

[3]Weinstein (1980) coined the term *unrealistic optimism*, which refers to individual probabilistic beliefs about exogenous events in which there is a favored outcome. Our notion of optimism as a pure mistake is most similar to the notion of *wishful thinking* studied in Mayraz (2011) and can be rationalized with *rank-dependent utility* (Quiggin, 1982).

first is to examine optimism and performance over-estimation as isolated phenomena and to test whether they are correlated at the individual level. The second is to determine how they interact to jointly influence decisions in more realistic settings in which individuals simultaneously face performance uncertainty and preferences over outcomes.

The experiment is described in detail in Section 4, but its main features can be explained using the example of a coin toss. First, we elicit an individual's probabilistic belief that a fair coin lands on heads. Second, we inform the individual that he will receive a side payment if heads occurs and again elicit his probabilistic belief about the likelihood of heads. Optimism is identified by comparing the individual's probabilistic beliefs when heads is payoff favorable versus when it is not. A subject that reports $\frac{1}{2}$ when heads is not payoff favorable and reports a number greater than $\frac{1}{2}$ when heads is payoff favorable is classified as optimistic. The reason is that the reported probability of the outcome was higher when that outcome was payoff favorable.

To identify performance overestimation, we use a similar technique. First, we elicit the likelihood of heads in a fair coin toss. Second, the individual answers a trivia question. A correct answer results in a coin that has $\frac{2}{3}$ chance of heads and $\frac{1}{3}$ chance of tails, whereas an incorrect answer results in the fair coin. In other words, the individual's performance directly affects the distribution he faces—in this case increasing the probability of heads. Without feedback on his performance, the individual forms a belief about his performance and we elicit his belief about the probability of heads. Suppose he did not answer the trivia question correctly, but thinks there is some possibility that he might have. Then, when asked to report his belief of the probability of heads, his uncertainty may be reflected by a report between $\frac{1}{2}$ and $\frac{2}{3}$. If so, then we say he over-estimates his performance. Again, we rely upon an individual-level control. Had the individual answered $\frac{2}{3}$ regardless of the role of performance, he would not be counted as having over-estimated his performance.

The economic environments we are interested in (e.g., entrepreneurship, investment strategies and decisions to compete) are those where the individual's performance affects the likelihood that a preferred outcome is realized. Thus, individual decisions may be influenced by performance over-estimation and optimism simultaneously. We simulate an environment in the laboratory where the individual can increase the probability of high payoff outcomes through better performance. This environment is called the *combined setting* since the subject can increase the probability that the coin lands on heads through his performance (as in the performance treatment) and receives an additional side payment when heads is realized (as in the payment treatment). Having examined performance over-estimation and optimism in isolation, we can examine to what degree beliefs in the combined setting are attributable to either optimism or performance over-estimation (or both).

In our experiment, to operationalize the intuition provided by the coin toss, we elicit beliefs over a more general set of distributions. We ask subjects to report beliefs about the number of white balls that will be drawn from jars containing various compositions of white and black balls. This experiment is described in detail in Section 4.[4] Further, our technical definitions of optimism and over-confidence appeal to stochastic dominance so that they can accommodate distributions more general than the coin toss. These definitions are presented in Section 2. According to this design, an individual is labeled optimistic if his subjective beliefs on the number of white balls drawn first-order stochastically dominate the objective distribution only when he is induced to prefer white balls because of the side payment. This generalizes the notion of overweighting exogenously determined, favorable outcomes. Similarly, performance over-estimation is isolated by studying changes under the two levels of the performance treatment.

Our study results in three key findings. First, individuals who over-estimate their performance are more likely to be optimistic. The within-subject correlation is 0.48. This finding could explain puzzling behavior like entrepreneurial entry despite low returns. Individuals who over-estimate their performance as business-owners are more likely to be optimistic, which may nudge them further towards entrepreneurship. Second, we examine to what degree shifts in beliefs in the combined setting can be explained by shifts in the payment and performance treatments. This amounts to decomposing responses in the combined setting into optimism versus performance over-estimation and we find that both have significant explanatory power. Moreover, since they are correlated, it would be misleading to explain responses in the combined setting solely as performance over-estimation or optimism.[5] To understand why, suppose a policy aimed at providing information to aspiring entrepreneurs is based on the observation that entrepreneurs are more likely to over-estimate their performance as business owners. This policy would be misguided since it ignores how individuals who over-estimate their performance are also likely to be optimistic, which may also encourage entry into self-employment. Third, although we find no evidence of gender differences in average treatment effects, we do find a striking gender difference in the decomposition of beliefs in the combined setting, i.e., in the degree to which responses in the combined setting can be explained by optimism versus performance over-estimation. For males, we find that over-estimation of high payoff outcomes is explained by performance over-estimation. For females, over-estimation is better explained by optimism.

[4] The counterpart to the coin toss—one white ball and one black ball with a single draw—is one of the distributions subjects face.

[5] More precisely, if we regress treatment effects in the combined setting onto treatment effects in either the payment and performance treatments, estimated coefficients are biased since they capture the positive correlation in optimism and performance over-estimation.

Over-estimation of high payoff outcomes may lead to costly decisions. Thus, information that helps the individual make better-informed decisions could be valuable. Given the proper information, the future failed bio-tech entrepreneur or restauranteur, for example, could avoid the psychological trauma of failure, unemployment, a slowed accumulation of human capital and financial ruin.[6] In fact, an emerging literature has shown that information interventions can be highly effective in several contexts, including students' perceptions of the returns to schooling (Manski, 1993; Jensen, 2010; Wiswall and Zafar, 2011), 401(K) participation (Clark et al., 2013), and entrepreneurship (Gompers et al., 2005; Nanda and Sørensen, 2010).[7]

However, in many of these contexts, the exact type of information that is transmitted is unclear. Lerner and Malmendier (2011) look at the effect of exposure to peers with entrepreneurial experience within the randomly assigned sections in the Harvard MBA program. They find that exposure to a higher share of classmates with an entrepreneurial background decreases entry into entrepreneurship and that decreased entry is driven by fewer unsuccessful endeavors. They speculate that the information provided addresses performance beliefs via feedback on potential business ideas, but cannot rule out that greater exposure to macro-level information on the widespread challenges of owning a business plays a role. In fact, Minniti (2005) suggests that the exposure to entrepreneurial peers reduces the ambiguity of the market environment and subsequently nurtures further entrepreneurial activity. Given our results, the lack of clarity about which type of information is most valuable in fostering successful entrepreneurship is not surprising. Individuals who over-estimate their performance may also be more likely to believe the market environment is in their favor. Moreover, our results suggest that men and women may respond best to very different types of information.

The remainder of this paper proceeds as follows. In Section 2, we discuss how optimism and overconfidence have been treated in previous literature. In Section 3, we provide technical definitions. In Section 4, we outline our experiment. Section 5 presents results and Section 6 discusses implications of our results and directions for future research.

[6] We must take care in discussing welfare, however, since some research claims that failure in entrepreneurship, in part due to optimistic beliefs, is an integral part of an economy with a high growth rate (Kaldor, 1954). Therefore, it is unclear whether information that ameliorates inflated beliefs about success would benefit anyone beyond the entrepreneur himself—and, perhaps, his investors.

[7] Consistent with our findings on gender differences, Nguyen (2010) shows that the effectiveness of the information intervention depends on the characteristics of the student. This point is echoed in Fréchette et al. (2011), who show that personality traits can affect the type of information individuals demand.

2 Overconfidence and Optimism

This section provides an overview of previous work on optimism and overconfidence and how our work contributes to it. Economists have ascribed a wide range of behaviors to individual overconfidence (Moore and Healy, 2008).[8] One of the most widely-cited findings on overconfidence is attributed to Svenson (1981), who shows that 81% of subjects report themselves to be safer drivers than the median driver. Hoelzl and Rustichini (2005) move away from stated responses and study overconfidence using revealed beliefs. They allow subjects to vote on the method of payment—to win money if their performance is ranked in the top half of their group or to be paid based on a random draw—and find that 55% of subjects preferred to be paid based on their rank. Building on Hoelzl and Rustichini (2005), Blavatskyy (2009) and Urbig et al. (2009) also find evidence of overconfidence in a setting where subjects can choose to be paid according to either their performance or a random lottery, where the two options have equal expected value.[9] It is important to note, that these designs rely on the previously untested hypothesis that overconfidence and optimism are not correlated at the individual level.

A related set of papers examines *unrealistic optimism* which focuses on the relative assessment of favorable events (Weinstein, 1980).[10] A general conclusion is that individuals tend to believe favorable events are more likely to occur to them than to their peers, although this bias is attenuated when individuals have less control over the events (Alicke, 1985). More similar to our approach, Ito (1990) and Mayraz (2011) suggest that there is empirical and experimental evidence in support of pure wishful thinking. Our paper aims to unify the literature on optimism and overconfidence by studying both concepts within a single decision-making context, which highlights that they are separate phenomena that potentially interact to drive decisions.

3 Definitions

The agent in our setting faces two dimensions of uncertainty: uncertainty about his performance represented by a random type $\theta \in \Theta$ and uncertainty about the outcome of a lottery

[8] Following (Moore and Healy, 2008), it is widely acknowledged that there are three ways to describe overconfidence: over-estimation (in an absolute sense), over-placement (over-estimation relative to a reference group) and over-precision, where individuals think their information is more precise than it actually is.

[9] Further contributions include: Russo and Schoemaker (1992); Klayman et al. (1999); Kirchler and Maciejovsky (2002); Moore and Healy (2008); Clark and Friesen (2009) and Fang and Moscarini (2005).

[10] This line of work differs from studies where optimistic beliefs and biased expectations are rationalized (Van den Steen, 2004; Brunnermeier and Parker, 2005; Santos-Pinto and Sobel, 2005; Köszegi, 2006).

X with support $[0, N]$, where N is a positive integer. The outcome of the lottery, denoted x, depends on the agent's performance type and is described by a distribution, which we denote:

$$G_X(x|\theta) = P(X \leq x|\theta). \quad (1)$$

We assume that the measure of performance and the outcome of the lottery are ordered and discrete. The link between performance type, θ, and the distribution $G_X(x|\theta)$ is given by the following: if $\theta' > \theta''$ then $G_X(x|\theta = \theta')$ first order stochastically dominates $G_X(x|\theta = \theta'')$. In other words, better performance is associated with a "better" distribution, in the sense of stochastic dominance. The agent has a true performance type, $\theta_o \in \theta$, and the distribution Θ is degenerate at θ_o, but θ_o is unknown to the agent. The agent has beliefs about his true performance type, θ_o, given by the distribution $\tilde{\Theta}$ over θ.

There is also a non-decreasing map from lottery outcomes to monetary payoffs, $m(\cdot)$, meaning that higher values of x have weakly larger monetary payoffs. Moreover, we assume that the agent is an expected utility maximizer and with preferences over monetary outcomes, such that if $x' > x''$ then $u(m(x')) > u(m(x''))$. The agent has subjective beliefs about G that are given by

$$F\left(x|m(\cdot); \tilde{\Theta}\right), \quad (2)$$

where we explicitly condition on $m(\cdot)$ to capture that an agent's beliefs may be affected by the payoff function.

Consider the case in which there is a single type, θ_o, and so the agent does not face performance uncertainty and cannot affect the distribution he faces via his performance. In other words

$$G_X(x|\theta) = G_X(x), \quad (3)$$

which captures the idea that the distribution is exogenous to the agent's performance. Further, consider the following two cases. First, if there is no performance uncertainty and the monetary payoff is constant (does not depend on x), we write $m(x) = m$. Then the agent has subjective beliefs given by

$$F\left(x|m(x) = m; \tilde{\Theta}\right) = F(x), \quad (4)$$

where we do not conditional on $m(\cdot)$ since the payoff is independent of x and we do not condition on $\tilde{\Theta}$ since there is no performance uncertainty and thus no uncertainty regarding the distribution the agent faces. Call this set of beliefs the *neutral-exogenous beliefs*, where we say *neutral* since the agent's payoff is independent of the outcome x and we say *exogenous* since the agent's belief is about an exogenously given distribution, i.e., one that is not influenced by his performance. Second, if there is no performance uncertainty and the monetary payoff is a weakly increasing in x, then the agent has subjective beliefs given by

$$F\left(x|m(\cdot);\tilde{\Theta}\right) = F(x|m(\cdot)), \tag{5}$$

where we explicitly conditional on $m(\cdot)$ since the payoff changes with x and might therefore influence beliefs about x. Call this set of beliefs the *payoff induced preference-exogenous beliefs*, where *payoff induced preference* means that the payment scheme has induced the agent to prefer one set of outcomes over others. An agent is optimistic (pessimistic) if he overweights (underweights) outcomes only because the outcomes are economically favorable. Formally:

Definition 1. *An agent is **optimistic** if:*

1. *$F(x|m(\cdot))$ stochastically dominates $G_X(x)$ and*
2. *$F(x)$ does not stochastically dominate $G_X(x)$.*

Definition 2. *An agent is **pessimistic** if:*

1. *$F(x|m(\cdot))$ is stochastically dominated by $G_X(x)$ and*
2. *$F(x)$ is not stochastically dominated by $G_X(x)$.*

Our definitions of optimism and pessimism require a very specific change in beliefs about $G_X(x)$ when higher outcomes are economically favorable versus when they are not. The definitions capture the idea that the optimistic agent places higher weight on larger outcomes x only when they are economically favorable to the agent. If the agent's neutral-exogenous and induced preference-exogenous beliefs both stochastically dominate $G_X(x)$, then the tendency to overweight higher outcomes is not induced by the favorable payoff of the larger outcomes and the agent is not deemed optimistic.

To fix ideas, we return to our coin toss example. Consider two independent tosses of a fair coin, where $x \in \{0, 1\}$ and $x = 1$ represents the outcome "heads" and hence, for any given toss,

$$\begin{aligned} G(0) &= 0.50 \\ G(1) &= 1. \end{aligned}$$

For the first toss, the agent is paid $m(x) = 0 \ \forall \ x$ and his corresponding beliefs are given by

$$\begin{aligned} F(0|m(x) = 0) &= 0.50 \\ F(1|m(x) = 0) &= 1. \end{aligned}$$

For the second toss, $m(x) = x, \ \forall \ x$ and his beliefs are given by

$$\begin{aligned} F(0|m(x) = x) &= 0.40 \\ F(1|m(x) = x) &= 1. \end{aligned}$$

Thus, the optimistic agent believes heads is more likely only when heads is payoff favorable. Alternatively, compare this to the following set of beliefs. Suppose the neutral-exogenous beliefs are given by

$$\begin{aligned} F(0|m(x) = 0) &= 0.45 \\ F(1|m(x) = 0) &= 1 \end{aligned}$$

and the induced preference-exogenous beliefs given by

$$\begin{aligned} F(0|m(x) = x) &= 0.40 \\ F(1|m(x) = x) &= 1. \end{aligned}$$

An agent with this set of beliefs over-weights the likelihood of heads relative to $G(x)$ regardless of whether heads is payoff favorable. Hence, he is not classified as optimistic under our definition. However, the agent's beliefs do shift in response to a change in $m(\cdot)$. This change in closely related to our definition of optimism and will be further discussed in Section 3.1.

Next we formalize beliefs when there is performance uncertainty. We drop the assumption that there is a single performance type and allow for subjective uncertainty regarding performance type. Recall, θ_o denotes the agent's true type, which is unknown to the agent. $\tilde{\Theta}$ denotes the agent's beliefs about the degenerate distribution over θ, Θ. The agent is tasked with forming beliefs over:

$$G(x|\theta_o) \tag{6}$$

In this case we allow G to change with performance as described above: if $\theta' > \theta''$ then $G(x|\theta = \theta')$ stochastically dominates $G(x|\theta = \theta'')$. In other words, when an agent performs better, he faces a better distribution in the sense of stochastic dominance. The distribution is therefore endogenously determined by the agent's level of performance. When the agent faces performance uncertainty and a monetary payoff function that is constant, his beliefs are given by

$$F\left(x|m(x) = m; \tilde{\Theta}\right) = F\left(x|\tilde{\Theta}\right) \tag{7}$$

where we do not condition on $m(\cdot)$ as payoffs are constant and do not depend on x. These beliefs are called *neutral-endogenous beliefs* as they arise when an agent is not induced to have preferences over any outcomes, but faces a distribution that is endogenously determined by his performance. An agent can over- or under-estimate his true performance. Formally:

Definition 3. *An agent exhibits* ***performance over-estimation*** *if:*

1. $F\left(x|\tilde{\Theta}\right)$ *stochastically dominates* $G\left(x|\theta_o\right)$ *and*

2. $F\left(x\right)$ *does not stochastically dominate* $G\left(x\right)$.

In other words, an agent over-estimates his performance when his beliefs stochastically dominate G when G is a function of his performance, but his beliefs do not stochastically dominate G when G is independent of his performance. Similarly:

Definition 4. *An agent exhibits* ***performance under-estimation*** *if:*

1. $F\left(x|\tilde{\Theta}\right)$ *is stochastically dominated by* $G\left(x|\theta_o\right)$ *and*

2. $F\left(x\right)$ *is not stochastically dominated by* $G\left(x\right)$.

Note that in defining beliefs about performance the payoff function is held constant. This allows us to isolate beliefs about performance from optimism or pessimism.

Finally, when the agent simultaneously faces performance uncertainty and a monetary payoff that is weakly increasing in x, then he must form beliefs over

$$G\left(x|\theta_o\right). \tag{8}$$

His beliefs are given by

$$F\left(x|m\left(\cdot\right), \tilde{\Theta}\right). \tag{9}$$

These beliefs are called *payoff induced preference-endogenous beliefs* since they reflect the agent's beliefs when both performance uncertainty and a preference for larger realizations x are present. It is important to note that in this setting, the agent can affect the lottery he faces with his performance and that larger values of x are payoff favorable. This means that the agent increases the likelihood of a larger monetary payoff through better performance.

3.1 Shifts in Beliefs

Our definitions of optimism, pessimism, performance over- and under-estimation are dichotomous and cannot capture important changes in beliefs that may occur. Consider again the following set of beliefs from the coin toss example: the neutral-exogenous beliefs

$$\begin{aligned} F\left(0|m\left(x\right)=0\right) &= 0.45 \\ F\left(1|m\left(x\right)=0\right) &= 1 \end{aligned}$$

and the induced preference-endogenous beliefs

$$\begin{aligned} F(0|m(x)=x) &= 0.40 \\ F(1|m(x)=x) &= 1 \end{aligned}$$

This agent is not classified as optimistic according to definition 1, and yet his beliefs shift when larger outcomes become economically favorable. To capture this additional variation in beliefs, we will consider a set of measures based on average agent deviations from the neutral belief and we will refer to these average deviations as "shifts." The shift measurements capture changes in the agent's beliefs that occur in response to payoff induced preferences, performance uncertainty or both. Formally:

Definition 5. *An agent displays a shift towards optimism (pessimism) when*

$$\frac{1}{N}\Sigma_{x=0}^{N}[F(x|m(x)=m)-F(x|m(\cdot))] > (<)\, 0. \tag{10}$$

Definition 6. *An agent displays a shift towards performance overestimation (underestimation) when*

$$\frac{1}{N}\Sigma_{x=0}^{N}[F(x|m(x)=m)-F\left(x|m(x)=m;\tilde{\Theta}\right)] > (<)\, 0. \tag{11}$$

Definition 7. *An agent displays a shift towards overestimation and/or optimism (underestimation and/or pessimism) in the combined setting when*

$$\frac{1}{N}\Sigma_{x=0}^{N}[F(x|m(x)=m)-F\left(x|m(\cdot),\tilde{\Theta}\right)] > (<)\, 0. \tag{12}$$

3.2 Within-Subject Variation

Finally, we address how an agent's behavior might change across lotteries and across payment maps $m(\cdot)$. Essentially, we put no restriction on agent behavior across any two lotteries. Thus, an agent can be optimistic when facing a lottery over X and not optimistic over lottery X'. However, one could consider the following sort of restriction on agent behavior across payment maps. Specifically, consider two payment schemes $\tilde{m}(x)$ and $\hat{m}(x)$, where

$$\frac{\partial \hat{m}(x)}{\partial x} > \frac{\partial \tilde{m}(x)}{\partial x}.$$

Thus, $\hat{m}(x)$ is more sensitive to increases in x. A natural hypothesis to test might be the following: if an agent is optimistic under $\tilde{m}(x)$ then he is also optimistic under $\hat{m}(x)$.[11]

[11] The payment scheme used in this paper is not continuous in x, but the underlying intuition remains the same.

Similarly, if

$$\frac{1}{N}\Sigma_{x=0}^{N}[F\left(x|\tilde{m}\left(x\right)=m\right)-F\left(x|\tilde{m}\left(\cdot\right)\right)]=\tilde{k}>0$$

, then

$$\frac{1}{N}\Sigma_{x=0}^{N}\left[F\left(x|\hat{m}\left(x\right)=m\right)-F\left(x|\hat{m}\left(\cdot\right)\right)\right]\geq\tilde{k}>0.$$

That is, if an agent shifts towards optimism under $\tilde{m}(\cdot)$, then his optimistic shift under $\hat{m}(\cdot)$ must be at least as large. Restrictions on pessimistic beliefs would be analogous.

Our discussion of our results will include remarks on within-agent variation across lotteries. However, since we do not vary the payment scheme in the the experiment we conducted for this paper, we cannot directly test agent behavior across payment maps. Instead, we see the hypotheses discussed in this section as an interesting and straightforward extension of our existing project.[12]

4 Experimental Design

We employ a 2×2 within-subject experimental design, which means that each subject faces each of the four treatment combinations. The design is illustrated in Figure 1. In each of the four treatment combinations, subjects face computerized "jars" containing various compositions of white and black balls.[13] Subjects are told the number of white and black balls in each jar and the number of balls that will be drawn from the jar and they are asked to report cumulative probabilities. On each screen, the computerized jar is displayed on the left side and a series of questions about the jar on the right side. Subjects move the cursor to indicate a percent chance of a certain number of white balls being drawn from the jar. The numerical value indicated by position of the cursor is displayed next to the number line. To ensure truthful reporting, subjects are always paid according to the widely employed quadratic scoring rule (QSR) (Brier, 1950; Murphy and Winkler, 1970), where the score is

[12] It is important to note that Mayraz (2011) finds that optimism in his subject pool did not significantly increase when the monetary rewards increased, which suggests that subject behavior may not react strongly to changes in $m(x)$.

[13] Figure 1 at the end of the instructions is a screen shot of the computerized interface the subject uses during the experiment.

computed as follows:[14]

$$\text{SCORE} = \begin{cases} 10 - 10 * [\text{reported belief} - 1]^2 & \text{if event occurs} \\ 10 - 10 * [\text{reported belief} - 0]^2 & \text{if event does not occur.} \end{cases}$$

Following procedures in Sonnemans and Offerman (2001), subjects are given a table that explains how reported beliefs translate to payoffs via the QSR, and are required to demonstrate full comprehension of the QSR by correctly answering comprehension questions before proceeding with the experiment. Subjects are also always informed that they can expect to make the most money by accurately reporting their probabilistic beliefs.[15] Subjects receive paper copies of the instructions. The instructions are read aloud to the subjects and projected on the screen at the front of the room. Each of the four tasks is presented separately, followed by a quiz to ensure full comprehension. Subjects cannot proceed to the task until all comprehension questions are answered correctly. In what follows, we describe each treatment separately and then go on to discuss the treatment effect measures we construct for use in subsequent analysis.

The two treatments are the *payment treatment* and the *performance treatment*. The payment treatment concerns the function $m(x)$, which maps outcomes of a lottery, x, to monetary payoffs, varying whether or not white balls are payoff favorable. The performance treatment concerns θ and varies whether performance affects the distribution faced by the subject or whether there is no performance uncertainty and the distribution is exogenously given. The 2×2 design exploits within-subject variation and allows for an individual-level control, the Neutral Exogenous setting, where neither the payment nor performance treatment is applied. In this setting, we measure subject beliefs in the absence of both performance uncertainty and preferences over outcomes. Treatment effects are changes in beliefs relative to beliefs elicited in the Neutral Exogenous setting. Table 1 maps treatments to beliefs and to the measured experimental treatment effects.

[14] The quadratic scoring rule is only incentive compatible for a risk-neutral expected utility maximizer. Risk aversion causes subjects' probabilistic reports to tend towards 0.5. This tendency towards 0.5 would occur in each treatment as subjects are incentivized with the QSR throughout the experiment. The use of an individual-level control in our analysis, as Section 4.4 will make clear, means that any factor that affects individual reports uniformly across treatments does not drive our results. An alternative, a binary lottery implementation of the quadratic scoring rule, is, in theory, incentive compatible and robust to risk preferences (McKelvey and Page, 1990), but evidence suggests that it may not induce risk-neutrality as suggested (Selten et al., 1999) and the cognitive burden imposed on subjects may result in less reliable reports than use of the deterministic quadratic scoring rule (Rabin and Thaler, 2001).

[15] A copy of the full instructions can be found at www.stephanieheger.com/research.html. The "General Setting" describes features that are common to all tasks.

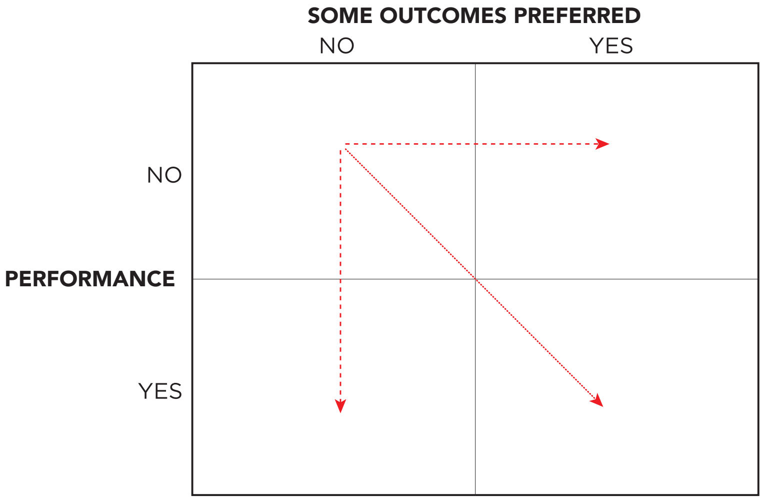


Figure 1: Experimental Treatments. Each cell in the box represents a level of one of the experimental treatments and the arrows represent the treatments. The *payment treatment* varies whether white balls are payoff favorable. The *performance treatment* varies whether the distribution is exogenously determined or endogenously determined by the subject's performance (correctly answered trivia questions).

Table 1: Treatment, Beliefs and Treatment Effects

Treatments	Beliefs	Treatment Effects
No Payment & No Performance	Neutral-Exogenous	
Payment Only	Induced Preference-Exogenous	Payment Effect
Performance Only	Neutral-Endogenous	Performance Effect
Payment & Performance	Induced Preference-Exogenous	Combined Effect

This table summarizes how the four experimental settings map into the beliefs we elicit and resulting treatment effects. Treatment effects are changes in beliefs relative to beliefs elicited in the Neutral Exogenous setting. For example, if only the performance treatment is applied (Performance Only), we elicit Neutral-Endogenous beliefs. By comparing these elicited beliefs with the Neutral Exogenous beliefs (No Payment & No Performance), we identify the Performance Effect.

4.1 The Payment Treatment

The *payment treatment* varies whether subjects have a preference over the outcome of the random draw from the jar. The neutral payment scheme occurs when the payment treatment is not applied and $m(x) = 0$. Thus, subjects are not induced to prefer that white balls, rather than black balls, be drawn from the jar because their payment is independent of the outcome. Alternatively, the induced preference scheme occurs when the payment treatment is applied and $m(x)$ is weakly increasing in x, and thus, subjects are induced to have preferences over the outcome. Specifically, subjects prefer that white balls be drawn from the jar since their expected payoff is now increasing in the number of white balls.

$$\text{m(x)} = \begin{cases} 100 & \text{with probability } \frac{\text{total white}}{\text{max white}} \\ 0 & \text{with probability } \left(1 - \frac{\text{total white}}{\text{max white}}\right). \end{cases}$$

The induced preference payment scheme differs only in that it pays an additional lottery ticket in which the probability of winning the prize increases with the number of white balls drawn during the treatment. This additional payment induces subjects to prefer higher realizations of white balls without interfering with the incentive to accurately report beliefs. Under the neutral payment scheme, subjects are not induced to prefer white balls. The neutral scheme therefore provides an individual-level control against which we can measure the effect of inducing preferences over realizations on agent beliefs about a given distribution. In other words, changes in probabilistic reports for a given distribution between the neutral and the induced preference payment scheme are attributable to induced preferences for white balls to be drawn.[16]

4.2 The Performance Treatment

In the *performance treatment* subjects form beliefs about distributions, the parameters of which are either *exogenous* or *endogenous*. If the distribution is exogenous, the experimenter determines the composition of white and black balls in the jar and informs subjects. In the endogenous treatment, subjects are likewise informed of the composition, but can increase the number of white balls in their jar through their performance on a trivia task. For example, a subject faces a jar with 1 white ball and 1 black ball (where 1 ball will be drawn). The subject then answers a trivia question. If he answers the question correctly then another

[16] Under the induced preference payment scheme, a risk-averse subject would want to use the additional side payment for white balls to hedge their risk. This would result in an under-reporting of probabilities and consequently a subjective distribution of white balls that was shifted further to the left. This means that agents may actually be more optimistic than we find.

white ball is added to his jar. If he answers the question incorrectly then no white ball is added to his jar. Subjects are not told whether their answer is correct, but are informed of the possible distributions that can be faced given each level of performance.[17] Table 2 outlines how a subject's performance determines the distribution he faces. A correct trivia answer always adds one additional white ball to the jar.

Table 2: Experimental Design and Distributions

Distribution Details				Performance
Label	White	Black	Draws	Treatment
1	1	1	1	↓
2	2	1	1	
3	1	3	3	↓
4	2	3	3	
5	3	3	3	
6	4	3	3	

In the *performance treatment*, one white ball is added to a starting distribution for each correctly answered trivia question such that subjects effectively face a different distribution. Subjects starting with distribution 1 answer 1 trivia question, adding 1 white ball if correct and therefore facing distribution 2. If they answer the question incorrectly, no white ball is added and they face distribution 1. Subjects starting with distribution 3 answer 3 trivia questions and may add 0, 1, 2 or 3 white balls, therefore facing distributions 3, 4, 5, or 6.

To gauge performance in the endogenous setting, subjects answer multiple choice questions from the Mensa Quiz book (Grosswirth et al., 1999).[18] Multiple choice questions are chosen to avoid open-ended questions and subject confusion. The Mensa Quiz book also reports the percentage of quiz takers that answered a given question correctly. This allows us to select questions of similar difficulty level, controlling for any complications that may arise from the documented "hard-easy" effect (Lichtenstein and Fischhoff, 1977).[19]

[17]Denote $\hat{p}$ as the subject's beliefs that he answers the trivia question correctly. Then he faces a jar with 2 white balls and 1 black with the probability $\hat{p}$ and a jar with 1 white ball and 1 black ball with probability $1-\hat{p}$. We note that a subject might purposefully choose an answer that he knows is incorrect rather than choose from several possibly correct options as doing so would increase her certainty over the distribution he faces at the cost of losing the payoff from a possible correct answer. Payoffs are structured so that this behavior is not incentivized, i.e., gains from correctly answered questions in the performance treatment are large relative to losses from incorrect guesses about the numbers of balls drawn.

[18]This task was also used in Owens et al. (2012); Grossman and Owens (2012).

[19]In particular, we choose answers that were answered correctly by between 45% and 55% of quiz-takers, according to the Mensa Quiz book.

4.3 Combining the Payment and Performance Treatments

Returning to Figure 1, it should be clear that the two treatments at two levels results in four sets of beliefs which correspond to our definitions stated in Section 3: neutral-exogenous, induced preference-exogenous, neutral-endogenous and induced preference-endogenous. The fourth combination, induced preference and endogenous, identifies the combined effects of the payment and performance treatments. Simultaneous exposure to both treatments is meant to mimic real-world decision-making scenarios where individual beliefs are simultaneously influenced by preferences over outcomes and by uncertainty about their own performance. Focusing solely on the combined effects, however, precludes identification of the effect of the independent application of the payment and the performance treatments. For example, if under the combined effect a subject reports an increase in the probability that white balls are drawn, it is unclear why. Perhaps the individual over-estimated his performance in answering trivia questions and therefore believes that there are more white balls in the jar than are actually there. Alternatively, the individual may be optimistic and therefore over-weights the probability that his payoff-favorable outcome (white balls being drawn) is realized. It is possible that both influence his report. Our design allows for the identification of the source of the over-estimation—both in isolation and then together with the combined effect—and thereby permits a credible analysis of the composition of beliefs observed in real-world settings.

4.4 Identification and Outcome Variables

The data collected during the experiment are subjective (reported) cumulative distributions under decision-making scenarios corresponding to the levels of the aforementioned treatments.[20] Using these data, we construct two sets of variables that exploit our within-subject control. The first is a categorical variable denoted $T_{i,d,\tau}$, which directly tests the definitions offered in Section 3 for subject i facing distribution $d \in \{1, 2, 3, 4, 5, 6\}$ under treatment $\tau \subseteq \{\text{Payment, Performance}\}$. τ is an improper subset of the set {Payment, Performance} to account for how the payment and performance treatments can be applied separately or simultaneously, the latter case corresponding to the combined setting. If $\tau = \text{Payment}$, $T_{i,d,\tau}$ classifies the subject as optimism, pessimistic or neither based on the stochastic dominance relation between the subjective and objective distributions. If $\tau = \text{Performance}$, $T_{i,d,\tau}$ clas-

[20]A total of 42 observations (2.14% of the sample) describing subjective distributions that violated monotonicity were therefore discarded. Subjective distributions with corresponding probability mass functions that did not sum to 1 (429 observations or about 22.37% of the sample after discarding non-monotonic subjective distributions) were normalized as propensity scores to ensure summation to 1.

sifies the subject as a performance over-estimator, under-estimator or neither. In particular,

$$T_{i,d,\tau} = \begin{cases} > 0 & \text{if} \quad \text{beliefs first order stoch. dominate under } \tau\text{, but do } \textit{not} \text{ under neutral \& exogenous} \\ < 0 & \text{if} \quad \text{beliefs are first order stoch. dominated under } \tau\text{, but are } \textit{not} \text{ under neutral \& exogenous} \\ 0 & \text{if} \quad \text{otherwise.} \end{cases}$$

Although $T_{i,d,\tau}$ classifies individuals according to our definitions, as discussed in Section 3 it does not capture relevant shifts in beliefs in response to the two treatments. Moreover, it would be difficult to interpret coefficients using $T_{i,d,\tau}$ as an outcome variable in regression analysis since it is a categorical variable. Therefore, we construct a second outcome variable that measures the average deviation of the elicited distribution when each treatment is applied (payment, performance or both combined) from the elicited distribution under the neutral-exogenous level (when no treatment is applied). This measure is constructed at the individual-distribution-treatment level and compares the average shift in each subject's beliefs when facing the same distribution under different treatment levels. More specifically, we define $\Delta_{i,d,\tau}$ for each individual i, distribution d and treatment $\tau \subseteq \{\text{Payment, Performance}\}$, where:

$$\Delta_{i,d,\tau} \equiv \frac{1}{N} \Sigma_{x=0}^{N-1} (F(x|\tau = \text{Neutral-Exogenous}) - F(x|\tau)).$$

If $d \in \{1, 2\}$ (distributions 1 or 2), then there is 1 moment ($N = 1$) where $x \in \{0\}$. If $d \in \{3, 4, 5, 6\}$, then there are 3 moments ($N = 3$) where $x \in \{0, 1, 2\}$. Here, notice that $\Delta_{i,d,\tau=\text{Payment}} > 0$ is indicative of an optimistic shift: absent performance uncertainty, the subjective distribution elicited when white balls are preferred is to the right of the subjective distribution elicited when there white balls are not preferred. The analogous statements are true for performance over-estimation when $\Delta_{i,d,\tau=\text{Performance}} > 0$.

4.5 Distribution Details

At the exogenous levels, each subject faces six different distributions, distinguished by the color of balls in a jar and the number of draws. These distributions are described in Table 2. When the performance treatment is applied, the distribution the subject faces is endogenous to his performance and ex-ante, we do not know how the subject will perform and thus which distributions he will ultimately face. But, since our analysis relies heavily upon the comparability of beliefs across treatments, it is necessary that the subject face each potential distribution in the exogenous setting that he may face in the endogenous setting. As Table 2 makes clear, the six distributions that the subject faces in the exogenous setting is the full set of distributions that the subject may also ultimately face in the endogenous setting.

4.6 Outline of Experimental Procedures

Finally, and for clarity, we now offer a detailed description of the experimental procedure from the perspective of the subject. Subjects enter the laboratory and are assigned to a computer terminal. At the terminal, subjects find copies of the "General Setting" instructions, a quadratic scoring rule "cheat sheet", and blank paper and pen for their own calculations. Once the experiment begins, the General Setting instructions are read aloud by an experimenter and projected on a screen at the front of the room. Subjects are informed that throughout the session they will complete 4 tasks, all of which will include answering questions about the chance that white balls are drawn from jars containing black and white balls. Further, they are informed that each of the four tasks differs slightly from the other three and that those differences will be explained before each new task begins. They are also informed that they can always expect to earn the most money by "making their best guess." In fact, this statement is repeated during the instructions for each task. Subjects are then told that their responses from one of the four rounds will be randomly chosen to determine their payment, but they will not know which round is chosen until the end of the session.

While the General Setting instructions are read, subjects are familiarized with the computer interface. The experimenter also discusses basic probability rules with the subjects. After the General Setting instructions are finished, subjects complete a comprehension quiz on the instructions. Once all subjects have entered the correct answers, the experimenter explains the correct answers aloud. A comprehension quiz is given after every set of task-specific instructions to highlight the similarities and differences across tasks. The experimenter then hands out the set of instructions for the first task, the task which elicits beliefs in the neutral-exogenous setting. Participants are informed that the first task exactly mirrors the General Setting. Subjects begin the first task and face each of the six distributions in a random order. They have 2 minutes to answer questions about each distribution. No feedback is given during the first task. When the first task is completed, the Task 1 instructions and the working paper are collected and Task 2 instructions and clean working paper are distributed.

The experimenter then reads Task 2 instructions aloud. Subjects are informed that the only difference between Task 1 and Task 2 is that they would receive an additional payment based on the number of white balls that are drawn from the jar. The method of payment is described in detail. After completing the comprehension quiz, subjects face the exact same distributions in Task 2 as they did in Task 1. The order of the distributions is again randomized and subjects have 2 minutes per distribution. Upon completion, instructions for Task 2 and used working paper are collected and Task 3 instructions and new working paper are distributed.

Next, Task 3 instructions are read aloud. Subjects are informed that they will not receive an extra payment for white balls, but that they will answer a set of trivia questions and then face a screen that is similar to the screens faced in Tasks 1 and 2. However, now they can add a white ball to their jar for each correct trivia answer. Subjects are also told that they will not be told how many they answer correctly and that they will make additional money for each correct trivia answer. Subjects subsequently practice with the slightly modified interface. In Task 3, subjects face either 1 trivia question or 3 trivia questions. When they face 1 trivia question they are told that if they answer 0 (1) correctly they will face a jar with 1 (2) white ball(s), 1 black ball and 1 draw. When they face 3 trivia questions they are told that if they answer 0 (1,2,3) correctly there would a jar with 1 (2,3,4) white ball(s), 3 black and 3 draws.

Upon completion of Task 3, the experimenter collects Task 3 instructions and used working paper and distributes Task 4 instructions and clean working paper. Subjects are informed that Task 4 is identical to Task 3 except now they can earn an additional payment when more white balls are drawn. Subjects practice with the interface, complete a comprehension quiz and then complete Task 4. Upon completing Task 4, the computer randomly chooses one of the four tasks and calculates the subject's payment based on responses for that task. The subjects are informed which task determines their payment and their earnings. Subjects are individually escorted out of the room to insure that payments are confidential and the experiment concludes.

Notice that the order of tasks reflects an increasing level of cognitive demand. It is also possible to reverse the order to test whether treatment effects are robust to possible order effects. When we reverse the order, we again appeal to the similarities of each of the previous tasks when describing a new task. The aim is to keep the instructions as comparable and simple as possible.

5 Results

In this section, we present results from the experiment outlined in the previous section. We begin with summary statistics that directly test the definitions proposed in Section 3. Second, we examine within-subject correlations in the payment, performance and combined treatment effects. In presenting our findings, to highlight our three main sets of results, we will label them as such (Results 1–3).

5.1 Summary Statistics and Treatment Effects

We conducted 11 experimental sessions in February and April of 2013 in the MISSEL laboratory at Washington University in Saint Louis. 98 subjects participated (52 females and 46 males) including both undergraduate and graduate students.[21] Table 3 lists the number of observations for each experimental effect. To be counted, subjects needed to face the same distribution for both the treatment level and the neutral-exogenous level. It is important to understand that not all subjects faced all distributions under all treatments. This occurs because the performance treatment effectively endogenizes the distribution. For example, suppose a subject never answers a trivia question correctly when he starts with Distribution 1 under the performance treatment. Then, he will never face Distribution 2 under the performance treatment. Similarly, a subject might face the same distribution under the same treatment more than once, in which case answers were averaged. Given this design, there are 311 subject distribution pairs under the neutral-endogenous level and 322 under the induced preference-endogenous level. For comparison across treatments, a given subject needed to face a given distribution under all treatment levels to be compared. Given endogenous distributions, there are 259 observations where subjects faced the same distribution for both the induced preference-exogenous and the neutral-endogenous levels. There are 160 observations where subjects faced the same distribution for the induced preference-exogenous, neutral-endogenous and induced preference-endogenous treatment levels.

Table 3: Summary Statistics: Sample Size

	Total	Male	Female
Subjects	98	46	52
Observations			
Total	1143	537	606
Payment Effect	508	240	268
Performance Effect	313	146	167
Combined Effect	322	151	171
Payment and Performance Effects	259	125	134
Combined, Payment and Performance Effects	160	79	90

[21] We ran 8 in the order described at length above and 3 in the complete reverse order. This was done to insure that results were robust to order effects, which we show in Appendix B.

5.1.1 Test of Definitions: Optimists and Performance Over-estimators

To directly test definitions 1-4 we classify subjects using the categorical variable, $T_{i,d,\tau}$, which scores whether a subject's beliefs first-order stochastically dominate the objective distribution in treatment level τ versus the neutral-exogenous in each distribution. In Table 3, subjects are categorized in each distribution based on their classification for $T_{i,d,\tau=\text{Payment}}$ and $T_{i,d,\tau=\text{Performance}}$. Table 3 highlights three key points. First, for every distribution, except distribution 1, optimism is more prevalent than pessimism, however the results suggest that under our strict definitions optimistic beliefs are still far less common than neutral beliefs.

Table 4: Summary Statistics: Within-Subject Correlations

			Distributions					
$T_{i,d,Payment}$		$T_{i,d,Performance}$	(1)	(2)	(3)	(4)	(5)	(6)
Optimistic	&	Over-estimate	0	.08	.06	.04	.08	0
Optimistic	&	Neutral	0	0	0	.15	.02	.27
Optimistic	&	Under-estimate	0	0	0	0	0	0
Neutral	&	Over-estimate	.57	.47	.31	.19	.02	.02
Neutral	&	Neutral	.13	.25	.63	.54	.73	.66
Neutral	&	Under-estimate	.21	.16	0	.04	.06	.18
Pessimistic	&	Over-estimate	.08	.03	0	0	.03	0
Pessimistic	&	Neutral	.01	0	0	.04	.04	0
Pessimistic	&	Under-estimate	.01	.02	0	0	.01	.06
Number of Reports (by Distribution):			92	103	16	26	97	50

For each distribution, subjects are categorized according to their average (categorical) responses to the payment treatment ($T_{i,d,\tau=\text{Payment}}$) and the performance treatment ($T_{i,d,\text{Performance}}$).

Second, the results on performance overestimation appear to be consistent with performance uncertainty.[22] Subjects are relatively more likely to over-estimate their performance when they perform particularly badly and to under-estimate their performance when they perform particularly well. To see this, consider the classifications for distributions 3-6. Performance over-estimation declines monotonically from distributions 3-6 (37%, 23%, 21%, 13% and 2%, respectively) and performance under-estimation increases monotonically from distributions 3-6 (0%, 4%, 7% and 24%, respectively). Third, even with these very strict categorical variables, there appears to be some interesting correlation patterns. Perhaps most striking, no subject simultaneously under-estimates performance and exhibits optimism. Fur-

[22]Benoît and Dubra (2011) and Benoît et al. (2013) discuss the possibility of over- or under-estimation of performance as reflecting uncertainty rather than as a mistake.

ther, a small proportion of subjects simultaneously exhibit pessimism and over-estimation.

Next, we examine definitions 5-7 and exploit the additional variation in our data that cannot be captured by the categorical variable. We consider the continuous variable $\Delta_{i,d,\tau}$, which captures treatment-induced shifts in beliefs. Recall, $\Delta_{i,d,\tau}$ measures the average difference between the neutral-exogenous subjective distribution and the subjective distribution under the treatment effect τ in distribution d for subject i. This variable gives a measure of shifts in beliefs that result from either the introduction of induced preference, performance uncertainty or both. First, we generate histograms of treatment effects, which are presented in Figure 2.[23] These histograms highlight a similar pattern found in our examination of the categorial variables. In particular, subjects tend to be correct and over-estimation tends to occur more often than under-estimation.

To determine whether treatment effects are significant, we use regression analysis to examine how individual level treatment effects relate to observed subject characteristics. In particular, we model the outcome variable $\Delta_{i,d,\tau}$ as follows:

$$\begin{aligned}
\Delta_{i,d,\tau} &= \beta_{\text{Payment}} \times \mathbf{1}[\text{Payment Effect}] \\
&+ \beta_{\text{Performance}} \times \mathbf{1}[\text{Performance Effect}] \\
&+ \beta_{\text{Combined}} \times \mathbf{1}[\text{Combined Effect}] \\
&+ \gamma_d + \delta_1 X_i + \varepsilon_{i,d,\tau},
\end{aligned}$$

where $\mathbf{1}[\cdot]$ are treatment indicator variables. $\beta_{Payment}$, $\beta_{Performance}$ and $\beta_{Combined}$ are treatment effects, γ_d is a distribution fixed effect and X_i contains individual-level characteristics, including session fixed effects and coefficients δ_1. Errors ($\epsilon_{i,d,\tau}$) are clustered at the individual level.

OLS regression estimates are presented in Table 5, where positive estimates of treatment effects indicate that, on average, subjects exhibit optimistic shifts or shift towards performance over-estimation. Specifically, column (2) shows that subjects report cumulative distributions that are shifted, on average, to the right by 12 percentage points when they have preferences over the outcomes relative to when they do not have such preferences. Further, our estimate of β_{Combined}, the response to the combined effects, is less than the sum of the payment and performance effects. This sub-additivity suggests a correlation in responses to the payment and performance effects, a possibility that will be explored in the following section. Our estimates are robust when we control for the total number of trivia questions that a subject answered correctly, thereby circumventing possible selection bias in

[23] In Appendix B, similar histograms are generated for the the categorical measure, where we average across distributions for each agent. See Figure B1.

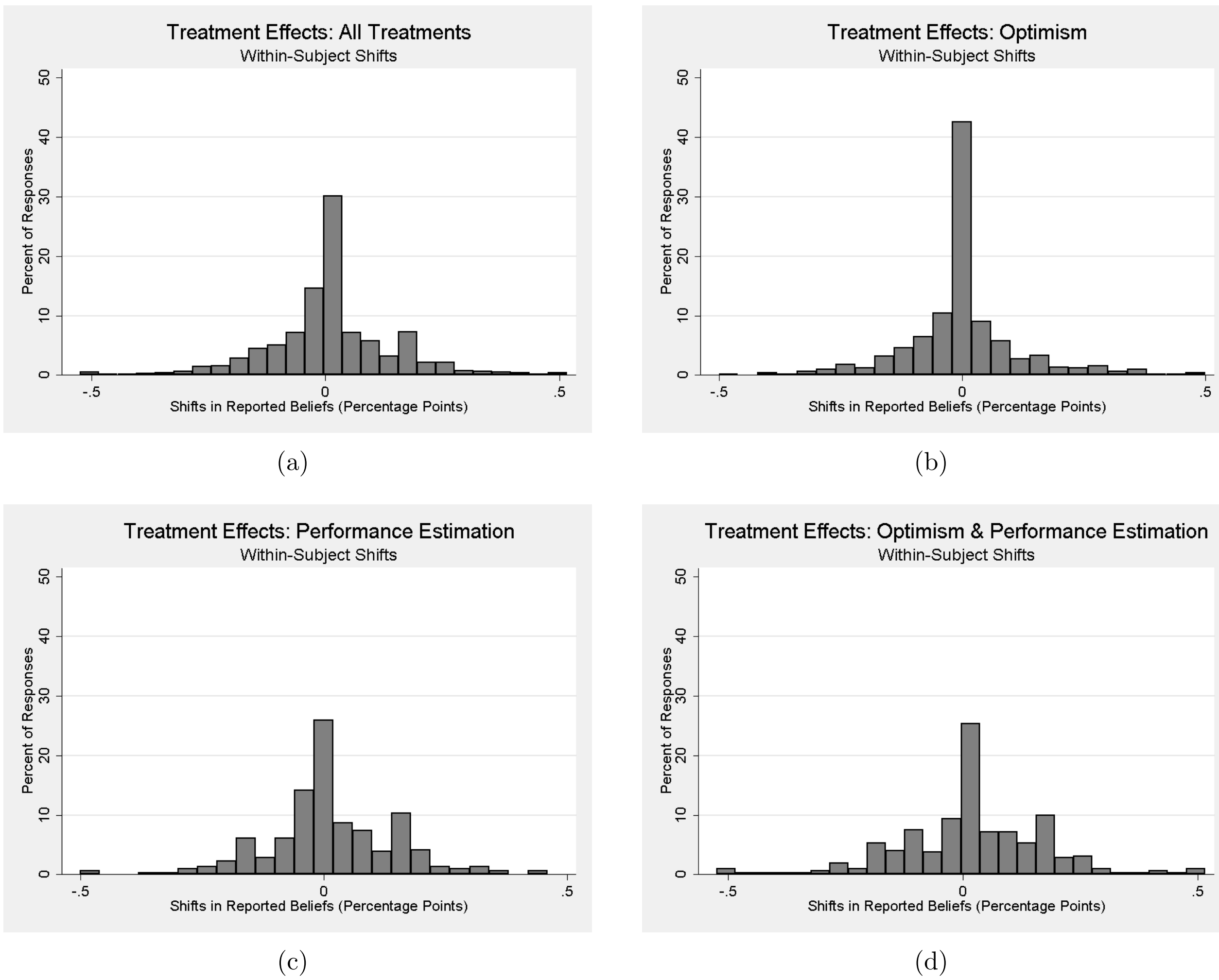


FIGURE 2: DISTRIBUTION OF TREATMENT EFFECTS: Treatment effects for individual i and distribution d are the difference between elicited beliefs under the baseline-exogenous and treatment level $\tau \subseteq \{\text{Payment, Performance}\}$ and denoted $\Delta_{i,d,\tau}$. Histograms are provided for all treatment effects (Panel 2(a)), payment treatment effects (Panel 2(b)), performance treatment effects (Panel 2(c)) and combined treatment effects (Panel 2(d)).

Table 5: Subject Characteristics and Treatment Effects: Definitions 5-7

	(1)	(2)	Male Only (3)	Female Only (4)	Gender Interactions (5)
Payment Effect	0.08***	0.12***	0.1***	0.14***	0.1***
Performance Effect	0.1***	0.13***	0.11***	0.15***	0.11***
Combined Effect	0.1***	0.14***	0.12***	0.15***	0.12***
Female	.	-0.001	.	.	.
Total Trivia Correct	.	-0.004	-0.003	-0.006	-0.003
Pay. Effect × Fem.	.	.	.	.	0.04
Perf. Effect × Fem.	.	.	.	.	0.05
Comb. Effect × Fem.	.	.	.	.	0.03
Observations	1143	1143	537	606	1143

Pooled treatment effects are regressed onto observed subject characteristics. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the neutral-exogenous and treatment level $\tau \subseteq$ {Payment, Performance} and denoted $\Delta_{i,d,\tau}$. Session and distribution fixed effects are included in all regressions and Column (5) interacts gender with the experimental treatment, where the omitted category is the female indicator variable. Errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

the distributions that subjects face in the performance treatment.[24]

5.2 Relating Optimism and Performance Over-Estimation

Recall that results in Table 5 suggest that the combined effect is sub-additive. This finding suggests possible positive correlation between performance over-estimation and optimism. Our goal in this section is to more closely examine to what degree the different treatment effects are correlated within individuals.

First, we ask whether subjects who tend to engage in optimistic behavior also tend towards over-estimating their performance. To examine this relationship, we regress performance shifts ($\Delta_{i,d,Performance}$) on payment shifts ($\Delta_{i,d,Payment}$). Without including any controls, the raw correlation is approximately 0.48. Next, we run the following regression:

$$\Delta_{i,d,Performance} = \phi_1 * \Delta_{i,d,Payment} + \gamma_d + \delta_2 X_i + \eta_{i,d,\tau},$$

where γ_d are distribution fixed effects, X_i is a vector of subject level observed characteristics,

[24]These results are also robust if we perform regressions using the categorical variable $T_{i,d,\tau}$ rather than the continuous variable. Results are in Appendix B, Table B5.

including session fixed effects and coefficient vector δ_2. Errors are clustered at the subject level. Note that $\phi_1 > 0$ indicates that a subject that shifts towards performance over-estimation in distribution d is also likely to shift towards optimism in distribution d. In fact, the estimates in Table 6 show exactly this (where $\Delta_{i,d,Payment}$ is recorded in the table as Δ Payment Effect).

The positive correlation suggests that optimism and performance over-estimation are reinforcing. An individual who perceives his own performance to be better than it actually is will also tend to believe that the likelihood of success at some endeavor, given this misperceived level of performance, is greater yet. This reinforcing behavior compounds decision-making errors. This result also highlights the strength of our experimental design, which allows us to explore the relationship between two effects that are often confounded in the literature. In other words, in settings where subjects face uncertainty about their own performance and also favor some outcomes over others, attributing elicited beliefs to either optimism or over-estimation of performance ignores that these two beliefs are different and overlooks the possibility that they might be correlated.

Table 6: Subject Characteristics and Correlations in Treatment Responses

	(1)	(2) Male Only	(3) Female Only	(4) Gender Interactions
Δ Payment Effect	0.64***	0.51***	0.68***	0.51***
Female	-0.006	.	.	.
Total Trivia Correct	-0.003	-0.009**	0.003	-0.009**
Δ Pay. Effect $\times$ Fem.	.	.	.	0.17
Observations	259	125	134	259

Performance treatment effects are regressed onto payment treatment effects and observed characteristics. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the baseline-exogenous and treatment level $\tau \subseteq$ {Payment, Performance} and denoted $\Delta_{i,d,\tau}$. Session fixed effects are included and errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Result 1. *Performance over-estimation and optimism are positively correlated at the individual level.*

To further explore the relationship between performance over-estimation and optimism, we exploit our experimental design to study how each, measured in isolation, correlates to behavior in settings where both performance and a preferred outcome are present. This amounts to decomposing the beliefs in the combined effect. To accomplish this, we estimate the following model:

$$\begin{aligned}\Delta_{i,d,\tau=Combined} &= \alpha_1 * \Delta_{i,d,\tau=Payment} \\ &+ \alpha_2 * \Delta_{i,d,\tau=Performance} \\ &+ \gamma_d + \delta_3 X_i + \sigma_{i,d,\tau}.\end{aligned}$$

Results are presented in Table 7.[25] Columns (1) and (2) present estimates where the combined effects are regressed onto the payment and performance effects individually. We find a positive correlation since the combined effect is composed of the payment and performance effects. In Column (3), we include both the payment and performance treatment effects and find that the estimated coefficients are much lower (0.65 versus 0.80 for payment treatment effects and 0.40 versus 0.54 for performance treatment effects).[26] This difference reflects how positive correlation between the two leads to classic omitted variables bias. This highlights how explaining behavior as either optimism or over-estimation of performance in settings where both types of beliefs can occur is misleading since optimism correlates with performance over-estimation at the individual level.

Result 2. *Optimism and performance over-estimation explain over-estimation in the combined setting. Given Result 1, explaining over-estimation in the combined setting only as performance over-estimation induces omitted variables bias.*

5.3 Gender Differences in Optimism and Performance Estimation

Previous work studying performance over-estimation and biased beliefs has considered the possibility of gender differences. Our final set of results explores gender differences in optimism and performance over-estimation and finds that gender differences, though nuanced, do exist. We explore average gender differences in Columns (3)-(5) of Table 5. Columns (3) and (4) contain estimates of coefficients for regressions run separately by gender and show that females appear to have a slightly stronger treatment effect than males. In Column (5), however, we fully interact a gender dummy with all treatment, distribution and session

[25] Here, we note that a selection problem arises since we can only use observations where subjects performed equally well on the trivia task under the modified-endogenous and baseline-endogenous levels. In order to get around this problem, we average $\Delta_{i,d,\tau=Payment}$, $\Delta_{i,d,\tau=Performance}$ and $\Delta_{i,d,\tau=Combined}$ across distributions 1 and 2 and distributions 3 through 6. This results in, at most, two observations per subject. To be clear, denote these re-averaged quantities as $\bar{\Delta}_{i,d,\tau}$.

[26] The coefficients on the Payment Effect and the Performance Effect shown in Columns (3) are not significantly different ($p = .16$).

Table 7: Decomposing the Combined Payment & Performance Treatments

	(1)	(2)	(3)	(4) Male Only	(5) Female Only	(6) Gender Interact.
$\bar{\Delta}$ Payment Effect	0.8***	.	0.65***	0.36*	0.91***	0.36*
$\bar{\Delta}$ Performance Effect	.	0.54***	0.4***	0.48***	0.4***	0.48***
Female	-0.004	-0.006	-0.004	.	.	.
Total Trivia Correct	-0.004	-0.005	-0.002	0.0004	0.005	0.0004
$\bar{\Delta}$ Pay. Effect × Fem.	.	.	.	.	.	0.55**
$\bar{\Delta}$ Perf. Effect × Fem.	.	.	.	.	.	-0.07
Observations	169	169	169	79	90	169

Combined treatment effects are regressed onto payment treatment effects and performance treatment effects. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the baseline-exogenous and treatment level $\tau \subseteq \{\text{Payment, Performance}\}$ and denoted $\Delta_{i,d,\tau}$. Session fixed effects are included and errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

variables. We find no significant difference across genders in treatment effects.[27]

The lack of gender differences in treatment effects is misleading. In Table 7 we present estimates of the decomposition by gender. Results in Columns (4) and (5) show that, although total combined treatment effects do not differ by gender, their source does. In particular, we find that, for males, over-estimation induced by the combined treatment effect are mostly explained by performance over-estimation as opposed to optimism. The opposite is true for females: optimism is a more important determinant of response to the combined effects.

Result 3. *On average, men do not over-estimate their performance more than women and do not differ in their average level of optimism. However, male over-estimation in the combined setting is explained by over-estimation of performance. Female over-estimation in the combined setting is explained by both optimism and performance over-estimation, but is driven significantly more by their optimism.*

Though our findings on beliefs about performance estimation are generally consistent with previous research, our findings on gender differences in the source of mis-estimation are novel. A key difference that helps to contextualize this finding is that performance over-estimation involves beliefs over outcomes that the individual can influence through his

[27]We experimented with different specifications and continue to find no evidence of gender differences in treatment effects. For example, if we only interact treatment dummies with the female indicator and do not interact session and distribution dummies, results remain qualitatively unchanged.

actions, whereas optimism involves beliefs about payoff-favorable outcomes over which the individual exerts no control. Psychology research has studied a related distinction in what is called the *locus of control*. This concept distinguishes between individuals with an *internal locus*, who view life events as contingent on their own actions, and those with an *external locus*, who attribute events to characteristics over which they have no control.[28] Our findings are consistent with research suggesting that males have an internal locus of control, whereas females tend towards an external locus of control (Sherman et al., 1997).

6 Conclusion

This paper presents findings from a novel experiment in which subjects face two kinds of uncertainty: performance uncertainty and objective uncertainty where they are induced to have preferences over outcomes. This design enables us to study optimism and performance over-estimation—both as separate phenomena and as phenomena that potentially interact to drive decisions. Moreover, the subject-level control provides a natural belief-counterfactual where the subject's beliefs are elicited absent performance uncertainty and preferences over outcomes. This control is crucial to our identification strategy. Finally, tasks are uniform across treatment combinations. Thus, subjects' beliefs are reported in the same units so that comparing, correlating and decomposing the variance of treatment effects is straightforward.

First, we find that subjects who respond optimistically also tend to over-estimate their performance. Second, we show that this correlation induces omitted variables bias if one attempts to attribute probabilistic assessments of success in real-world settings solely to beliefs about one's own performance. Thus, to understand subjective assessments about success when performance can increase the likelihood of success, it is not sufficient to focus solely on beliefs about performance. Third, we show that males and females, on average, respond similarly to experimental treatments. Both exhibit optimism and performance over-estimation. Further, when beset by performance uncertainty and induced to have preferences over outcomes, men and women over-estimate the likelihood of success in similar magnitudes. However, the source of their mis-estimation is different. Male mis-estimation is explained by performance over-estimation, whereas female over-estimation is better explained by optimism.

Understanding what fosters successful entrepreneurship—and what leads to failure—has large implications for economic growth, but remains elusive. Communities strive to replicate

[28]Koellinger et al. (2007) discuss a positive relationship between optimism and internal versus external locus of control to explain cross-country variations in entrepreneurial activity.

the entrepreneurial hotbeds found in Silicon Valley and Route 128 in Massachusetts. In order to succeed, the community must understand how those places came to be. This requires pinpointing what sort of useful information is shared among the local entrepreneurs. Are entrepreneurs learning new technical skills that help them perfect a new product (Gompers et al., 2005), which would bring beliefs about expected performance more in line with reality? Alternatively, does exposure to other entrepreneurs simply resolve some of the ambiguity about the challenges of a startup (Minniti, 2005), which may attenuate optimistic beliefs? We also know that projects funded by venture capital firms with activist investors, who have previous entrepreneurial experience, are more likely to succeed (Bottazzi et al., 2008). Here, again, it is unclear what is contained in the information they transmit to the entrepreneurs in whom they invest.

Focusing on individual beliefs, our study sheds light on why fostering successful entrepreneurship remains an elusive objective. We find that individuals exhibit patterns in their beliefs that could explain entrepreneurial high rates of entry and subsequent failure. Our findings also show that men and women may be driven by different types of beliefs, which suggests that they would benefit from different types of information and advice. Men, according to our results, would find performance feedback and training more helpful, while women would benefit more from advice on the general challenges of entrepreneurship that are beyond their control.

References

Alicke, M. D., "Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives.," *Journal of Personality and Social Psychology* 49 (1985), 1621.

Benoît, J.-P. and J. Dubra, "Apparent Overconfidence," *Econometrica* 79 (2011), 1591–1625.

Benoît, J.-P., J. Dubra and D. Moore, "Does the Better-Than-Average Effect Show That People Are Overconfident?: Two Experiments," (2013).

Blavatskyy, P., "Betting on Own Knowledge: Experimental Test of Overconfidence," *Journal of Risk and Uncertainty* 38 (2009), 39–49.

Bottazzi, L., M. Da Rin and T. Hellmann, "Who Are the Active Investors?: Evidence from Venture Capital," *Journal of Financial Economics* 89 (2008), 488–512.

Brier, G. W., "Verification of Forecasts Expressed in Terms of Probability," *Monthly weather review* 78 (1950), 1–3.

Brunnermeier, M. and J. Parker, "Optimal Expectations," *American Economic Review* 95 (2005), 1092–1118.

Camerer, C. and D. Lovallo, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review* 89 (1999), 306–318.

Card, D., A. Mas, E. Moretti and E. Saez, "Inequality at work: The effect of peer salaries on job satisfaction," *American Economic Review* 102 (2012), 2981.

Clark, J. and L. Friesen, "Overconfidence in Forecasts of Own Performance: An Experimental Study," *The Economic Journal* 119 (2009), 229–251.

Clark, R. L., J. A. Maki and M. S. Morrill, "Can Simple Informational Nudges Increase Employee Participation in a 401(K) Plan?," NBER working paper (2013).

Elston, J., G. Harrison and E. Rutström, "Experimental Economics, Entrepreneurs and the Entry Decision," *University of Central Florida working paper* (2006), 06–06.

Fang, H. and G. Moscarini, "Morale Hazard," *Journal of Monetary Economics* 52 (2005), 749–777.

Fréchette, G. R., A. Schotter and I. Trevino, "Personality and Choice in Risky and Ambiguous Environments: An Experimental Study," Mimeo, Dept. of Economics, New York University. (2011).

Gompers, P., J. Lerner and D. Scharfstein, "Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1999," *The Journal of Finance* 60 (2005), 577–614.

Grossman, Z. and D. Owens, "An unlucky feeling: Overconfidence and noisy feedback," *Journal of Economic Behavior & Organization* (2012).

Grosswirth, M., A. Salny and A. Stillson, *Match Wits with Mensa: The Complete Quiz Book* (Da Capo Press, 1999).

Hamilton, B., "Does Entrepreneurship Pay? An Empirical Analysis of the Returns of Self-Employment," *Journal of Political Economy* (2000), 604–631.

Hoelzl, E. and A. Rustichini, "Overconfident: Do You Put Your Money On It?," *The Economic Journal* 115 (2005), 305–318.

Ito, T., "Foreign Exchange Rate Expectations: Micro Survey Data.," *American Economic Review* 80 (1990), 434–49.

Jensen, R., "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics* 125 (2010), 515–548.

Kaldor, N., "The Relation of Economic Growth and Cyclical Fluctuations," *The Economic Journal* 64 (1954), pp. 53–71.

Kirchler, E. and B. Maciejovsky, "Simultaneous Over-and Underconfidence: Evidence from Experimental Asset Markets," *Journal of Risk and Uncertainty* 25 (2002), 65–85.

Klayman, J., J. Soll, C. González-Vallejo and S. Barlas, "Overconfidence: It Depends on How, What, and Whom You Ask," *Organizational Behavior and Human Decision Processes* 79 (1999), 216–247.

Koellinger, P., M. Minniti and C. Schade, ""I Think I Can, I Think I Can": Overconfidence and Entrepreneurial Behavior," *Journal of Economic Psychology* 28 (2007), 502–527.

Köszegi, B., "Ego Utility, Overconfidence, and Task Choice," *Journal of the European Economic Association* 4 (2006), 673–707.

Lerner, J. and U. Malmendier, "With a Little Help from my (Random) Friends: Success and Failure in Post-Business School Entrepreneurship," NBER working paper (2011).

Lichtenstein, S. and B. Fischhoff, "Do Those Who Know More Also Know More About How Much They Know?," *Organizational Behavior and Human Performance* 20 (1977), 159–183.

Malmendier, U. and G. Tate, "CEO Overconfidence and Corporate Investment," *The Journal of Finance* 60 (2005), 2661–2700.

Manski, C. F., "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?," in *Studies of supply and demand in higher education* (University of Chicago Press, 1993), 43–60.

Mayraz, G., "Wishful Thinking," Mimeo, University of Melbourne (2011).

McKelvey, R. D. and T. Page, "Public and Private information: An Experimental Study of Information Pooling," *Econometrica: Journal of the Econometric Society* (1990), 1321–1339.

Minniti, M., "Entrepreneurship and network externalities," *Journal of Economic Behavior & Organization* 57 (2005), 1–27.

Moore, D. and P. Healy, "The Trouble with Overconfidence.," *Psychological review* 115 (2008), 502.

Moskowitz, T. and A. Vissing-Jorgensen, "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," NBER working paper (2002).

Murphy, A. H. and R. L. Winkler, "Scoring Rules in Probability Assessment and Evaluation," *Acta Psychologica* 34 (1970), 273–286.

Nanda, R. and J. B. Sørensen, "Workplace peers and entrepreneurship," *Management Science* 56 (2010), 1116–1126.

Nguyen, T., "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," *MIT Working Paper* (2010).

Niederle, M. and L. Vesterlund, "Do Women Shy Away from Competition? Do Men Compete too much?," *Quarterly Journal of Economics* 122 (2007), 1067–1101.

Owens, D., Z. Grossman and R. Fackler, "The Control Premium: A Preference for Payoff Autonomy," (2012).

Quiggin, J., "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization* 3 (1982), 323–343.

Rabin, M. and R. Thaler, "Anomalies: Risk Aversion," *Journal of Economic Perspectives* (2001), 219–232.

Russo, J. and P. Schoemaker, "Managing Overconfidence," *Sloan Management Review* 33 (1992), 7–17.

Santos-Pinto, L. and J. Sobel, "A Model of Positive Self-Image in Subjective Assessments," *American Economic Review* (2005), 1386–1402.

Savage, L., *The Foundations of Statistics* (Wiley, 1954).

Selten, R., A. Sadrieh and K. Abbink, "Money Does not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse," *Theory and Decision* 46 (1999), 213–252.

Sherman, A. C., G. E. Higgs and R. L. Williams, "Gender Differences in the Locus of Control Construct," *Psychology and Health* 12 (1997), 239–248.

Sonnemans, J. and T. T. Offerman, "Is the Quadratic Scoring Rule Behaviorally Incentive Compatible?," *European Economic Review* (2001).

Svenson, O., "Are We All Less Risky and More Skillful than Our Fellow Drivers?," *Acta Psychologica* 47 (1981), 143–148.

Urbig, D., J. Stauf and U. Weitzel, "What Is Your Level of Overconfidence? A Strictly Incentive Compatible Measurement of Absolute and Relative Overconfidence," *Discussion paper series/Tjalling C. Koopmans Research Institute* 9 (2009), 1–40.

Van den Steen, E., "Rational Overoptimism (and Other Biases)," *American Economic Review* 94 (2004), 1141–1151.

Weinstein, N. D., "Unrealistic Optimism about Future Life Events," *Journal of Personality and Social Psychology* 39 (1980), 806–820.

Wiswall, M. and B. Zafar, "Determinants of College Major Choice: Identification Using an Information Experiment," *FRB of New York Staff Report* (2011).

Appendix A Robustness of Results to Order Effects

Because we rely on a within-subject variation our results may be confounded by order effects. We ran our experiment in two sensible orders. The first order follows the description in Section 4.6. This is the most natural order as it allows us to appeal to the similarity across each treatment adding a degree of complexity as the experiment progresses. We also ran sessions in the reverse order. This again allows us to appeal to the similarities of each task, but instead begin with the more complex task and reduce complications as we move through the experiment. In this section we test whether our results are confounded with the order of the tasks. To do this, we walk through the main tables presented in Section 5 and account for the order in which the tasks were completed.

TABLE A1: ORDER EFFECTS ROBUSTNESS: SUBJECT CHARACTERISTICS AND TREATMENT EFFECTS

	(1)	(2)
Payment Effect	0.07^{**}	0.1^{***}
Performance Effect	0.09^{***}	0.07^{**}
Combined Effect	0.09^{***}	0.06^{*}
Pay. Effect × Order	.	-0.005
Perf. Effect × Order	.	0.06^{***}
Comb.Effect × Order	.	0.08^{***}
Female	-0.003	-0.003
Total Trivia Correct	-0.005^{*}	-0.005^{*}
Order Dummy Variable	0.04^{**}	.
Observations	1143	1143

This table shows estimates testing whether results presented in Table 5 are robust to order effects. Pooled treatment effects are regressed onto observed subject characteristics. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the neutral-exogenous and treatment level $\tau \subseteq$ {Payment, Performance} and denoted $\Delta_{i,d,\tau}$. Column (1) contains a dummy variable that takes a value of 1 if the treatments were completed in the order described throughout the paper and a value of 0 when the order is reversed. Column (2) fully interacts the Order dummy with treatment dummies. Distribution fixed effects included. Errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

First, we look at the average treatment effects reported in Table 5. Column (1) estimates the same model as Column (2) of Table 5, but replaces session fixed effects with a dummy variable that takes a value of 1 (Order=1) when the tasks are completed in the order described in Section 4.6 and 0 (Order=0) when the tasks were completed in the reverse order. Average treatment effects are still positive, although slightly smaller in magnitude. Column (2) interacts the Order dummy with the treatment dummies to test whether average treatment responses differed by order. The coefficients in Column (2) suggest subjects who faced the performance treatment during the final two tasks of the session were more likely to over-estimate. Remember, over-estimation is more prevalent at lower levels of performance and this order effect may be explained by fatigue. If subjects who faced the performance treatment during the final half of the session performed worse than subjects who faced the treatment in the first half, then we would except to see more performance over-estimation. In fact, this seems to be the case: the average number of correct trivia questions per subject is 9.14 (sd=2.41) in Order=1 and 9.43 (sd=2.90) in Order=0.

Table A2 presents a robustness check on the correlations between performance over-estimation and optimism. Column (1) replicates the first column of Table 6, but replaces session fixed effect with the Order dummy variable. The coefficients are nearly identical. Column (2) goes a step further and presents the estimates when the Order dummy is interacted with the Payment Effect. The interaction term is positive but not significant.

TABLE A2: ORDER EFFECT ROBUSTNESS: SUBJECT CHARACTERISTICS AND CORRELATIONS IN TREATMENT RESPONSES

	(1)	(2)
Δ Payment Effect	0.62***	0.38*
Δ Pay. Effect $\times$ Order	.	0.26
Female	-0.007	-0.006
Total Trivia Correct	-0.003	-0.003
Order Dummy Variable	0.05***	0.05***
Observations	259	259

This table shows estimates testing whether results presented in Table 6 are robust to order effects. Performance treatment effects are regressed onto payment treatment effects and observed characteristics and in Column (2) the payment treatment effect is interacted with the Order dummy to check for order effects. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the neutral-exogenous and treatment level $\tau \subseteq \{\text{Payment, Performance}\}$ and denoted $\Delta_{i,d,\tau}$. Distribution fixed effects are included and errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Tables A3 and A4 report robustness effects for the decomposition by male and female, respectively. The first columns of each table replicate Column (3) and (4) of Table 7 for reference. Column (2) replaces session fixed effects with the Order dummy variable and Columns (3) interacts the Order dummy variable with the two treatment effects. The story remains the same: male mis-estimation in the combined setting is driven by performance over-estimation whereas female mis-estimation is driven by optimism.

Table A3: Order Effects Robustness: Decomposing the Combined Payment & Performance Treatments (Males)

	(1)	(2)	(3)
$\bar{\Delta}$ Payment Effect	0.36*	0.18	0.26
$\bar{\Delta}$ Performance Effect	0.48***	0.44***	0.36**
$\bar{\Delta}$ Pay. Effect × Order	.	.	-0.11
$\bar{\Delta}$ Perf. Effect × Order	.	.	0.15
Total Trivia Correct	0.0004	-0.007	-0.006
Order Dummy Variable	.	0.01	0.01
Observations	79	79	79

This table shows estimates testing whether results presented in the first three columns of Table 7 are robust to order effects. Combined treatment effects are regressed onto payment treatment effects and performance treatment effects. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the baseline-exogenous and treatment level $\tau \subseteq \{\text{Payment, Performance}\}$ and denoted $\Delta_{i,d,\tau}$. Column (1) replicates results from Table 7. Column (2) includes an Order dummy variable and Column (3) interacts order with the two treatment effects. Distribution fixed are included and errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table A4: Order Effects Robustness: Decomposing the Combined Payment & Performance Treatments (Females)

	(1)	(2)	(3)
$\bar{\Delta}$ Payment Effect	0.91^{***}	0.86^{***}	0.97^{***}
$\bar{\Delta}$ Performance Effect	0.4^{***}	0.41^{***}	0.58^{**}
$\bar{\Delta}$ Pay. Effect $\times$ Order	.	.	-0.2
$\bar{\Delta}$ Perf. Effect $\times$ Order	.	.	-0.2
Total Trivia Correct	0.005	0.003	0.004
Order Dummy Variable	.	0.1^{***}	0.1^{***}
Observations	90	90	90

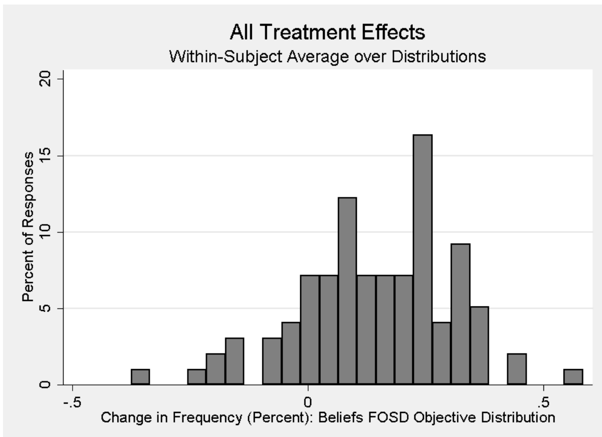
This table shows estimates testing whether results presented in the second three columns of Table 7 are robust to order effects. Combined treatment effects are regressed onto payment treatment effects and performance treatment effects. Treatment effects for individual i and distribution d are the difference between elicited beliefs under the baseline-exogenous and treatment level $\tau \subseteq \{\text{Payment, Performance}\}$ and denoted $\Delta_{i,d,\tau}$. Column (1) replicates results from Table 7. Column (2) includes an Order dummy variable and Column (3) interacts order with the two treatment effects. Distribution fixed are included and errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Appendix B Additional Tables and Figures

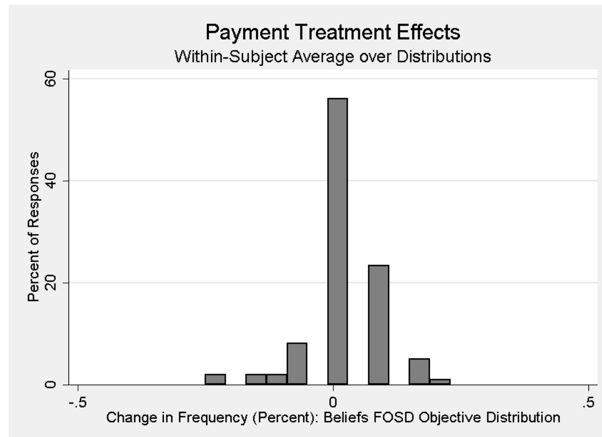
Table B5: Subject Characteristics and Treatment Effects

	(1)	(2)	Male Only (3)	Female Only (4)	Gender Interactions (5)
Payment Effect	0.22***	0.21***	0.19	0.22**	0.16
Performance Effect	0.4***	0.39***	0.41***	0.35***	0.38***
Combined Effect	0.5***	0.49***	0.48***	0.49***	0.46***
Female	.	-0.04	.	.	.
Total Trivia Correct	.	0.005	-0.002	0.003	0.0007
Pay. Effect × Fem.	.	.	.	.	0.08
Perf. Effect × Fem.	.	.	.	.	-0.007
Comb. Effect × Fem.	.	.	.	.	0.05
Observations	1181	1181	567	614	1181

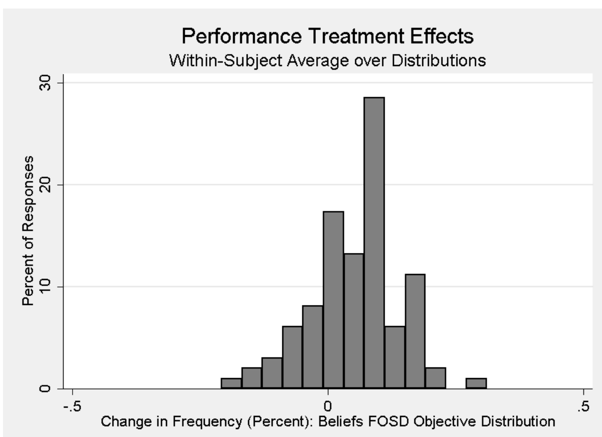
Pooled treatment effects are regressed onto observed subject characteristics. Treatment effects (denoted $T_{i,d,\tau}$) for individual i and distribution d are categorical, taking a value of 1 if elicited beliefs FOSD the objective under τ and not under Baseline-Exogenous, -1 if beliefs are first-order stochastically dominated under τ, but not under Baseline-Exogenous and a value of 0 if neither occurs. Session and distribution fixed effects are included in all regressions and Column (5) interacts gender with the experimental treatment, where the omitted category is the female indicator variable. Errors are clustered at the individual level and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.



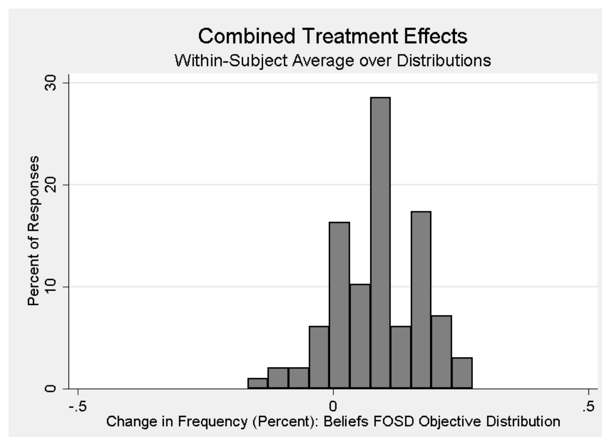
(a)



(b)



(c)



(d)

FIGURE B1: DISTRIBUTION OF TREATMENT EFFECTS: Treatment effects (denoted $T_{i,d,\tau}$) for individual i and distribution d are categorial, taking the value of 1 if elicited beliefs FOSD the objective under $\tau \subseteq \{Payment, Performance\}$ and not under the baseline-exogenous and equal to -1 if beliefs are first-order stochastically dominated under τ but not under the baseline-exogenous and a value of 0 if neither occurs. These treatment effects are then averaged across distributions. Histograms are provided for all treatment effects (Panel 1(a)), payment treatment effects (Panel 1(b)), performance treatment effects (Panel 1(c)) and combined treatment effects (Panel 1(d)).