We Should *Totally* Open a Restaurant: How Optimism and Overconfidence Affect Beliefs*

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

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

October 6, 2015

Abstract: Overestimating the probability of high-payoff outcomes is a widely-documented phenomenon that can affect decision-making in numerous domains, including finance, management, and entrepreneurship. In this paper, we design an experiment to distinguish between two types of biases in beliefs, optimism and overconfidence, that can explain why individuals over-estimate the probability of high-payoff outcomes. Optimism occurs when individuals, independent of their own performance, over-estimate the probability of outcomes they prefer. Overconfidence occurs when individuals believe they perform better than they actually do. We find that optimism and overconfidence are positively correlated at the individual level and that both optimism and overconfidence help to explain why individuals over-estimate high-payoff outcomes. These findings challenge previous work that tends to focus solely on overconfidence to explain puzzling economic behavior, such as excess entry into self employment. The reason is that optimism is ignored, which can lead to an upwardly biased estimate of overconfidence. To illustrate the magnitude of the problem, we show that 30% of our observations are misclassified as under- or over-confident if optimism is omitted from the analysis.

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 Eraslan, 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

A robust finding in economics and psychology is that individuals tend to over-estimate the probability of high-payoff outcomes (De Bondt and Thaler, 1995). An example that has received considerable attention is the systematically biased belief among potential business owners that their business venture will be successful. Evidence suggests that the magnitude of this type of bias is considerable: 80% of nascent entrepreneurs believe their chances of success are at least 70%, but roughly two-thirds fail within the first few years (Cooper, Woo, and Dunkelberg, 1988). If individuals start businesses that are unprofitable or destined to fail—and if they do so because they systematically over-estimate the probability of success—then they may be making inefficient use of their personal resources and possibly forgo higher earnings in paid employment (Hamilton, 2000).[1]

Previous literature explaining why individuals over-estimate the probability of high-payoff outcomes has typically focused on overconfidence, which is the tendency to over-estimate one's own performance.[2] An alternative explanation is optimism or "wishful thinking", defined as a tendency to over-estimate the probability of preferred outcomes (Irwin, 1953; Weinstein, 1980). We argue that previous work confounds optimism and overconfidence. The reason is that most scenarios in which we discuss overconfidence, such as entrepreneurship or financial decision-making, also contain the possibility for optimism.[3]

In this paper, we present results from an experiment designed to study optimism and overconfidence as distinct but potentially related phenomena. We make three key contributions, which highlight the problems associated with studying optimism and overconfidence in isolation. First, we show that optimism and overconfidence are positively correlated at the individual level.[4] Second, we show that optimism and overconfidence jointly explain beliefs in settings where (i) individuals must assess their own performance and (ii) their performance affects their payoffs. Together, these results mean that inferring overconfidence in settings where individuals are optimistic can suffer from omitted variables bias. Our third result is

[1] Excess entry into self-employment may also waste government resources designed to encourage entrepreneurship. Governments and non-profits spend billions of dollars every year to subsidize self-employment Fairlie, Karlan, and Zinman (2015). However, Hamilton, Pande, and Papageorge (2015) provide evidence that subsidies tend to encourage relatively low-quality business ideas.

[2] Moore and Healy (2008) identify three types of overconfidence: (1) over-estimation: believing one's own performance or ability is better than it actually is; (2) over-placement: over-estimating one's own performance or ability relative to a reference group; and (3) over-precision: over-estimating the precision of one's knowledge. Throughout this study, we focus on the first type of overconfidence.

[3] This is consistent with the discussion in Åstebro et al. (2014) regarding the joint role of optimism and over-estimation in driving entrepreneurs' inflated beliefs of success.

[4] Åstebro, Jeffrey, and Adomdza (2007) report evidence of a similar type of correlation in cognitive biases in the context of entrepreneurship.

to show that ignoring the role of optimism in such settings results in the misclassification of overconfidence for nearly one-third of our observations.

Our experimental design is outlined in detail in Section 2 and further details are found in Appendix A.[5] For now, we explain the main features of the experiment using a simple urn example. We also emphasize how our experiment distinguishes our study from earlier work on overconfidence. In particular, we argue that previous research on overconfidence assesses beliefs in settings that are similar to one of our experimental treatments in which overconfidence and optimism are confounded.

Individuals participating in the experiment face an urn containing 1 white ball and 1 black ball and we elicit a probabilistic belief that a single draw from the urn will be white and pay the subject for the accuracy of his belief. This treatment is the "Baseline Treatment", where the elicited beliefs are referred to as the "Baseline Beliefs", which we denote z_{base}. To identify optimism, we vary whether there is a payoff-favorable outcome (the "Payment Treatment"), but still incentivize accuracy. We inform the individual that he will receive a side payment if a white ball is realized and again elicit his probabilistic belief about the single draw. This belief is denoted z_{pay}. Optimism (pessimism) is identified by comparing the individual's probabilistic beliefs when white is payoff-favorable versus when it is not. A subject is optimistic (pessimistic) if his reported belief increases (decreases) when white becomes payoff-favorable, $z_{pay} > (<) z_{base}$. Conceptually, optimism can occur if the decision-maker experiences anticipatory utility (Brunnermeier and Parker, 2005; Caplin and Leahy, 2001) or affective decision-making (Bracha and Brown, 2012), has preferences given by a rank dependent utility (Quiggin, 1982), or preferences consistent with subjective expected utility but with differing priors (Van den Steen, 2004) and differing technologies (Santos-Pinto and Sobel, 2005).[6]

To identify overconfidence, we incorporate the individual's beliefs about their own performance (the "Performance Treatment"). The individual answers an IQ question, where 1 white ball is added to the urn if the individual answers correctly (resulting in an urn with 2 white, 1 black and 1 draw) and does not add any ball otherwise. In other words, the individual's performance directly affects the distribution he faces—in this case increasing the probability that a white ball is drawn. Without feedback on his performance (whether or not the IQ question was correctly answered), the individual forms a belief about his performance and, therefore, the number of white balls he is facing. We elicit his belief about

[5] A copy of the full online appendix for this paper can be found at Online Appendix

[6] The experimental and empirical literature on optimism deals explicitly with environments similar to this experimental treatment and systematically finds beliefs consistent with optimism in that subjects overweight the probability of payoff-favorable outcomes that are independent of performance (Coutts, 2014; Irwin, 1953; Ito, 1990; Mayraz, 2011; Weinstein, 1980).

the probability that the single draw results in white.[7] We denote this belief z_{perf}. We say that the subject is overconfident (under-confident) in his performance if $z_{perf} > (<) z_{base}$.[8] Overconfidence, can occur, for example, if the decision-maker has belief-based preferences that permit ego-utility or self-image concerns (Bénabou and Tirole, 2002; Köszegi, 2006) or limited capability to process information (Rabin and Schrag, 1999; Bénabou and Tirole, 2002; Enke and Zimmermann, 2013).

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 overconfidence and optimism simultaneously. To simulate this type of environment in the laboratory, we combine the Payment and Performance Treatments into a "Combined Treatment", where a subject can increase the probability that a white ball is drawn through his performance (as in the Performance Treatment) and receives an additional side payment when a white ball is drawn (as in the Payment Treatment). Having examined overconfidence and optimism in isolation, we are able to examine to what degree beliefs in the Combined Treatment are explained by optimism, overconfidence (or both).

Previous studies examine overconfidence in environments that resemble the Combined Treatment, but attribute overestimation in these settings solely to overconfidence (Kirchler and Maciejovsky, 2002; Malmendier and Tate, 2008; Koellinger, Minniti, and Schade, 2007; Niederle and Vesterlund, 2007). Alternatively, they rely on the untested assumption that optimism and overconfidence are independent (Blavatskyy, 2009; Camerer and Lovallo, 1999; Hoelzl and Rustichini, 2005). Specifically, over-willingness to compete (Niederle and Vesterlund, 2007), suboptimal investment strategies (Malmendier and Tate, 2008) and market over-entry (Camerer and Lovallo, 1999) are all studied in settings, which embody the key features of the Combined Treatment: not only are individuals tasked with making a probabilistic assessment of their own performance, but their performance directly affects their payoffs.

In contrast, our experimental design allows us to assess whether optimism and overconfidence are independent. It also allows us to assess whether biases in the Combined Treatment are a good measure of overconfidence. As we show, the independence assumption is rejected. Moreover, optimism and overconfidence are significant components of beliefs biases in the Combined Treatment. This means that measures of bias in settings similar to the Combined

[7]Following Benoît and Dubra (2011) and as described in Section 2, we elicit beliefs about full distributions, which gives a more accurate description of subjects' beliefs.

[8]In Appendix B, which contains a number of robustness checks, we show that our main results are robust to using a the more standard measure of miscalibration as our measure of overconfidence (Lichtenstein, Fischhoff, and Phillips, 1977).

Treatment are not only capturing overconfidence, but are also capturing optimism. By comparing Performance Treatment Beliefs, which isolate overconfidence, to Combined Treatment beliefs, we are able to show that relying on the Combined Treatment to infer overconfidence results in the misclassification of nearly one third of our observations.

The remainder of this paper is organized as follows. Section 2 describes our experiment and Section 3 describes the experimental data. In Section 4, we present our main results and Section 5 concludes.

2 Experimental Design

The purpose of our experimental design is to separately measure optimism and overconfidence and to relate these measures to individual decisions in a setting where both biases can occur, i.e., a setting where subjects form beliefs about outcomes that they have preferences over and that can be influenced by their performance.

Each subject completes a common task facing each of the four treatments (described below). The common task consists of reporting probabilistic beliefs about realizations from six distributions, which are summarized in Table 1. The distributions are presented as computerized jars with various compositions of white and black balls. Subjects know 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 about the likelihood of a certain number of white balls being drawn. We elicit subjects' beliefs about the entire distribution.[9]

To incentivize reports of probabilistic beliefs, we pay subjects according to the quadratic scoring rule (QSR) (Brier, 1950; Murphy and Winkler, 1970).[10]

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

[9] In Appendix A, we provide additional details on the experiment, including screen shots of the computerized interface along with discussion of the comprehension quizzes and order of treatments. 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.

[10] We chose to incentivize beliefs using the QSR because of its simplicity, although it is only incentive-compatible under the assumption of risk-neutrality. 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. A binary lottery implementation of the quadratic scoring rule is theoretically incentive compatible and robust to risk preferences (McKelvey and Page, 1990), but experimental evidence suggests that it does not successfully induce risk-neutrality (Selten, Sadrieh, and Abbink, 1999) and the cognitive burden imposed on subjects may result in less reliable reports than the deterministic quadratic scoring rule (Rabin and Thaler, 2001).

Each of the four treatments described below were presented as separate Tasks. Subjects completed all of the questions in a single Task before moving on to the next Task. Within Task, the order of the distributions was randomized and a single Task (i.e. treatment) was randomly chosen at the end for payment.

The experiment was conducted at Washington University in Saint Louis in the MISSEL laboratory. Subjects were recruited via ORSEE (Greiner, 2004) and the experiment was conducted in Z-Tree (Fischbacher, 2007). In total 125 subjects participated in 15 sessions. On average, sessions lasted approximately 90 minutes and subjects earned $25 USD.

2.1 Treatments

The experiment is a within-subject 2×2 design, meaning subjects face all four treatment combinations: the Baseline Treatment, the Payment Treatment, the Performance Treatment and the Combined (Payment + Performance) Treatment. The Baseline Treatment elicits beliefs about known distributions and pays subjects for the accuracy of those probabilistic beliefs. Similarly, the Payment, Performance and Combined Treatments also pay subjects for the accuracy of beliefs, but each have an additional feature that will be described below.

In the Payment Treatment, in addition to being paid for the accuracy of beliefs, subjects are induced to prefer that white balls, instead of black balls, be drawn from the jar. This is operationalized by giving subjects a lottery ticket (side payment) that is independent of their payment for belief accuracy and increases in expected value when more white balls are drawn.

$$= \begin{cases} 10 & \text{with probability } \frac{\text{total white}}{\text{max white}} \\ 0 & \text{with probability } \left(1 - \frac{\text{total white}}{\text{max white}}\right). \end{cases}$$

In the Performance Treatment, subjects' performance on a trivia task influences which distribution they face. To explain: subjects start out facing Distribution 1 or 3. When starting with Distribution 1 (1 white, 1 black and 1 draw; see Table 1), the subject is given an IQ question and told that if he answers it correctly then another white ball will be added to his jar, in which case he faces Distribution 2 (2 white, 1 black, and 1 draw). Without feedback on the IQ question (i.e., subjects do not know whether their answer is correct), subjects report their belief about the likelihood that one draw from the jar is a white ball. Similarly, when subjects start in Distribution 3, they answer 3 IQ questions and a white ball is added to the jar for each correct IQ question, resulting in a final distribution that corresponds to Distribution 3, 4, 5 or 6. Without feedback, subjects are asked about the likelihood that 0, 1, 2, or 3 of the draws from the jar consist of white balls.

To gauge performance in the Performance Treatment, we ask subjects to answer multiple choice IQ questions from the Mensa Quiz book (Grosswirth, Salny, and Stillson, 1999).[11] 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 "hard-easy" effect (Lichtenstein and Fischhoff, 1977). We pay subjects \$2 for each correct IQ question so that subjects do not have an incentive to purposely give an incorrect answer to an IQ question to increase his certainty about the distribution he faces.[12]

In the Combined Treatment, we simultaneously apply both the Payment and Performance Treatment. Not only can subjects expect to make more money when more white balls are drawn from the jar (via the same side payment as in the Payment Treatment), but they can also influence the number of white balls in the jar by correctly answering IQ questions in the same manner as in the Performance Treatment. Thus, in the Combined Treatment, subjects can increase the likelihood of a higher payoff outcome. In this sense, the Combined Treatment contains the elements that are similar to scenarios outside of the laboratory. In many contexts, such as starting a business, individual performance increases, but does not guarantee, the likelihood of higher payoff outcomes.

Finally, return to Table 1, which summarizes the six distributions and how they relate to each other. Table 1 illustrates how we chose the six distributions used in the experiment so that subjects in the Performance Treatment or the Combined Treatment will always end up facing one of the six distributions that they face in the Baseline Treatment and the Payment Treatment. This means that we can make within-subject comparisons for the same distribution and the same treatment.

2.2 Key Features of the Experimental Design

In this section, we elaborate on two key features of our experiment: (1) the within-subject design and (2) the single unified task used across treatments. The within-subject design allows us to study belief changes at the individual level. In particular, we are interested in whether and how beliefs change when the potential for optimism or overconfidence is present. In order to do this, we must have an accurate measure of the individual's belief when neither bias is present, which is achieved in the Baseline Treatment. Thus, our measures

[11] This task was also used in Owens, Grossman, and Fackler (2012); Grossman and Owens (2012).

[12] Subjects start in Distribution 1 twice and Distribution 3 twice in both the Performance Treatment and the Combined Treatment. Also, the order of the distributions is randomized within each treatment.

of optimism and overconfidence will be a comparison between Payment, Performance and Combined Treatment Beliefs and the subject's Baseline Beliefs for the same distribution.

We elicit a set of Baseline Beliefs as a comparison (rather than comparing against the objective distribution) since there may be systematic departures from the objective distribution.[13] The use of an individual-level control means that any factors that affects individual reports uniformly across treatments, including poor mathematical skills or curvature of the utility function, do not drive our results as they are effectively netted out. For example, if poor math skills lead a subject to over-estimate probabilities, then the over-estimation induced by poor math skills occurs in all treatments, including the Baseline Treatment. Thus, we are able to net out its impact on reported beliefs and to isolate the experimental treatment effects.[14]

Finally, we note that although a within-subject design allows for clean identification of optimism and overconfidence, there is a potential for order effects due to the sequence in which subjects face each of the four treatments. Accordingly, we have run sessions in 5 different orders, which allows us to control for order and test for robustness of our results when the experiment is run in different orders. A discussion of the orders, along with other robustness checks, are found in Appendix B. Within each of the five treatment orders, the order of the distributions were randomized at the subject-level.

A second key feature of our experimental design is that the variable of interest in all treatments is the subject's probabilistic belief about white balls drawn from a jar with one of six mixtures of white and black balls. This commonality across treatments means that it is straightforward to compare magnitudes of optimism and overconfidence at the individual-level, as well as to directly relate beliefs reported in the Combined Treatment to beliefs reported in the Payment and Performance Treatments.[15]

[13] In Appendix B, where we present robustness checks, we show that the absolute differences between the Baseline Beliefs and the objective are larger for subjects with lower IQ and, unsurprisingly, when subjects face complex distributions, where three white balls are drawn. However, we also show that misclassification of overconfidence, which constitutes one of our key findings, does not vary significantly by IQ, gender or type of distribution.

[14] In Appendix C, we show how relying on within-subject shifts helps to dispel concerns about risk aversion. Still, we note that it is possible for individuals to hedge in the payment treatment. If so, then we would under-estimate the degree to which they are optimistic.

[15] We never directly ask subjects about their own assessment of their performance on the IQ questions since our experimental design allows us to impute the subject's belief about the probability of having correctly answered. Alternatively, it is possible to elicit the subject's belief about his performance and then impute z_{perf}. However, doing so would mean we lose uniformity of the experimental task across treatments and also raise the subjects' cognitive burden. In Appendix B, as part of a robustness test, we impute the subject-specific $\hat{p}$ in the single-draw distribution.

3 Data Description and Preliminary Data Analysis

Before describing features of the data, we will explain how we obtained the sample sizes for our analyses (see Appendix A for additional details on experimentally generated data). 125 individuals across 15 sessions were asked to report beliefs about the number of white balls being drawn from six different distributions and under four different experimental treatments. Beliefs were elicited 20 times for each individual, resulting in 2,500 observations (subject-distribution pairs). In the Performance and the Combined Treatments, subjects start in Distribution 1 (3) twice and can move to Distribution 2 (4,5 or 6) if they answer 1 (1,2 or 3) IQ questions correctly. However, if the individual answers 0 IQ questions correctly in both rounds, then he reports beliefs for Distribution 1 (3) twice. In such cases, we average over the individual's responses, leaving 2,296 observations.[16] Additionally, we drop 52 observations in which individual reports are not consistent with positive marginal probabilities, leaving 2,244 observations.[17]

3.1 Measuring Optimism and Overconfidence

For each individual and for each distribution, we elicit beliefs under the Baseline Treatment. The Baseline Treatment establishes a subject-level control, against which we compare beliefs in the other treatments. We measure optimism and overconfidence as within-subject shifts (relative to Baseline Beliefs) in the Payment and Performance Treatments, respectively.

Formally, for individual i facing distribution d under treatment τ, we define shifts relative to Baseline Beliefs as follows:

$$\overline{shift}_{i,d,\tau} \equiv \frac{1}{M_d} \sum_{m=1}^{M_d} \left[z_{i,d,m,\tau} - Truth_{d,m} \right] - \left[z_{i,d,m,\tau=B} - Truth_{d,m} \right], \tag{1}$$

where M_d is the number of moments for distribution d, $z_{i,d,m,\tau}$ are beliefs reported by individual i facing moment m of distribution d under treatment τ. $Truth_{d,m,\tau}$ is the objective probability for moment m of distribution d, and $\tau = B$ refers to beliefs elicited under the Baseline Treatment. Notice that we can rewrite equation (1) as

$$\overline{shift}_{i,d,\tau} = \frac{1}{M_d} \sum_{m=1}^{M_d} \left[z_{i,d,m,\tau} - z_{i,d,m,\tau=B} \right]. \tag{2}$$

[16]Alternatively, we could randomly choose one set of beliefs. Main results are robust to these changes.

[17]For example, if a subject reports that the probability of drawing either one or two white balls is 20% and that the probability of drawing one white ball is 40%, answers are not consistent with probabilistic beliefs. Results are unchanged if these observations are re-coded to be consistent with non-negative probabilities.

The variable $\overline{shift}$ is therefore the average difference between beliefs reported under treatment τ and beliefs reported under the Baseline Treatment ($\tau = B$) for the same individual and distribution. In the Payment Treatment, $\overline{shift}$ captures optimism by measuring within-subject shifts in beliefs due to the presence of a side payment for white balls. In the Performance Treatment, $\overline{shift}$ captures overconfidence by measuring within-subject shifts in beliefs due to changes in how IQ answers affect the distribution of white balls. In the Combined Treatment, $\overline{shift}$ captures shifts in beliefs when subjects face uncertainty, can affect the distribution through their IQ performance and also receive side payments for each white ball that is drawn.

Our definition of $\overline{shift}$ nets out Baseline Beliefs. Therefore, we only define it for

$$\tau \in \{Payment, Performance, Combined\}$$

Shifts relative to the Baseline Beliefs under the Payment Treatment are called "Optimistic Shifts". Shifts under the Performance Treatment are called "Overconfident Shifts". Shifts under the Combined Treatment are called "Combined Treatment Shifts".[18]

We note that netting out Baseline Beliefs reduces the sample size since, for each distribution, each individual loses the Baseline Beliefs observation. The result is a sample reduction of 726 observations, leaving 2,244−726=1,518 observations for which $\overline{shift}$ is defined. These observations are comprised of 738 within-subject shifts comparing the Payment Treatment to the Baseline Beliefs control, 385 for the Performance Treatment and 395 for the Combined Treatment. Of the 125 individuals in the sample, 59 are male and 66 are female.

Finally, we note that $\overline{shift}$ constitutes one way to construct an experimentally-induced beliefs shift variable. We consider alternative definitions in Appendix B and find that results remain largely unchanged. For example, in Appendix B, we show that our main results are robust if we specify that a shift occurs when only if elicited beliefs first-order stochastically dominate the objective distribution.

3.2 Measuring Within-Subject Correlation

The main analysis in Section 4 focuses on how optimism, overconfidence and beliefs in the Combined Treatment are related at the individual-level. Doing so places additional burden onto the data since we need to observe the same individual in multiple treatments for the same distribution, which is not always possible in the Performance and Combined Treatments

[18] In Appendix D, we provide formal definitions of optimism and overconfidence. Next, we use these definitions to derive definitions for Optimistic Shifts and Overconfident Shifts.

given that the distribution subjects face depends on his answers to the IQ questions. 383 individuals are observed in the same distribution in both the Payment and the Performance Treatments and 247 individuals are observed in the same distribution in all three treatments. These are the sample sizes we use when comparing within-subject correlations in responses to treatments.[19]

3.3 Average Treatment Effects

In this section, we study average treatment effects for the 1,518 observations of the variable $\overline{shift}$ defined in equation (1).[20] In Table 2, we present estimates from OLS regressions where the outcome variable is $\overline{shift}$ and explanatory variables include experimental treatments (payment, performance or combined), gender, correctly answered IQ questions. The specification we use is

$$\overline{shift}_{i,d,\tau} = \sum_{\tau} \mathbf{1}[treatment = \tau]\psi_{\tau} + X_{i,d}\delta + e_{i,d,\tau} \tag{3}$$

We also include distribution and order dummy variables. In the first specification (Column [1]), we only control for treatments and find that the Payment Only treatment induces no average shift, but that the Performance and Combined Treatments lead people to over-estimate the number of white balls (relative to the Baseline Beliefs).[21]

However, once we control for possible order effects by adding dummy variables for each order (column [2]) and distribution dummy variables (Column[3]) we find a significant increase in the estimated treatment effects. Finally, in Columns [4] and [5] we control for gender and IQ, measured as the total number of correctly answered IQ questions during the entire experiment.[22] We find that, on average, men and women are equally likely to over-estimate the number of white balls that are drawn and that the treatment effects decrease in the number of correctly answered IQ questions. One possibility is that answering more IQ questions

[19]In Appendix A, which contains more detailed information on the data generated by the experiment, we show that these samples are composed nearly evenly of males versus females and of low versus high IQ individuals, which helps to dispel concerns that our main results are identified off of a non-randomly selected sample.

[20]In Appendix A, we plot histograms of $\overline{shift}$ for all 1,518 observations and then separately for the payment, Performance and Combined Treatments. We demonstrate that the mode in all four panels is zero, meaning that a plurality of subjects shows no significant departures in their reported beliefs across treatments. However, there are non-zero observations, which means that some subjects are optimistic (pessimistic) and overconfident (under-confident).

[21]This is consistent with other laboratory studies of optimism, which find small average treatment effects in pure optimism (Barron, 2015).

[22]Throughout the experiment, subjects are asked to answer 16 IQ questions. The mean number of correct answers across subjects is 9.23 and the standard deviation is 0.53.

correctly is indicative of a stronger ability to calculate probabilities, which means that individuals might be less prone to shift in response to experimental treatments.[23] Column [6] is consistent with this interpretation since answering more IQ questions does not directly affect the distribution in the Payment Treatment, but subjects who answered more IQ questions correctly also display smaller treatment effects in the Payment Only Treatment.[24]

4 Main Results

In this section, we present our main results. They are (i) optimism and overconfidence are positively correlated; (ii) both optimism and overconfidence help to explain why individuals facing uncertainty over-estimate high-payoff outcomes; and (iii) misclassification of under and over-confidence occurs if these correlation patterns are ignored.

4.1 How Optimism Relates to Overconfidence

In this section, we study the relationship between optimism and overconfidence by relating $\overline{shift}$ for the Payment and Performance Treatments. The Payment Treatment captures optimism since performance plays no role, but individuals are induced to prefer that white balls be drawn instead of black balls. The Performance Treatment captures overconfidence since subjects must form beliefs over their ability to answer IQ questions, but are not induced to prefer that white balls are drawn. In Figure 1, we plot Optimistic Shifts against Overconfident Shifts, where each observation is for an individual facing the same distribution in both the Payment and the Performance Treatments. The figure shows clear evidence of a positive convex relationship. The interpretation is that optimistic individuals tend to be overconfident. Conversely, pessimistic individuals tend to be under-confident, though the latter relationship is weaker.

Next, we ask if this correlation is robust when we control for different sets of covariates. We use OLS to estimate equations of the following form:

$$\overline{shift}_{i,d,performance} = \overline{shift}_{i,d,payment}\phi_1 + X_{i,d}\beta_1 + \epsilon_{i,d,\tau} \tag{4}$$

[23]A second possibility is that subjects who answer more IQ questions correctly are also more likely to (correctly) believe that they have given the right answer and thus uncertainty over which urn they face is reduced. However, controlling for distribution should rule out the latter possibility.

[24]In Appendix B, we provide further evidence that our main results our robust to selection into distributions due to variation in how many IQ questions are correctly answered. In particular, we show that average treatment effects are similar if we focus on the subsample of observations used in our main analysis on Combined Treatment Shifts, where we require that the individual face the given distribution across all experimental treatments.

We estimate this model for the 383 observations where the individual is observed in the Baseline, Payment and Performance Treatments for the same distribution.[25] Results are presented in Table 3. In the first four specifications, we add varying sets of controls and find that the correlation between optimism and overconfidence is positive and significant. Column [1] of Table 3 suggests that a 10 percentage point increase in a subject's optimism is associated with 6.6 percentage point increase in his overconfidence. Next, we show that gender and IQ do not affect the relationship.[26] In the fourth column, we limit attention to the 247 observations that we use in subsequent analysis, where we observe the individual in the same distribution for each of the four treatments. Finally, in Column [5], we permit a second-order polynomial and find that the coefficient is significant and positive.[27] This pattern is consistent with the relationship evident in Figure 1, which shows that the positive correlation is stronger at higher levels.[28]

One possible concern is that the correlation between optimism and overconfidence is driven by variation in the Baseline Belief, rather then differences in the Payment Treatment or Performance Treatment beliefs. We address this possibility in Appendix B by examining the correlation between Payment Treatment and the Performance Treatment beliefs and use the Baseline Belief as a control (rather than netting it out as with the $\overline{shift}$ variable) and obtain qualitatively equivalent results. In Appendix B.3, we show that optimism is significantly positively correlated with over-calibration (an alternative measure of overconfidence).

4.2 How Optimism and Overconfidence Affect Beliefs

Next, we relate the same measures of overconfidence and optimism to shifts in the Combined Treatment. Here, the goal is to assess how optimism and overconfidence relate to beliefs in a setting where both may be present. We begin by separately plotting optimism and overconfidence with shifts in the Combined Treatment (Figure 2 and Figure 3, respectively). In both cases, there is clear evidence of positive correlation. Next, we regress shifts in the Combined Treatment onto optimism and overconfidence and are therefore limited to the 247 observations where the same individual is observed in the same distribution for all four

[25] We allow clustered by individual to control for individual level heterogeneity in variance.

[26] We also fully interact optimistic shifts with the various orders that subjects performed each treatment. While there is some variation, only in the fifth order do we find a weaker correlation that is not significantly different from zero.

[27] In results not shown, but available from the authors, we also show that results are robust to the inclusion of individual fixed effects.

[28] An alternative specification would be to include an interaction of $\overline{shift}_{i,d,payment}$ when it is positive. When we do this, we find that the coefficient on the interaction is positive. However, it is insignificant at conventional levels ($p = 0.12$).

treatments.

$$\overline{shift}_{i,d,combined} = \overline{shift}_{i,d,payment}\phi_2 + \overline{shift}_{i,d,performance}\phi_3 + X_{i,d}\beta_2 + \eta_{i,d,\tau} \quad (5)$$

Columns [1] and [2] of Table 4 show that overconfidence and optimism are positively correlated to shifts in the Combined Treatment, respectively. In Column [3], we regress shifts in the Combined Treatment onto both optimism and overconfidence. Since optimism and overconfidence are positively correlated, the estimates shrink in comparison to Column [1] and [2] due to omitted variables bias. Specifically, when we regress shifts in the Combined Treatment onto overconfidence, the estimated coefficient on overconfidence is upwardly biased since it also captures the positive correlation between optimism and overconfidence. This finding is concerning as it means that inferring overconfidence in scenarios where individuals also have preferences over outcomes is susceptible to omitted variables bias due to the presence of optimism.[29]

Further, in Appendix B.3, we find equivalent results when we restrict our sample to subjects who are sufficiently miscalibrated.

4.3 Misclassification of Overconfidence

A key motivation underlying our experimental design is our claim that overconfidence can be confounded with optimism. According to our results, in scenarios where individuals are prone to both overconfidence and optimism (i.e., our Combined Treatment), attributing overestimation solely to overconfidence results in an omitted variable bias due to the presence of optimism. Nevertheless, as we pointed out in Section 1, previous literature uses beliefs akin to those measured in our Combined Treatment (where both optimism and overconfidence may be present) to identify overconfidence. In this section, we explore the degree of misclassification that results from this approach: how wrong are we, as the researcher, if we ignore optimism and only elicit beliefs in "Combined Treatments" and call it overconfidence?

First, we formally test whether beliefs shifts in the Combined Treatment are equal to beliefs shifts in the Performance Treatment (overconfidence) using an F-test of the joint hypothesis that $\phi_2 = 1$ and $\phi_3 = 0$ in equation (5). We conduct this test for each of the models presented in Table 4 and, in each case, we reject the null hypothesis at the

[29] In results that are available upon request, we check for non-linearities. We permit second order polynomials in both treatment effects, interactions between treatments effects and also interact each treatment effects with a dummy for positive treatment effects. We do this for each treatment effect separately and also together. We cannot reject the null hypothesis of linearity in any case, which includes joint tests of significance of parameters on non-linearities.

1% level. The F-statistic ranges between 10.85 and 16.96, thus we can soundly reject the hypothesis that belief biases in the Combined Treatment are equal to belief biases in the Performance Treatment. In other words, our experiment casts doubt on the reliability of the common practice of measuring overconfidence using beliefs in scenarios where individuals have preferences over outcomes and so optimism is possible.

Second, Figure 4 plots beliefs in the Combined Treatment against beliefs in the Performance Treatment, where the x-axis is belief shifts in the Performance Treatment (overconfidence) and the y-axis is beliefs shifts in the Combined Treatment. If overconfidence is equivalent to beliefs in the Combined Treatment (as previous research implicitly implies), then the data would fall exclusively on the 45-degree line. We can reject the null hypothesis that the 45-degree line provides a good fit for the data (p-value<0.01).

We identify four types of misclassification; observations in (1) the upper-left quadrant; (2) the lower-right quadrant; (3) along the x-axis (excluding the origin); and (4) along the y-axis (excluding the origin). In the upper-left (lower-right) quadrant, the bias in the Combined Treatment is positive (negative), but the bias in the Performance Treatment is negative (positive). That is, if the researcher relies on beliefs in the Combined Treatment to infer overconfidence, then in 9% of the observations an under-confident individual is mis-classified as over-confident (p-value<0.001) and in 6% of the observations an over-confident individual is mis-classified as under-confident (p-value<0.001). Thus, 15% of observations are mis-classified with an *opposite* belief bias.

A related error occurs if an observation lies on the y-axis (but not including the origin). In these cases, the individual is neither overconfident nor under-confident when we account for the role of optimism, but is classified as such if the researcher relies on beliefs in the Combined Treatment to infer overconfidence. This occurs in 7% of cases. The fourth error occurs for observations on the x-axis (not including the origin), where the individual is not classified as over- or under-confident using beliefs in the Combined Treatment, even though the individual is. This classification error occurs in 7% of cases. In total, misclassification occurs in 29% of observations.[30]

The remaining 71% of the data lie in the upper-right and lower-left quadrant, indicating that the direction of the bias in the Combined Treatment and Performance Treatment are the same. However, even if we restrict our analysis to just this subset we can reject the null hypothesis that the 45-degree line in Figure 4 fits the data (p-value<0.01). Thus, for

[30] If we repeat the exercise for different groups (e.g., by gender), the proportions are similar. However, the magnitude of the difference between beliefs in the treatment groups does appear to vary by observable groups. As we show in Appendix E the magnitude of the difference between beliefs shifts in the performance versus the Combined Treatment is relatively large for females, low IQ individuals and for complex distributions.

this set of observations, we would correctly classify the individual as overconfident or underconfident, but the magnitude of the bias is still mis-estimated.

Together, these results illustrate the value of our experimental approach. By isolating optimism and overconfidence, we can assess whether the beliefs measured in the Combined Treatment, which better mimics real-world scenarios, permit effective identification of overconfidence. We show that in nearly one third of cases misclassification occurs and that even for the correctly classified observations the degree of bias is, on average, mis-estimated. This casts doubt on earlier literature, which elicits beliefs of individuals in scenarios where they must assess their own performance and have preferences over outcomes. Upward biases in beliefs in these scenarios are deemed "overconfidence". In contrast, we show that the researcher is likely capturing a mixture of overconfidence and optimism and, further, that about a third of individuals are misclassified.

5 Conclusion

Overconfidence has been identified as a widespread phenomenon, affecting financial, entrepreneurial and managerial decision-making. We show that the correlation between optimism and overconfidence implies that individuals can be classified as overconfident when they are not. Beyond the issue of mis-measuring overconfidence, our findings raise concerns regarding interventions aimed at de-biasing beliefs to encourage informed decision-making, e.g., information interventions.[31] Presumably, the effectiveness of information interventions relies on the type of bias being corrected and the information being provided to correct it. Previous work has demonstrated that the way in which individuals incorporate new information may be a function of whether their biases are due to overconfidence (Eil and Rao, 2011; Ertac, 2011; Charness, Rustichini, and van de Ven, 2011; Mobius et al., 2011; Barron, 2015). Mobius et al. (2011) show that when subjects receive information about their IQ score, they heavily weight their prior belief, update asymmetrically (weighting favorable information more heavily than unfavorable information) and are increasingly averse to receiving correct information the more overconfident they are. Therefore, incorrectly measuring overconfidence, in particular, failing to distinguish it from optimism, may lead poorly designed information interventions with limited effectiveness at improving decision-making.

[31] Information to encourage and improve entrepreneurial prospects can come from a variety of sources, including peers (Minniti, 2005; Nanda and Sørensen, 2010; Lerner and Malmendier, 2011; Field et al., 2015), employers (Gompers, Lerner, and Scharfstein, 2005) and government programs (Fairlie, Karlan, and Zinman, 2015).

References

Åstebro, Thomas, Holger Herz, Ramana Nanda, and Roberto A Weber. 2014. "Seeking the Roots of Entrepreneurship: Insights from Behavioral Economics." *The Journal of Economic Perspectives* 28 (3):49–69.

Åstebro, Thomas, Scott A Jeffrey, and Gordon K Adomdza. 2007. "Inventor Perseverance after Being Told to Quit: The Role of Cognitive Biases." *Journal of Behavioral Decision Making* 20 (3):253–272.

Barron, Kai. 2015. "Belief Updating: An Experimental Test of Bayes Rule and the 'Good-News, Bad-News' Asymmetry." Working Paper.

Bénabou, R. and J. Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics* 117 (3):871–915.

Benoît, Jean-Pierre and Juan Dubra. 2011. "Apparent Overconfidence." *Econometrica* 79 (5):1591–1625.

Blavatskyy, P.R. 2009. "Betting on Own Knowledge: Experimental Test of Overconfidence." *Journal of Risk and Uncertainty* 38 (1):39–49.

Bracha, Anat and Donald J Brown. 2012. "Affective Decision Making: A Theory of Optimism Bias." *Games and Economic Behavior* 75 (1):67–80.

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

Brunnermeier, M.K. and J.A. Parker. 2005. "Optimal Expectations." *American Economic Review* 95 (4):1092–1118.

Camerer, C. and D. Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review* 89 (1):306–318.

Caplin, Andrew and John Leahy. 2001. "Psychological Expected Utility Theory and Anticipatory Feelings." *Quarterly Journal of Economics* :55–79.

Charness, Gary, Aldo Rustichini, and Jeroen van de Ven. 2011. "Overconfidence, Self-Esteem, and Strategic Deterrence." Working Paper.

Cooper, Arnold C, Carolyn Y Woo, and William C Dunkelberg. 1988. "Entrepreneurs' Perceived Chances for Success." *Journal of Business Venturing* 3 (2):97–108.

Coutts, Alexander. 2014. "Testing Models of Belief Bias: An Experiment." Working Paper.

De Bondt, Werner FM and Richard H Thaler. 1995. "Financial Decision-Making in Markets and Firms: A Behavioral Perspective." *Handbooks in Operations Research and Management Science* 9:385–410.

Eil, David and Justin M Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." *American Economic Journal: Microeconomics* 3 (2):114–138.

Enke, Benjamin and Florian Zimmermann. 2013. "Correlation Neglect in Belief Formation." CESifo Working Paper Series.

Ertac, Seda. 2011. "Does Self-Relevance Affect Information Processing? Experimental Evidence on the Response to Performance and Non-Performance Feedback." *Journal of Economic Behavior & Organization* 80 (3):532–545.

Fairlie, Robert W, Dean Karlan, and Jonathan Zinman. 2015. "Behind the GATE Experiment: Evidence on Effects of and Rationales for Subsidized Entrepreneurship Training." *American Economic Journal: Economic Policy* 7 (2):125–61.

Field, Erica, Seema Jayachandran, Rohini Pande, and Natalia Rigol. 2015. "Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship?" *American Economic Journal: Economic Policy, Forthcoming* .

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171–178.

Gompers, P., J. Lerner, and D. Scharfstein. 2005. "Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1999." *The Journal of Finance* 60 (2):577–614.

Greiner, Ben. 2004. "An Online Recruitment System for Economic Experiments." University of Cologne Working Paper Series.

Grossman, Zachary and David Owens. 2012. "An Unlucky Feeling: Overconfidence and Noisy Feedback." *Journal of Economic Behavior & Organization* .

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

Hamilton, Barton H, Nidhi Pande, and Nicholas W Papageorge. 2015. "The Right Stuff? Personality and Entrepreneurship." Mimeo, Johns Hopkins University.

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

Hoelzl, E. and A. Rustichini. 2005. "Overconfident: Do You Put Your Money On It?" *The Economic Journal* 115 (503):305–318.

Irwin, Francis W. 1953. "Stated Expectations as Functions of Probability and Desirability of Outcomes." *Journal of Personality* 21 (3):329–335.

Ito, Takatoshi. 1990. "Foreign Exchange Rate Expectations: Micro Survey Data." *American Economic Review* 80 (3):434–49.

Kirchler, E. and B. Maciejovsky. 2002. "Simultaneous Over-and Underconfidence: Evidence from Experimental Asset Markets." *Journal of Risk and Uncertainty* 25 (1):65–85.

Koellinger, P., M. Minniti, and C. Schade. 2007. ""I Think I Can, I Think I Can": Overconfidence and Entrepreneurial Behavior." *Journal of Economic Psychology* 28 (4):502–527.

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

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

Lichtenstein, Sarah and Baruch Fischhoff. 1977. "Do Those Who Know More Also Know More About How Much They Know?" *Organizational Behavior and Human Performance* 20 (2):159–183.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D Phillips. 1977. *Calibration of Probabilities: The State of the Art*. Springer.

Malmendier, U. and G. Tate. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction." *Journal of Financial Economics* 89 (1):20–43.

Mayraz, Guy. 2011. "Wishful Thinking." Mimeo, University of Melbourne.

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

Minniti, Maria. 2005. "Entrepreneurship and Network Externalities." *Journal of Economic Behavior & Organization* 57 (1):1–27.

Mobius, M.M., M. Niederle, P. Niehaus, and T.S. Rosenblat. 2011. "Managing Self-Confidence: Theory and Experimental Evidence." NBER working paper.

Moore, D.A. and P.J. Healy. 2008. "The Trouble with Overconfidence." *Psychological review* 115 (2):502.

Murphy, Allan H and Robert L Winkler. 1970. "Scoring Rules in Probability Assessment and Evaluation." *Acta Psychologica* 34:273–286.

Nanda, Ramana and Jesper B Sørensen. 2010. "Workplace Peers and Entrepreneurship." *Management Science* 56 (7):1116–1126.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete too much?" *Quarterly Journal of Economics* 122 (3):1067–1101.

Owens, David, Zachary Grossman, and Ryan Fackler. 2012. "The Control Premium: A Preference for Payoff Autonomy."

Quiggin, John. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization* 3 (4):323–343.

Rabin, M. and J.L. Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *Quarterly Journal of Economics* 114 (1):37–82.

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

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

Selten, R., A. Sadrieh, and K. Abbink. 1999. "Money Does not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse." *Theory and Decision* 46 (3):213–252.

Van den Steen, E. 2004. "Rational Overoptimism (and Other Biases)." *American Economic Review* 94 (4):1141–1151.

Weinstein, Neil D. 1980. "Unrealistic Optimism about Future Life Events." *Journal of Personality and Social Psychology* 39 (5):806–820.

6 Tables and Figures

TABLE 1: EXPERIMENTAL DESIGN AND DISTRIBUTIONS

	Distribution Details			Performance
Distribution	# of White	# of Black	# of 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	

Subjects face 6 distributions in the Baseline Treatment and the Payment Treatment. In the Performance Treatment and the Combined Treatment subjects start in Distribution 1 twice and Distribution 3 twice and one white ball is added to the starting distribution for each correctly answered IQ question. Subjects starting in Distribution 1 answer 1 IQ 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. When subjects start with Distribution 3, they answer 3 IQ questions and may add 0, 1, 2 or 3 white balls, therefore facing Distributions 3, 4, 5, or 6.

Table 2: Average Treatment Effects: Regressions

	[1]	[2]	[3]	[4]	[5]	[6]
Payment Treatment Beliefs	0.003	0.01*	0.06***	0.07***	0.11***	0.1***
Performance Treatment Beliefs	0.03***	0.04***	0.09***	0.09***	0.14***	0.16***
Combined Treatment Beliefs	0.02***	0.04***	0.09***	0.09***	0.14***	0.16***
Male	.	.	.	-0.007	0.0005	0.0006
Correct IQ Answers	.	.	.	.	-0.006***	.
IQ × Pay.	.	.	.	.	.	-0.004*
IQ × Perf.	.	.	.	.	.	-0.008**
IQ × Combined	.	.	.	.	.	-0.008**
Observations	1518	1518	1518	1518	1518	1518
R^2	0.02	0.03	0.07	0.07	0.08	0.08
Order Dummies	[N]	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Baseline Beliefs minus beliefs reported in the remaining three treatments, i.e., the $\overline{shift}$ variable pooled across the Payment, Performance and Combined Treatments. Robust Standard Errors in parentheses and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3: Correlation between Optimism and Overconfidence

	[1]	[2]	[3]	[4]	[5]
Optimistic Shift	0.66***	0.65***	0.64***	0.62***	0.51***
Optimistic Shift-Squared	.	.	.	.	0.62***
Male	.	0.01	0.006	0.02*	0.02
Correct IQ Answers	.	-0.009***	-0.003	-0.004	-0.002
Constant	0.03***	0.11***	0.12***	0.13***	0.11***
Observations	383	383	383	247	247
R^2	0.32	0.34	0.41	0.46	0.48
Order Dummies	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]
Smaller Sample	[N]	[N]	[N]	[Y]	[Y]

This table shows estimates from an OLS regression where the outcome variable is Baseline Beliefs minus Performance Treatment Beliefs ($\overline{shift}_{i,d,performance}$) and the explanatory variable of interest is Baseline Beliefs minus Payment Treatment Beliefs ($\overline{shift}_{i,d,payment}$), which we denote "Optimistic Shift". Robust Standard Errors in parentheses and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 4: The Role of Optimism and Overconfidence

	[1]	[2]	[3]	[4]	[5]
Overconfident Shift	0.65^{***}	.	0.5^{***}	0.56^{***}	0.5^{***}
Optimistic Shift	.	0.59^{***}	0.27^{***}	0.25^{***}	0.28^{***}
Male	.	.	.	0.02	0.02
Correct IQ Answers	.	.	.	-0.002	0.001
Constant	0.05^{***}	0.11^{***}	0.06^{***}	0.04	0.04
Observations	247	247	247	247	247
R^2	0.51	0.41	0.54	0.52	0.54
Order Dummies	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[N]	[N]	[Y]
F-test statistic ($\phi_2 = 1$ & $\phi_3 = 0$)	10.85	12.37	16.26	16.96	16.60

This table shows estimates from an OLS regression where the outcome variable is Baseline Beliefs minus Combined Treatment Beliefs ($\overline{shift}_{i,d,combined}$) and the explanatory variables of interest are Baseline Beliefs minus Payment Treatment Beliefs ($\overline{shift}_{i,d,payment}$) which we denote "Optimistic Shift", and Baseline Beliefs minus Performance Treatment Beliefs ($\overline{shift}_{i,d,performance}$) which we denote "Overconfident Shift". The F-test statistic is for the joint hypothesis that the coefficient on Overconfident Shifts is equal to 1 and that the coefficient on Optimistic Shifts is equal to zero, i.e., the null hypothesis is that biases in beliefs in the Performance Treatment (which isolates overconfidence) are equal to biases in the Combined Treatment. The null hypothesis is rejected in all model specifications at the 1% level. Robust Standard Errors in parentheses and *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

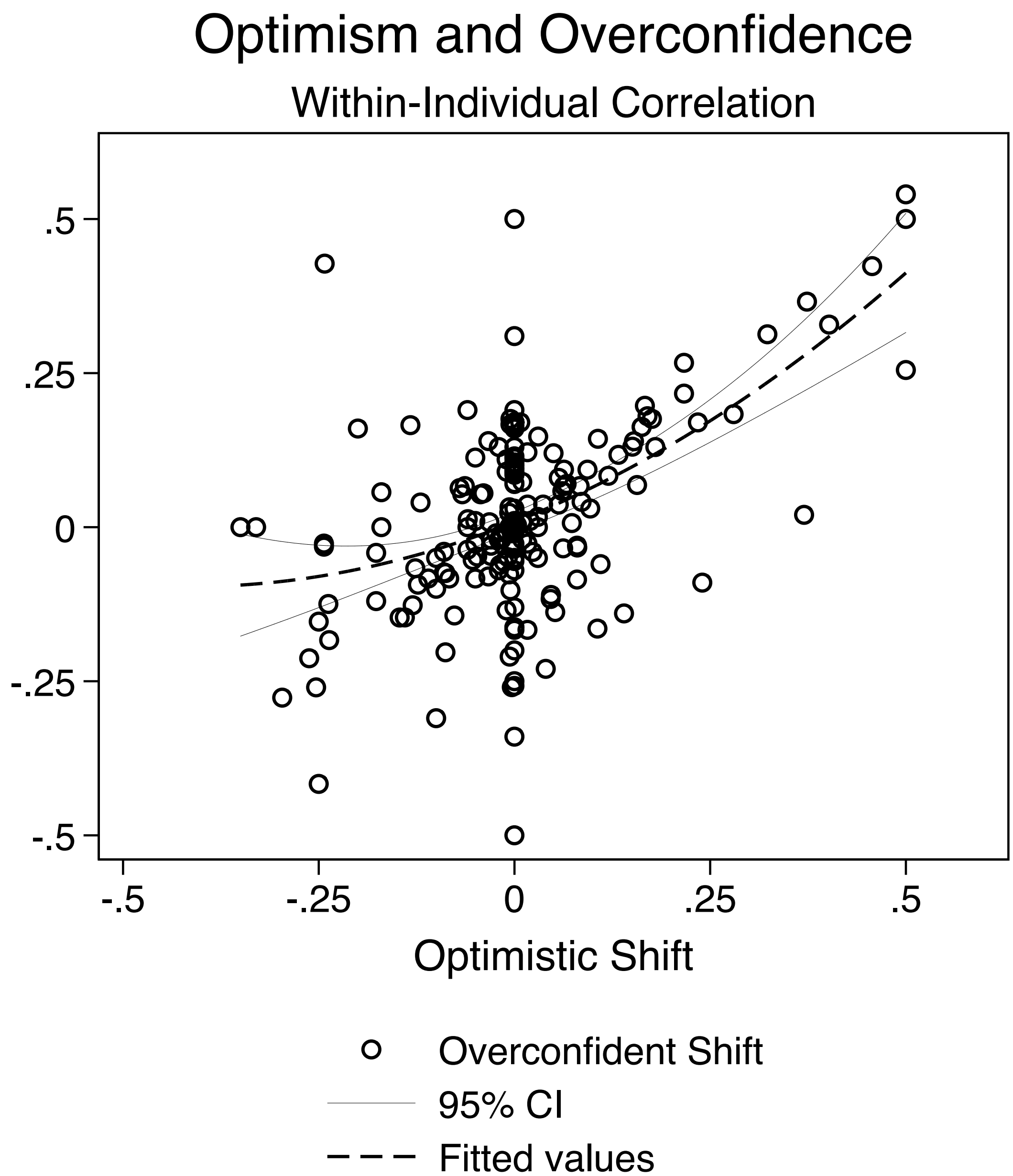

Figure 1: Optimism and Overconfidence: This figure relates shifts in beliefs in the Payment and Performance Treatments. This plot shows evidence that optimistic individuals also tend to be overconfident.

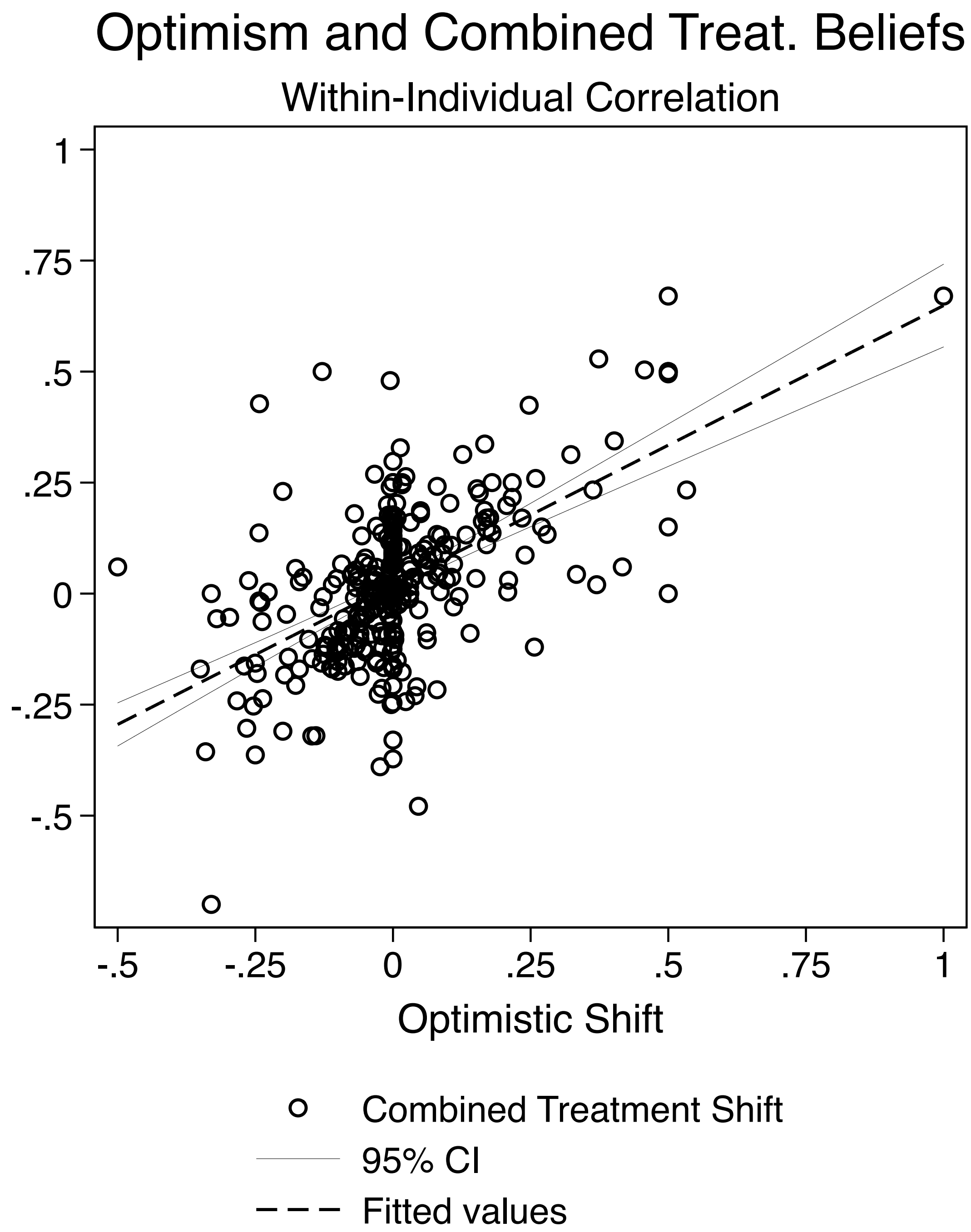

FIGURE 2: OPTIMISM AND BELIEFS IN THE COMBINED TREATMENT: This figure relates shifts in beliefs in the Payment and Combined Treatments. This plot shows suggestive evidence that optimistic individuals also tend to over-estimate high-payoff outcomes in the Combined Treatment, where individuals have preferences over outcomes and performance plays a role in the distribution individuals face.

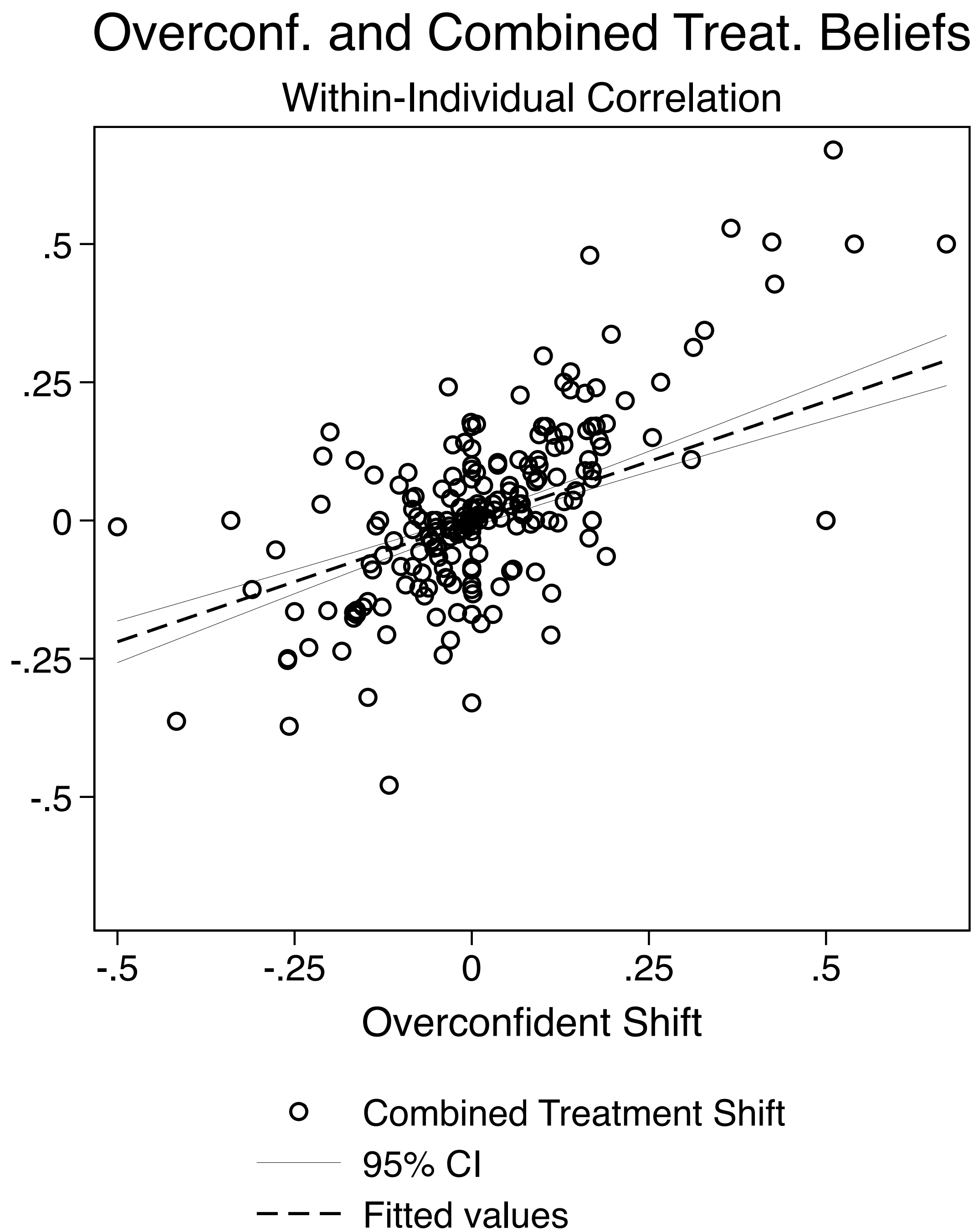

Figure 3: Overconfidence and Beliefs in the Combined Treatment: This figure relates shifts in beliefs in the Performance and Combined Treatments. This plot shows suggestive evidence that overconfident individuals also tend to over-estimate high-payoff outcomes in the Combined Treatment, where individuals have preferences over outcomes and performance plays a role in the distribution individuals face.

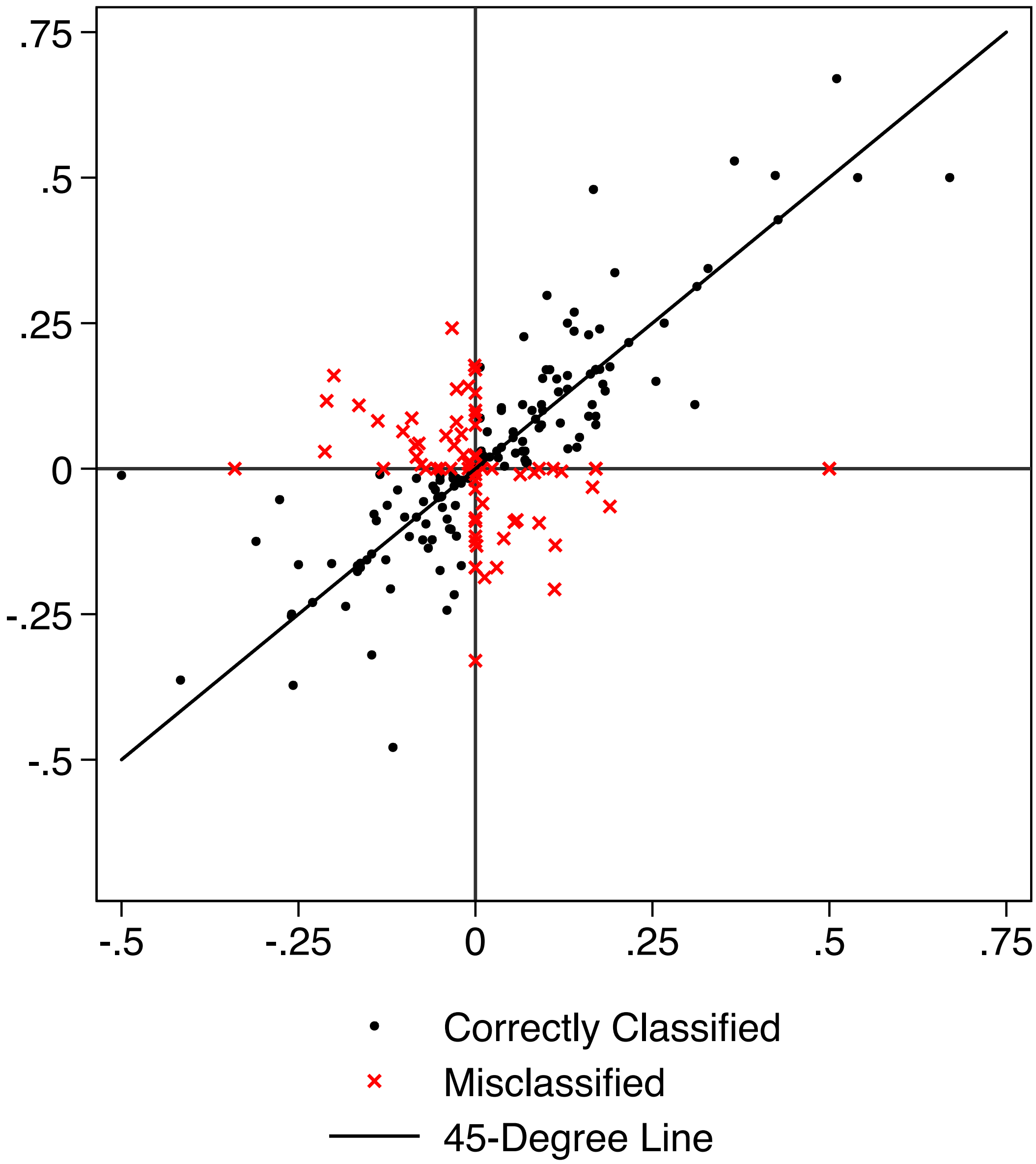

Figure 4: Misclassification of Overconfidence: This figure summarizes misclassification. The 45-degree line illustrates where shifts in the Performance Treatment are equal to shifts in the Combined Treatment. For observations above (below) the 45-degree line, shifts in the Combined Treatment are larger (smaller) than shifts in the Performance Treatment. Note that we can reject the null hypothesis that shifts in both settings are equal (equivalent to the 45-degree line representing the data) at the 1% level. Observations lying in different quadrants reflect correct or incorrect classification. In the upper-right (lower-left) quadrant, overconfident (under-confident) individuals are correctly classified. Observations in the upper-left quadrant: under-confident individuals are misclassified as overconfident. Observations in the lower-right quadrant: overconfident individuals are misclassified as under-confident.