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

Stephanie A. Heger† Nicholas W. Papageorge‡

Abstract: Wishful thinking, defined as the tendency to over-estimate the probability of high-payoff outcomes, is a widely-documented phenomenon that can affect decision-making across numerous domains, including finance, management, and entrepreneurship. We design an experiment to distinguish and test the relationship between two easily-confounded biases, optimism and overconfidence, both of which can contribute to wishful thinking. We find that optimism and overconfidence are positively correlated at the individual level and that both help to explain wishful thinking. These findings suggest that ignoring optimism results in an upwardly biased estimate of the role of overconfidence in explaining wishful thinking. To illustrate this point, we show that 30% of our observations are misclassified as under- or over-confident if optimism is omitted from the analysis. Our findings have potential implications for the design of information interventions since how agents incorporate information depends on whether the bias is ego-related.

Keywords: Subjective beliefs, overconfidence, optimism, information, experiments.
JEL Classification: C91 D03 D8 L26.

*Current version: June 26, 2018. This project benefitted from many helpful comments and conversations. We particularly thank Barton Hamilton and David Levine for their continual support. Additionally, we thank: Thomas Åstebro, Sascha Baghestanian, Jorge Balat, Robert Barbera, Hülya Eraslan, Andrew Gray, Glenn Harrison, PJ Healy, Dan Levin, Filippo Massari, Glenn MacDonald, Patrick McAlvanah, 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] School of Economics, University of Sydney. Email: `saheger@gmail.com`.
‡Department of Economics, Johns Hopkins University, IZA and NBER. Email: `papageorge@jhu.edu`.

1 Introduction

A robust empirical pattern, documented in research from both economics and psychology, is that individuals tend to engage in "wishful thinking," defined as over-estimating the probability of high-payoff outcomes (De Bondt and Thaler, 1995). Wishful thinking may lead to sub-optimal decisions in labor markets (Larkin and Leider, 2012), insurance markets (Sandroni and Squintani, 2013), financial markets (Barber and Odean, 2001; Biais et al., 2005), management (Malmendier and Tate, 2005, 2015) and entrepreneurship (Camerer and Lovallo, 1999; Koellinger, Minniti, and Schade, 2007; Åstebro et al., 2014). In each of these examples, there are two distinct biases that could drive wishful thinking. The first, *overconfidence*, is ego-related and is defined as an individual's over-estimation of their own performance (Moore and Healy, 2008).[1] The second, *optimism* is not performance or ego-related, but instead embodies a tendency to over-estimate higher-payoff or preferred outcomes (Irwin, 1953; Weinstein, 1980). Thus, neither bias implies the other, but both could drive wishful thinking in contexts where a high-payoff event is performance-dependent. Previous research has generally studied overconfidence or optimism as isolated phenomena. Doing so ignores how both biases could jointly contribute to economic behavior via wishful thinking, potentially conflating the two, and also overlooks the possibility that the biases are correlated.

In this paper, we present results from an experiment designed to study optimism and overconfidence as two distinct biases, both of which can drive wishful thinking. We present three main findings, which highlight possible shortcomings associated with studying optimism and overconfidence in isolation. First, we show that optimism and overconfidence are positively correlated at the individual level.[2] Second, we show that optimism and overconfidence jointly explain wishful thinking in settings where (i) individuals must assess their own performance and (ii) their performance affects their payoffs. Together, these results raise concerns about measurement and inference. For example, limiting attention to overconfidence in settings where individuals are optimistic can lead to an omitted variables bias. Our

[1]Regarding terminology, Moore and Healy (2008) identify three distinct phenomena that have been called overconfidence in previous literature, all of which are related to an individual's assessment of their own performance: (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 concerning performance in general. In this study, our definition of overconfidence corresponds to the first definition in Moore and Healy (2008).

[2]Previous experimental literature has cleanly identified overconfidence (e.g., Burks et al. (2013)) or optimism (e.g., Weinstein (1980); Ito (1990); Mayraz (2017); Coutts (2015); Barron (2016)). Closer to our study, Tasoff and Letzler (2014) make a distinction between optimism and overconfidence and Åstebro and Gutierrez (2017) discuss the importance of disentangling confounded biases to understand behavior, in their case entry into self employment, while Åstebro, Jeffrey, and Adomdza (2007) report evidence of correlated cognitive biases in the context of entrepreneurship.

third result illustrates this point, demonstrating that ignoring the role of optimism in such settings results in the misclassification of overconfidence for nearly one-third of our observations. In other words, ignoring optimism can lead to a misdiagnosis of the true source of wishful thinking. This is a potentially widespread issue since many findings in previous work relating overconfidence to behavior are conducted in settings where optimism might arise, but is ignored, which relies on the implicit assumption that overconfidence is independent of optimism (Camerer and Lovallo, 1999; Hoelzl and Rustichini, 2005; Blavatskyy, 2009).

We distinguish between overconfidence and optimism on the basis that overconfidence is ego-related (e.g., related to beliefs about one's own performance) whereas optimism is not. This distinction is important for two reasons. First, there is mounting evidence that belief updating significantly differs depending on whether the prior concerns a belief that is ego-related (Ertac, 2011; Mobius et al., 2014; Charness, Rustichini, and van de Ven, 2014). For example, Charness, Rustichini, and van de Ven (2014) and Ertac (2011) find evidence that individuals abandon Bayes' rule when incorporating ego-related information (even though they update using Bayes' rule when incorporating information that is not related to the self).[3] These differences in updating matter for the design of information interventions (i.e., information campaigns aimed at helping individuals make better decisions), which may depend crucially on whether the information being provided is aimed at improving decisions in an ego-related context targeting overconfidence versus in a context where performance plays no role and the intervention targets optimism.[4] Second, the distinction matters for incorporating subjective expectations into structural models of decision-making (de Bruin, Van der Klaauw, and Topa, 2011; Van der Klaauw, 2012; Delavande, 2008; Delavande and Zafar, 2014). These models are often used to simulate behavior in counterfactual environments that potentially vary by the importance of individual performance. Misdiagnosing the source of wishful thinking could potentially lead to inaccurate predictions of counterfactual beliefs and thus behavior.

Indeed, the importance of ego is found throughout the literature in models designed to rationalize overconfidence, but not in models of optimism. Conceptually, optimism may be motivated by host of different models, including anticipatory utility (Brunnermeier and Parker, 2005; Caplin and Leahy, 2001), affective decision-making (Bracha and Brown, 2012), rank dependent utility (Quiggin, 1982), preferences consistent with subjective expected util-

[3]Similarly, Eil and Rao (2011) find differences in updating when subjects receive good news as opposed to bad news. Hoffman (2016) also finds evidence of a negative relationship between overconfidence and a willingness-to-pay for information in a framed field experiment.

[4]There is a burgeoning literature on information interventions, designed to mitigate bias in beliefs by providing information. Information interventions have been studied in a variety of scenarios, including education (Jensen, 2010; Fryer, 2016; Wiswall and Zafar, 2015), health (Dupas, 2011) and entrepreneurship (Fairlie, Karlan, and Zinman, 2015).

ity but with differing priors (Van den Steen, 2004) or decision-makers with differing technologies (Santos-Pinto and Sobel, 2005). Overconfidence, however, is typically characterized using belief-based preferences that permit ego-utility or self-image concerns (Bénabou and Tirole, 2002; Köszegi, 2006) or models of strategic over-estimation (Charness, Rustichini, and van de Ven, 2014; Burks et al., 2013). Experimentally, Ewers and Zimmermann (2015) show evidence that over-stating performance may be a way to seek approval, and Schwardmann and van der Weele (2016) find that upwardly biased beliefs about one's own performance or aptitudes can be strategically beneficial if shared with opponents. Our study is not designed to distinguish among these models, but to build on their implicit insight that overconfidence permits a role for individual performance while optimism does not. This fundamental difference underscores the importance of distinguishing between the two forms of bias. For example, rationalizing beliefs by appealing to the strategic benefits of over-estimating own performance is likely to be an inappropriate model of beliefs formation if wishful thinking is due to optimism and is thus independent of performance.

To clarify how and why we distinguish optimism and overconfidence, we present a simple and concrete example. An aspiring restauranteur is tasked with computing expected returns from opening a restaurant. Earnings are a function of the individual's performance as a chef P and local demand for the cuisine he will cook D. Assume that expected earnings are $E = P + D$ and that opening the restaurant is a utility maximizing choice if $E \geq \lambda$, where we assume that $\lambda > 0$ captures earnings from alternative sources (e.g., paid employment) along with fixed costs associated with opening a new restaurant net of the non-pecuniary benefits of running a restaurant. Suppose beliefs are given by

$$\tilde{E} = P + \epsilon^P + D + \epsilon^D$$

where ϵ^P represents the bias in beliefs about performance and ϵ^D is bias about demand. In this setting, $\epsilon^P > 0$ could arise if the individual over-estimates performance — that is, if the individual is overconfident. However, it might also arise from optimism since a preferred outcome (higher earnings) would be generated by better performance, which is the case in many economic contexts. In contrast, $\epsilon^D > 0$ directly implies optimism since it represents an upwardly biased belief that is payoff-favorable, but independent of performance. The crucial difference is that, while optimism influences beliefs concerning the probability of a payoff, overconfidence affects beliefs when individuals assess their own performance. In this example, sub-optimal decision-making occurs in this model if $\tilde{E} \geq \lambda > E$, which means the individual is upwardly biased enough to open a restaurant when doing so is not the utility maximizing choice.

The example highlights a few points. First, either optimism or overconfidence (or some combination of the two) could promote sub-optimal decision-making, e.g., induce the aspiring restauranteur to open his restaurant when, if unbiased, he would prefer not to. To see this, note that an individual opens a restaurant if $\epsilon^P + \epsilon^D > \lambda - P - D$, and both optimism and overconfidence increase the value of the left-hand side of this inequality. Second, a positive correlation between optimism and overconfidence — a novel finding of this study — would suggest problems with inference. In the example, this would mean there is an increased positive correlation between ϵ^P and ϵ^D. The ϵ are necessarily positively correlated to some degree since performance generates a higher payoff, in which case optimism increases both; correlation between optimism and overconfidence increases the degree of this correlation. In the experiment we run, we construct treatments where only optimism or overconfidence is possible and then assess whether these biases are correlated.[5]

To understand the consequences of a correlation between optimism and overconfidence, suppose we measure beliefs that an aspiring, wishful thinking restauranteur holds about the likelihood of success opening a restaurant. If his wishful thinking is driven by optimism, but we mistakenly attribute it to overconfidence, then we would inaccurately conclude that his bias should be reduced in a setting where his own performance may play less of a role, such as deciding whether to play the lottery. In this case, an optimistic individual would over-estimate the probability of winning, but individuals who are merely overconfident will not perceive themselves as contributing to the outcome, and thus will not engage in wishful thinking. Whether a pattern of behavior will carry over to this situation from a context like that of the restauranteur depends on what generates wishful thinking in the latter context, and deciphering this is a step which has generally been skipped in the literature.

The correlation between the two biases has further implications for aggregate biases and behavior. Correlated biases mean that optimistic individuals are more likely to be overconfident, both of which could lead such individuals' beliefs to cross the threshold (captured by λ) above which they believe opening a restaurant is a good decision. Higher bias is particularly relevant when the barriers to entry are sufficiently high as it would induce aspiring restauranteurs whose true payoff will be negative to "take the plunge" even when less extreme wishful thinking might not be enough. In contrast, uncorrelated biases would mean that there are more individuals in the population who are either optimistic or overconfident, but not both, so their level of wishful thinking may not suffice to induce them to open the restaurant. By similar reasoning, if λ is small, over-entry would occur more often when more

[5]It is important to distinguish correlation in the ϵ from the possible correlation between P and D. We make no assumptions on the joint distribution of P and D. However, we note that one way to interpret them is that P is "skill" (how well the aspiring restauranteur prepares food or runs a business) and the other is "luck" (demand for cuisine the aspiring entrepreneur plans to offer).

individuals are biased but the size of that bias is small, which would be the case if the biases are uncorrelated.

Our experimental design is outlined in detail in Section 2. Here, we explain a simplified version. Subjects 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 (we call this Distribution 1), paying the subject for the accuracy of his belief. This is the "Baseline Treatment", where the elicited beliefs are referred to as the "Baseline beliefs" and are denoted $z_{d1,baseline}$. We repeat this same exercise with a second urn that contains 2 white balls and 1 black ball and where 1 ball is drawn (we call this Distribution 2).

To identify optimism, we vary whether there is a payoff-favorable outcome (the "Payment Treatment"), while still incentivizing 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 $z_{d1,payment}$. 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, $z_{payment} > (<) z_{baseline}$. Our definition of optimism is consistent with previous work that finds consistent evidence that subjects overweight the probability of payoff-favorable outcomes that are independent of performance (Irwin, 1953; Weinstein, 1980).

To identify overconfidence, we incorporate the individual's beliefs about their own performance (the "Performance Treatment"). In this treatment, optimism should not play a role because individuals are incentivized to correctly guess the number of white balls, but white balls are not payoff-favorable. 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, equivalent to Distribution 2) and does not add any ball otherwise (equivalent to Distribution 1). 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 the probability that the single draw results in white. We denote this belief $z_{performance}$. If the subject answers correctly, he faces Distribution 2 and we compare his $z_{payment}$ belief to $z_{d2,baseline}$ (which is the belief measured in the Baseline Treatment for Distribution 2). If he answers the IQ question incorrectly we compare $z_{payment}$ to $z_{d1,baseline}$. We say that the subject is overconfident (under-confident) in his performance if $z_{performance} > (<) z_{d*,baseline}$, where $d* = d1$ if answer is incorrect and $d* = d2$ if IQ answer is correct.

We are interested in wishful thinking in economic environments, such as evaluating returns to opening a restaurant, where an individual's performance affects the likelihood that a preferred outcome (in this case, high income generated by the restaurant) is realized. Here, decisions could 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 estimate how each individual's wishful thinking measured in the Combined Treatment is related to their optimism and overconfidence measured in the Payment and Performance Treatments.

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, including a series of sensitivity analyses in Section 4.3. Section 5 concludes.

2 Experimental Design

The purpose of our experiment is to separately measure optimism and overconfidence and to relate these measures to beliefs in a setting where both biases can occur, i.e., a setting where subject performance can influence the probability of preferred outcomes. To that end, the experiment follows a within-subject 2×2 ($Payment \times Performance$) design. This means that each subject faces four treatments, which arise from experimentally varying (1) whether some outcomes are preferred (through a side payment); and (2) whether performance affects outcomes. Figure 1 provides a graphical illustration of this design. According to the figure, the four treatments are the Baseline Treatment (no Payment and no Performance), the Payment Treatment, the Performance Treatment and the Combined (Payment + Performance) Treatment. Each treatment is described in greater detail below. From the perspective of experimental subjects, each of the four treatments were presented as separate Tasks. Subjects completed all of the questions in a single Task before moving on to the next Task. Finally, at the end of the session, we randomly chose one Task for payment.[6]

The experiment was conducted at Washington University in Saint Louis in the MISSEL

[6]We ran sessions where treatments are completed in 5 different orders, which allows us to control for order and test for robustness of our results to different Task orderings within the experiment. Within each of the five treatment orders, the order of the distributions was randomized at the subject level. In Section 4.3, after presenting our main results, we summarize a series of robustness checks, including robustness of main results to order effects. More detailed presentations of robustness checks are found in Appendix A.1.

laboratory. Subjects were recruited via ORSEE (Greiner, 2015) and the experiment was conducted in Z-Tree (Fischbacher, 2007). In total, 125 subjects (59 male and 66 female) participated across 15 sessions. On average, sessions lasted approximately 90 minutes and subjects earned $25 USD. In the appendix, 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.

2.1 Treatments

We begin by describing the Baseline Treatment and will then describe each of the 3 additional treatments, highlighting how they differ from the Baseline. For quick reference, Table 1 summarizes the characteristics of the four treatments.

Baseline Treatment In the Baseline Treatment (depicted as the upper-left box in Figure 1), subjects face a series of six distributions presented in a random order and are asked to report their belief about the probability that white balls will be drawn from an urn that contains white and black balls. Subjects know the number of white and black balls in the urn, as well as the number of draws from the urn. Importantly, we elicit the subject's beliefs about the entire distribution. The six distributions, listed in Table 2, can be divided into two classes: single-draw (where one ball is drawn) and three-draw (where three balls are drawn). For single-draw distributions, subjects report beliefs about the likelihood that the white ball is drawn. In three-draw distributions, subjects report beliefs that at least one, at least two and at least three balls are white.

To incentivize accurate reporting of probabilistic beliefs, we pay subjects according to the quadratic scoring rule (QSR) (Brier, 1950; Murphy and Winkler, 1970). [7]

Subjects face each distribution (i.e., urn) one at a time in a random order, and they receive no feedback on the accuracy of their reports or on the realization of the draw. When subjects complete one treatment, they then move on to complete the other treatments.

[7] 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).

Payment Treatment In the Payment Treatment, in addition to being paid for the accuracy of beliefs as in the Baseline Treatment, subjects receive a side payment that increases in expected value as more white balls are drawn from the urn. This treatment is depicted as the upper-right box in Figure 1 and the arrow from left to right illustrates that the Payment Treatment differs from the Baseline Treatment because individuals are experimentally induced to prefer white balls through the addition of the side payment, the value of which increases in the number of white balls drawn. Importantly, this side payment is independent of the payment for belief accuracy and pays 10 USD with a probability equal to $\frac{\text{total white drawn}}{\text{max white possible}}$ and 0 otherwise. The denominator "max white possible" is the maximum number of white balls that can be drawn in each experimental treatment and is discussed below.

Performance Treatment In the Performance Treatment, subjects' performance on a multiple choice IQ task determines which distribution they face. The treatment is depicted as the lower-left box in Figure 1 and the vertical arrow illustrates the additional role for performance relative to the Baseline Treatment.[8] Moreover, the way in which white balls are added links the six distributions and is depicted in Table 2. Subjects may start facing Distribution 1 (a single-draw distribution) and are then asked a single IQ question. The subject is 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., the subject does not know whether the answer is correct), the subject reports his belief about the likelihood that the one draw from the jar is a white ball. In the three-draw distributions the subject is asked to perform the exact same exercise. The only difference is that the subject starts in Distribution 3 (the starting distribution in the three-draw class) and answers three IQ questions. A white ball is added to the jar for each correctly answered IQ question, resulting in a final distribution that corresponds to either Distribution 3, 4, 5, or 6. Without feedback, the subject is asked about the likelihood that 0, 1, 2, or 3 draws from the jar consist of white balls. We use one-draw and three-draw distributions to assess whether beliefs' patterns shift when subjects face a more difficult computation.

We pay \$2 for each correctly answered question so that subjects do not have an incentive to purposely give an incorrect answer to a question to increase his certainty about the

[8] 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). This task was also used in Grossman and Owens (2012) and in Owens, Grossman, and Fackler (2014). 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).

distribution he faces. We acknowledge that this introduces the possibility for optimism in the Performance treatment. We attempted to balance this concern with the concern that subjects were properly incentivized to earnestly answer the IQ questions. A potential robustness check, which we do not conduct, would be to offer no incentives for the IQ questions.

Combined Treatment In the Combined Treatment, we simultaneously apply the Payment Treatment and the Performance Treatment. This treatment is depicted as the lower-right box in Figure 1 and the diagonal arrow illustrates that relative to the Baseline Treatment, the Combined Treatment adds a side payment along with a role for performance. In particular, 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), and 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.[9]

2.2 Key Features of the Experimental Design

In this section, we elaborate on two key features of our experiment: (1) the elicitation of the Baseline belief and (2) the single unified task used across treatments. We elicit a set of baseline beliefs for each subject in each of the 6 distributions. The Baseline belief then serves as an individual-level control to which we compare the individual's beliefs from the other three treatments. Returning to Figure 1, making within-subject comparisons of beliefs elicited in the Baseline Treatment to those from other treatments is illustrated by the three arrows originating in the upper-left box. Unlike using the objective distribution as a comparison, using the Baseline belief as an individual-level control means that any factors that affect reported beliefs uniformly across treatments, including poor mathematical skills, 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, and we are able to net out its impact on reported beliefs. This point relates to concerns about risk aversion. One

[9]If the Payment or Combined Treatments were chosen for payment to the subject, then the associated side payments were also paid to subjects.

concern about the side payment in the Payment and Combined Treatments is that risk-averse individuals might hedge their probabilistic beliefs, implying subjects would under-report the probability of a white ball being drawn. This could lead us to under-estimate the degree of optimism in our subject pool. However, in Appendix A.2, we show formally how netting out Baseline beliefs helps to ease concerns about risk aversion distorting reported beliefs.

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 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. Indeed, in a recent contribution, Coutts (2017) shows evidence that studying different biases by making comparisons across contexts (in his case, to identify asymmetries in how individuals update their beliefs) can potentially generate misleading results. We avoid this issue by studying different biases using a common task across treatments.

3 Data Description

In this section, we describe the data generated by the experiment and explain how we construct measures of optimism and overconfidence.

3.1 Data Generated by the Experiment

Across 15 sessions, we collected belief data from 125 subjects in four treatments. Panel A of Table 3 shows the number of observations per subject at the distribution-treatment level. Our data consist of 20 observations the subject-distribution-treatment level, totaling 2,500 observations at the distribution-treatment level (125×20).

Given our design, it is possible that we see the same subject in the same treatment and facing the same distribution more than once. For example, suppose the individual answers 0 IQ questions correctly both times he faces Distribution 1 in the Performance Treatment. Then he reports Performance Treatment beliefs for Distribution 1 twice. In such cases, we average over the individual's responses, which means we drop 204 observations, leaving us with 2,296. It is important to note that these are not inconsistent responses because subjects are answering different IQ questions and hence have different beliefs about the correctness of their answers. Additionally, we drop 52 distribution-treatment observations in which the reported beliefs violate the rules of probability, leaving 2,244 observations.[10]

Panel B of Table 3 shows how many of these observations correspond to each distribution-treatment dyad. Recall that starting distributions are not the same as the final distributions subjects actually face in the Performance and Combined Treatments, where performance on the IQ task determines the final distribution.

3.2 Constructing Measures of Optimism and Overconfidence

We define optimism as the belief that a desirable outcome is more likely to occur. This is consistent with existing definitions of optimism or wishful thinking (Weinstein, 1980; Ito, 1990; Tasoff and Letzler, 2014; Barron, 2016). Optimism is distinct from overconfidence (or performance overestimation) in that overconfidence is "ego-related" and is the belief that one's performance is better than it actually is. Our definition of the term is consistent with research using the term to describe over-estimation of own performance (Moore and Healy, 2008). To construct our measures of optimism and overconfidence, we use the Baseline belief as a subject-level control. To do this, our measure of optimism (overconfidence) is the average difference between the Baseline Treatment belief and the Payment (Performance) Treatment belief at the subject-distribution level.

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

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

where M_d is the number of draws for distribution d, $z_{i,d,m,\tau}$ are beliefs reported by individual i facing draw m of distribution d under treatment τ, where

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

The variable $\overline{shift}$ is defined at the subject-distribution-treatment level, and is the average difference between beliefs reported under treatment τ and beliefs reported under the Baseline Treatment ($\tau = Baseline$). $\overline{shift}_{i,d,Pay}$ captures optimism by measuring how a subject's belief changes due to the presence of a side payment for white balls. For example, suppose a subject's baseline beliefs for distribution 2 are given by $z_{i,2,1,base} = 0.3$ and his belief in the payment treatment is given by $z_{i,2,1,pay} = 0.4$. That is, he reports that the probability

[10] 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.

of 1 white ball being drawn from an urn composed 2 white balls, 1 black ball with 1 drawn is 0.3 in the baseline and 0.4 in the payment treatment, implying that $\overline{shift}_{i,d,Pay} = 0.1$. $\overline{shift}_{i,d,Perf}$ captures overconfidence by measuring how much a subject's belief changes when his performance affects the distribution of white balls. Finally, $\overline{shift}_{i,d,Combined}$ captures changes in beliefs when a subject can affect the distribution through his own performance and also receive side payments for each white ball that is drawn. We refer to each of these "shifts" as "Optimistic Shifts", "Overconfident Shifts"and "Combined Treatment Shifts", respectively.[11] Our data consist of 738 Optimistic Shifts, 385 Overconfident Shifts and 395 Combined Treatment Shifts, for a total of 1,518 shifts.

It is important to note that $\overline{shift}_{i,d,\tau}$ is only one possible way to measure optimism and overconfidence. Our current measure of overconfidence is identified from individuals who answer an IQ question incorrectly, but believe there is some probability of having answered it correctly. On the other hand, if a subject answers the question correctly and believes there is some probability of having answered it incorrectly, then he cannot be considered overconfident. We take two steps to ensure our results are robust to this strict definition. First, we choose questions that are sufficiently difficult that overconfidence can be identified using a larger number of individuals and observations.[12] Second, in our regression analysis we control for the individual's total number of correctly answered IQ questions to avoid conflating overconfidence with overall poor performance. We discuss alternative approaches and robustness checks in Section 4.3, which are presented in greater detail throughout Appendix A.1.[13]

Finally, we note that, in our main analysis in Section 4, we relate Optimistic, Overconfident and Combined Treatment Shifts at the subject-distribution level. Doing so places additional burden on the data because it requires that an individual be observed in the same distribution for each treatment, which does not necessarily occur in the Performance and Combined Treatments since the subject's IQ performance determines the distribution faced. There are 383 observations in both the payment and the performance treatments. We use these individuals to identify how Optimistic and Overconfident Shifts relate at the

[11]In Appendix A.4, we provide formal definitions of optimism and overconfidence. We use these definitions to derive definitions for Optimistic Shifts and Overconfident Shifts.

[12]Grosswirth, Salny, and Stillson (1999) provide the percentage of readers who answer the question correctly. We chose questions such that the probability of a correct answer was between 25%-75%.

[13]For example, 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. One possible alternative would be to elicit the subject's belief about his performance, denoted $\hat{p}$, and then impute $z_{performance}$. However, doing so would mean we lose uniformity of the experimental task across treatments and would also raise the subjects' cognitive burden. In Appendix A.1, as part of a robustness test, we impute the subject-specific $\hat{p}$ in the single-draw distribution to construct an alternative measure of overconfidence. We find that our main qualitative results remain unchanged.

individual-distribution level. Finally, there are 247 observations where the same individual faces the same distribution in the Payment, Performance and Combined Treatments. These observations are used to identify how Optimistic and Overconfidence Shifts explain Combined Treatment shifts at the individual-distribution level. One potential concern is that our main results are therefore estimated on a selected sample, which may bias our results. We return to this point in Section 4.3 when discussing robustness.

4 Findings

4.1 Basic Patterns in the Data

In Figure 2, we plot histograms of $\overline{shift}$ for all 1,518 by treatment. The figure demonstrates that there is substantial heterogeneity in terms of optimism/pessimism and overconfidence/underconfidence.

To further examine average shifts from the baseline across treatments, we regress the variable $\overline{shift}$ on treatment dummies (payment, performance or combined), gender, the number of correctly answered IQ questions and distribution and order fixed effects. Results are reported in Appendix A.3 (see Table S9). Absent controls, we find that the Payment 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). This is consistent with other laboratory studies of optimism, which find small average treatment effects in pure optimism (Barron, 2016). We also show that once we include order and distribution fixed effects, men and women are equally likely to over-estimate the number of white balls. The inclusion of distribution fixed effects helps to alleviate concerns that empirical patterns are driven by three-draw versus single-draw distributions or the possibility that subjects skilled at answering trivia questions (correctly) believe that they have given the right answer, in which case uncertainty over which distribution they face is reduced. Related to the previous point, we also find the magnitudes of the shifts decrease in the number of correctly answered IQ questions. One possibly reason is that answering more IQ questions correctly is indicative of a stronger ability to calculate probabilities, which means that individuals might be less prone to optimism and overconfidence. Consistent with this interpretation, we find that answering more IQ questions correctly, which does not directly affect the distribution in the Payment Treatment as it does in the Performance Treatment, is associated with lower levels of optimism. These empirical patterns help to motivate why, in our main analyses, we include specifications that condition on each individual's total number of correctly answered trivia questions for the entire experiment.

4.2 Main Results

We now present are main results. Recall, the experiment is designed not only to measure optimism and overconfidence, but also assess whether and how they relate to each other and, moreover, how they drive beliefs in settings where performance affects the likelihood of preferred outcomes. The three main results presented here characterize these relationships.

Result 1. *Optimism and overconfidence are positively correlated at the individual level; that is, individuals who over-estimate their own performance are also more likely to over-estimate the likelihood of high-payoff outcomes in contexts in which performance does not play a role.*

Figure 3 shows the positive correlation between optimism and overconfidence by relating Optimistic Shifts and Overconfident Shifts. The interpretation is that optimistic individuals tend to be overconfident, while pessimistic individuals tend to be under-confident.

We also show that this correlation is robust to various controls. In Table 4, we regress Overconfident shifts on Optimistic shifts using the following equation:

$$\overline{shift}_{i,d,performance} = \overline{shift}_{i,d,payment}\phi_1 + X_{i,d}\beta_1 + \epsilon_{i,d,\tau} \tag{2}$$

Columns [1]-[3] use all 383 observations where an individual is observed in both the Performance and the Payment treatment for the same distribution, while columns [4] and [5] consider the sample of 247 observations where an individual is observed in the Performance, Payment and the Combined treatment for the same distribution (which is the sample used for our next set of results where optimism and overconfidence are used to explain beliefs in the Combined Treatment). Column [1] only adjusts for order fixed effects and the estimated coefficient is 0.66 and significant at the 1%-level. Column [2] adds a gender dummy and total number of correctly answered questions, while column [3] adds distribution fixed effects. These additions barely affect the estimates, lowering it to 0.64. Limiting attention to the smaller sample in column [4] likewise has no appreciable effect on estimates. Together, estimates in columns [1]-[4] suggest that a 10 percentage point increase in optimism is associated with a 6.2-6.6 percentage point increase in overconfidence. Column [5] includes a second-order polynomial to capture the pattern seen in Figure 3: the coefficient on the squared term is significant and positive, suggesting that the positive correlation is stronger at higher levels of optimism and overconfidence.[14]

[14] Recall that Result 1 concerns optimism and overconfidence, measured as deviations of beliefs from Baseline levels in the Payment and Performance Treatments, respectively. In Appendix A.3, we show that beliefs in the Payment and Performance Treatments are positively correlated even before we net out Baseline beliefs.

Result 2. *Both optimism and overconfidence explain Combined Shifts. Thus, given Result 1, ignoring optimism results in an omitted variable bias.*

Next, we examine how overconfidence and optimism relate to Combined shifts, where both optimism and overconfidence may be present. Our sample consists of the 247 observations where an individual is observed in the Performance, Payment and Combined treatment for a given distribution. We regress Combined shifts onto overconfidence and optimism using the following equation.

$$\overline{shift}_{i,d,combined} = \overline{shift}_{i,d,performance}\phi_2 + \overline{shift}_{i,d,payment}\phi_3 + X_{i,d}\beta_2 + \eta_{i,d,\tau} \tag{3}$$

Estimates are reported in Table 5. Columns [1] and [2] restrict ϕ_2 and ϕ_3 to 0, respectively, while column [3] estimates the full model presented in equation 3. We do this to highlight the omitted variable bias that arises if optimism (overconfidence) is ignored when trying to explain beliefs in a setting where both may be present. Comparing columns [1] and [3], the coefficient on Overconfident Shift falls from 0.65 to 0.50 when we control for Optimistic Shifts. Similarly, comparing columns [2] and [3], the coefficient on Optimistic Shifts falls from 0.59 to 0.27 when we control for Overconfident shifts. Estimates remain similar if we adjust for distribution fixed effects, gender and total number of trivia questions correctly answered throughout the experiment in columns [4] and [5]. This finding is concerning as it means that inferring overconfidence in scenarios where individuals also have preferences over outcomes is susceptible to omitted variable bias due to the presence of optimism.

Result 3. *If observations are sorted into the classes "overconfident," "underconfident," and "not influenced by performance-related confidence," failing to account for optimism would result in the misclassification of 29% of observations. In addition, the magnitude of under- or overconfidence for the remaining 71% of observations is significantly mis-estimated.*

It is quite common in the literature to find beliefs akin to those measured in our Combined Treatment (where both optimism and overconfidence may be present) used to identify overconfidence. For example, in the case of Barber and Odean (2001) or Biais et al. (2005), investors may be optimistic about returns or overconfident in their ability to choose good investments; in the case of Malmendier and Tate (2005) or Malmendier and Tate (2015), CEOs may be optimistic about firm performance due to factors out of their control or overconfident about how well they can run a firm; and in (Camerer and Lovallo, 1999) or Koellinger, Minniti, and Schade (2007), it is possible that entrepreneurs, like the aspiring restauranteur in our example, over-estimate their performance as business owners or are optimistic about the returns to opening a business above and beyond their own performance.

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 Combined shifts are equal to overconfidence (i.e., Overconfident shifts) using an F-test of the joint hypothesis that $\phi_2 = 1$ and $\phi_3 = 0$ in equation (3). We conduct this test for each of the models presented in Table 5 and, in each case, we reject the null hypothesis at the 1% level. The F-statistic ranges between 10.85 and 16.96. Second, in Figure 4, we plot Combined shifts (y-axis) against overconfidence (x-axis). If overconfidence is equivalent to Combined shifts, then the data should fall 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).

To explore this point a bit further, we identify four types of misclassification in Figure 4: 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.

The remaining 71% of the data lie in the upper-right and lower-left quadrant, indicating that the direction of the biases are the same. However, even if we restrict our analysis to this subset we continue to reject the null hypothesis that the 45-degree line in Figure 4 fits the data (p-value<0.01). Thus, for this set of observations, we would correctly classify the individual as overconfident or underconfident, but the magnitude of the bias is still mis-estimated.

4.3 Robustness of Main Findings

In this section, we briefly outline a subset of the numerous robustness checks we conduct to show that our main results are not driven by our measures of overconfidence and optimism, our specification or the selected samples. A more detailed presentation of robustness is found in Appendix A.1.

Robustness to Order In Appendix A.1.1, we test for order effects. While there is some variation due to less precise estimates in some cases, our main findings are generally robust. For each order, optimism and overconfidence are positively correlated and both optimism and overconfidence explain beliefs in the Combined Treatment.

Robustness to Definition and Specification As mentioned earlier, our current measure of overconfidence is identified from individuals who answer an IQ question incorrectly, but believe there is some probability of having answered it correctly. When an individual is uncertain about how well they answered a particular IQ question, they may "fill-in" this uncertainty about their performance by either "rounding up" or "rounding down". In our definition, a subject who has a tendency to "round up" when faced with uncertainty is overconfident, while a subject who has a tendency to "round down" will be under-confident.[15] Nonetheless, to help mitigate concerns that overconfidence is identified solely off an individual's uncertainty about particular questions, in Appendix A.1.2, we show that our main results are also robust if we measure overconfidence as mis-calibration (Lichtenstein, Fischhoff, and Phillips, 1977). Within the context of our IQ task, Lichtenstein, Fischhoff, and Phillips (1977) define an individual as well-calibrated if their beliefs about the number of correctly IQ questions are on average correct, without requiring the individual to know exactly which of the IQ questions he answered correctly. Results show qualitatively similar relationships to results in Tables 4 and 5.

Another potential alternative to the variable $\overline{shift}$ to measure optimism and overconfidence is to use elicited beliefs to compute an expected value of the number of white balls. In Appendix A.1.3, we show that the variable $\overline{shift}$ is equivalent to computing changes in the expected number of white balls and then dividing by three for the three-draw distributions. In other words, $\overline{shift}$ can be interpreted as measuring individual-level shifts in the expected value, but suitably weighted to account for larger possible numbers of white balls in some distributions, which would otherwise artificially overweight their importance in our results.

[15]The experiment allows us to assess whether people who tend to "round up" when uncertain about their performance also tend to be optimistic, i.e., to "round up" when evaluating probabilities of events unrelated to their performance.

In Appendix A.1.4, we show that our results are robust when we estimate less restrictive versions of equations (2) and (3. In particular, we use the Baseline belief as an independent regressor, rather than subtracting it from the Payment, Performance and Combined beliefs (which effectively restricts the coefficient on Baseline beliefs to 1. To study the correlation between Optimism and Overconfidence, we regress the Payment belief on the Performance belief, Baseline belief and our vector of controls. To study how overconfidence and optimism relate to beliefs in the Combined Treatment, we regress beliefs in the Combined Treatment on Performance, Payment and Baseline beliefs, as well as our vector of controls. Results using these less restrictive specifications are qualitatively similar to our main results.

Robustness to Sample Selection The samples used to establish our main results rely on individuals who are observed in the same distribution across treatments. To assess robustness of our results we average the $\overline{shift}$ variable for *every* subject in the single-draw distributions and also in the three-draw distributions in Appendix A.1.5. Using these alternative shift variables, we replicate our main specifications from equations (2) and (3) and obtain qualitatively similar results.

Additional Exploratory Analyses Appendix A contains a number of additional exploratory analyses used to assess robustness and validity of our main results. For example, in one set of analyses, we plot the Optimistic Shift and Overconfidence Shift against the Baseline belief for each distribution. We find that the magnitude of the shift is independent of the magnitude of the Baseline belief, which allays concerns that our measures of optimism and overconfidence are driven by systematic errors in the Baseline beliefs. In general, these additional analyses provide support for our three main findings.

5 Conclusion

We investigate optimism and overconfidence as two distinct, but related, biases in beliefs. By offering a clear distinction between these biases and then studying them within the same experimental setting, we are able to examine how they relate to each other. We show that both biases can lead to wishful thinking and that they are positively correlated. Together, this suggests not only that they are easily confounded, but that measuring one while ignoring the other leads to incorrect inference.

As we discussed, both biases lead to wishful thinking and potentially have similar impacts on outcomes, such as excess entry into self employment. Thus, one could argue that distinguishing between them is unnecessary. There are two key reasons why the distinction is important, however. First, information interventions aimed at reducing wishful thinking may not be effective if they target the incorrect type of bias. Second, failing to distinguish between the two could lead to inaccurate predictions depending on the role of individual performance in the behavior or outcome being predicted. This is particularly important if the aim is to consider the impact of wishful thinking at the aggregate level since the role of performance varies across economic decision-making contexts.

In light of these concerns, we outline a number of directions for future research. One possibility is to augment the experiment we have run to include variation in the importance of performance. This would allow us to explicitly test whether optimism remains steady across contexts, while overconfidence varies. Moreover, it would allow us to more precisely quantify the consequences of mis-measurement due to the kind of contextual variation we are more likely to encounter in investigating economic behavior. For instance, suppose we observe behavior in a performance-light context, such as the work of a team on which one member's actions will have a small impact. What would be the consequences of assuming that an individual's measured positive bias about the outcome of this work is generated entirely by overconfidence and using that notion to project that individual's behavior in contexts where performance is important, like a complex individual task? How much more accurate would such projections be if the different sources of bias were correctly measured?

A second avenue for future research would be to look more closely at the correlation between the two types of bias and its impacts on aggregate behavior as barriers to entry are varied. *Ceteris paribus*, correlated biases imply that there are a few individuals with more extreme positive bias, since optimistic individuals are likely to also be overconfident. This may affect decision-making when entry costs are high. An optimistic or an overconfident entrepreneur might not be positively biased enough to open a restaurant, while an optimistic and overconfident entrepreneur may be biased enough to make the investment. As barriers to entry increase, failures may become more rare but also more spectacular, because those who enter mistakenly do so due to large, compounded biases in judgment.

Third, future research could assess how information interventions targeting optimism affect overconfidence (or vice versa). This would permit a test of whether targeting the wrong bias due to misclassification undermines information interventions. This type of research would also strengthen our understanding of both optimism and overconfidence by shedding light on why they are correlated, permitting a test of whether they are causally related or whether the correlation arises due to a common third factor.

References

Åstebro, Thomas and Cédric Gutierrez. 2017. "The Impact of Overconfidence on Market Entry." Mimeo, HEC Paris.

Åstebro, Thomas, Holger Herz, Ramana Nanda, and Roberto A Weber. 2014. "Seeking the Roots of Entrepreneurship: Insights from Behavioral Economics." *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.

Barber, Brad M and Terrance Odean. 2001. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics* :261–292.

Barron, Kai. 2016. "Belief Updating: Does the 'Good-news, Bad-news' Asymmetry Extend to Purely Financial Domains?" Working Paper.

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

Biais, Bruno, Denis Hilton, Karine Mazurier, and Sébastien Pouget. 2005. "Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market." *Review of economic studies* 72 (2):287–312.

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.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini. 2013. "Overconfidence and Social Signalling." *Review of Economic Studies* 80 (3):949–983.

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. 2014. "Self-confidence and Strategic Behavior." Working Paper.

Coutts, Alexander. 2015. "Testing Models of Belief Bias: An Experiment." Working Paper.

———. 2017. "Good News and Bad News are Still News: Experimental Evidence on Belief Updating." Mimeo, Universidade Nova de Lisboa.

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.

de Bruin, Wändi Bruine, Wilbert Van der Klaauw, and Giorgio Topa. 2011. "Expectations of Inflation: The Biasing Effect of Thoughts about Specific Prices." *Journal of Economic Psychology* 32 (5):834–845.

Delavande, Adeline. 2008. "Pill, Patch, or Shot? Subjective Expectations and Birth Control Choice." *International Economic Review* 49 (3):999–1042.

Delavande, Adeline and Basit Zafar. 2014. "University Choice: The Role of Expected Earnings, Non-Pecuniary Outcomes, and Financial Constraints." ASU working paper.

Dupas, Pascaline. 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 3 (1):1–34.

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.

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.

Ewers, Mara and Florian Zimmermann. 2015. "Image and Misreporting." *Journal of the European Economic Association* 13 (2):363–380.

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.

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

Fryer, Roland G. Jr. 2016. "Information, Non-Financial Incentives, and Student Achievement: Evidence from a Text Messaging Experiment." *Journal of Public Economics* 144:109–121.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1 (1):114–125.

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.

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

Hoffman, Mitchell. 2016. "How is Information Valued? Evidence from Framed Field Experiments." *Economic Journal* 126 (595):1884–1911.

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.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics* 125 (2):515–548.

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.

Larkin, Ian and Stephen Leider. 2012. "Incentive Schemes, Sorting, and Behavioral Biases of Employees: Experimental Evidence." *American Economic Journal: Microeconomics* :184–214.

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. 2005. "CEO Overconfidence and Corporate Investment." *Journal of Finance* 60 (6):2661–2700.

Malmendier, Ulrike and Geoffrey Tate. 2015. "Behavioral CEOs: The Role of Managerial Overconfidence." *Journal of Economic Perspectives* 29 (4):37–60.

Mayraz, Guy. 2017. "Wishful Thinking." Mimeo, University of Melbourne.

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

Mobius, M.M., M. Niederle, P. Niehaus, and T.S. Rosenblat. 2014. "Managing Self-Confidence." 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.

Owens, David, Zachary Grossman, and Ryan Fackler. 2014. "The Control Premium: A Preference for Payoff Autonomy." *American Economic Journal: Microeconomics* 6 (4):138–161.

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

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

Sandroni, Alvaro and Francesco Squintani. 2013. "Overconfidence and Asymmetric Information: The Case of Insurance." *Journal of Economic Behavior & Organization* 93:149–165.

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

Schwardmann, Peter and Joël van der Weele. 2016. "Deception and Self-Deception." Mimeo, University of Munich.

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.

Tasoff, Joshua and Robert Letzler. 2014. "Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion." *Journal of Economic Behavior & Organization* 107:107–122.

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

Van der Klaauw, Wilbert. 2012. "On the Use of Expectations Data in Estimating Structural Dynamic Choice Models." *Journal of Labor Economics* 30 (3):521–554.

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

Wiswall, Matthew and Basit Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *Review of Economic Studies* 82 (2):791–824.

Tables and Figures

Table 1: Summary of Experimental Treatments

Treatments	Payment for Belief Accuracy	Side Payment for White Balls Drawn	White Balls Added to Jar for Correct Answers
Baseline	Yes	No	No
Payment	Yes	Yes	No
Performance	Yes	No	Yes
Combined	Yes	Yes	Yes

This table summarizes the main features of each of the four experimental treatments.

Table 2: Summary of Distributions Subjects Face

	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	

This table summarizes the six distributions subjects face. Subjects face all six distributions in the Baseline and Payment Treatments. In the Performance and Combined Treatments, 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 3: Summary of Data Collected and Sample Sizes

Panel A: Data Collected From Each Subject					
Starting	Treatments				
Distribution	Baseline	Payment	Performance	Combined	
1	1	1	2	2	
2	1	1	0	0	
3	1	1	2	2	
4	1	1	0	0	
5	1	1	0	0	
6	1	1	0	0	
Σ	6	6	4	4	
Panel B: Number of Observations from Each Treatment-Distribution Dyad					
					Σ
1	125	125	86	88	424
2	125	125	92	102	444
3	110	117	22	24	273
4	119	122	40	62	343
5	123	125	94	74	416
6	124	124	51	45	344
Σ	726	738	385	395	2,244

Panel A: shows what information is gathered from each subject. There are six distributions and four treatments. Each subject is asked about a total of 20 distributions. In the Baseline Treatment, the subject is asked about all six distributions. In the Payment Treatment, the subject is asked about all six distributions. In the Performance Treatment and the Combined Treatment, the subject begins facing Distribution 1 twice and Distribution 3 twice. Recall that in the Performance Treatment and the Combined Treatment the actual distribution subjects face is endogenous to the number of IQ questions they answer correctly and so may differ from the starting distribution. Panel B: shows how many observations are collected for each distribution and treatment pair.

Table 4: Main Results: Correlation between Optimism and Overconfidence

	[1]	[2]	[3]	[4]	[5]
Optimistic Shift	0.66***	0.65***	0.64***	0.62***	0.51***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)
Optimistic Shift-Squared	.	.	.	.	0.62***
					(0.24)
Male	.	0.01	0.006	0.02*	0.02
		(0.01)	(0.01)	(0.01)	(0.01)
Correct IQ Answers	.	-0.009***	-0.003	-0.004	-0.002
		(0.003)	(0.004)	(0.004)	(0.004)
Constant	0.03***	0.11***	0.12***	0.13***	0.11***
	(0.007)	(0.03)	(0.03)	(0.04)	(0.04)
Observations	383	383	383	247	247
R^2	0.32	0.34	0.41	0.46	0.48
Independent Obs	125	125	125	121	121
Order Dummies	[Y]	[Y]	[Y]	[Y]	[Y]
Distribution Dummies	[N]	[N]	[Y]	[Y]	[Y]
Smaller Sample	[N]	[N]	[N]	[Y]	[Y]

This table presents estimates from OLS regressions where the outcome variable is Overconfident Shifts ($\overline{shift}_{i,d,performance}$) and the explanatory variable of interest is Optimistic Shifts ($\overline{shift}_{i,d,payment}$). *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

Table 5: Main Results: How Optimism and Overconfidence Affect Beliefs

	[1]	[2]	[3]	[4]	[5]
Overconfident Shift	0.65***	.	0.5***	0.56***	0.54***
	(0.1)		(0.09)	(0.08)	(0.09)
Optimistic Shift	.	0.59***	0.27***	0.25***	0.25***
		(0.11)	(0.09)	(0.09)	(0.09)
Male	.	.	.	0.02	0.02
				(0.02)	(0.02)
Correct IQ Answers	.	.	.	-0.002	-0.001
				(0.004)	(0.004)
Constant	0.05***	0.11***	0.06***	0.04	0.05
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)
Observations	247	247	247	247	247
R^2	0.51	0.41	0.54	0.52	0.53
Independent Obs	121	121	121	121	121
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 relates optimism and overconfidence to shifts in the Combined Treatment. We present estimates from OLS regressions where the outcome variable is Combined Treatment Shifts ($\overline{shift}_{i,d,combined}$) and the explanatory variables of interest are Overconfident Shifts ($\overline{shift}_{i,d,performance}$) and Optimistic Shifts ($\overline{shift}_{i,d,payment}$). The F-test statistic is for the joint hypothesis that the coefficient on Overconfident Shifts is equal to 1 and 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. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Standard errors are clustered by individual.

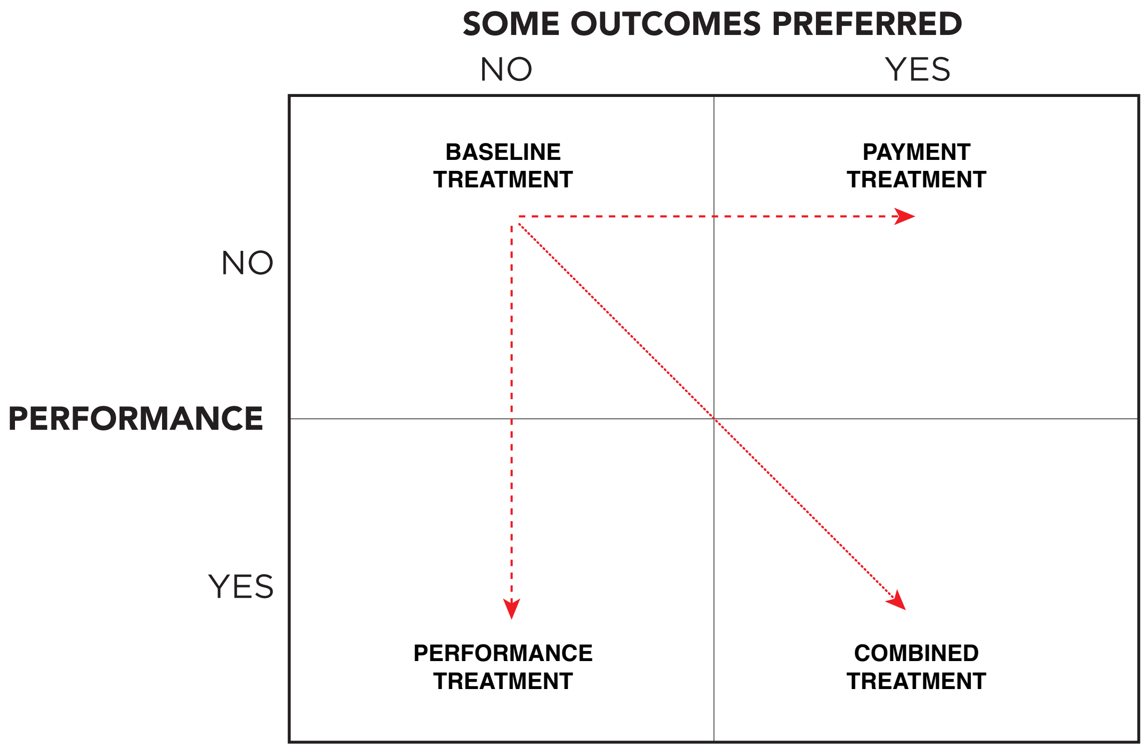
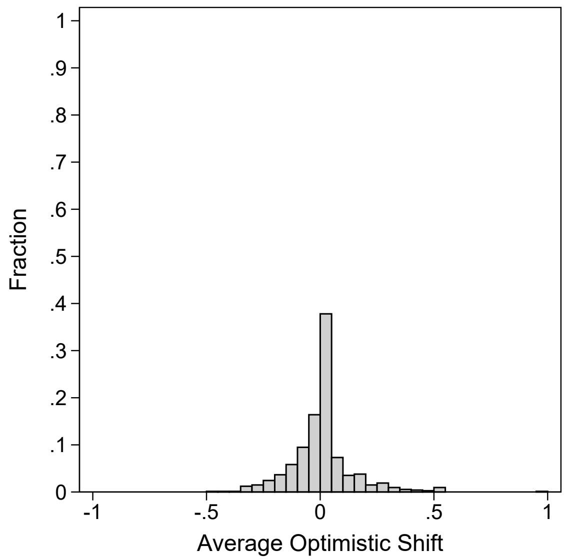
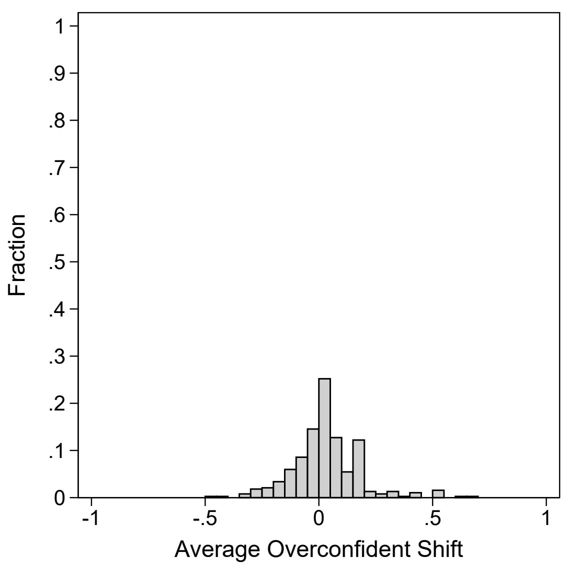


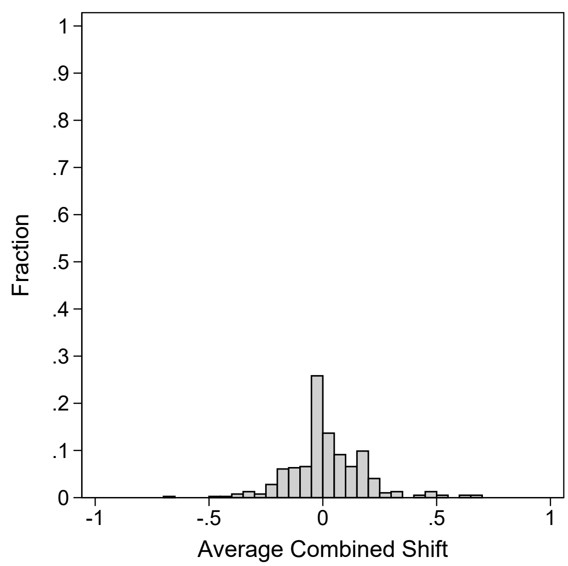
Figure 1: Experimental Design. This figure illustrates how the four experimental treatments are designed and how they relate to one another. In each treatment, subjects report beliefs about the probability of drawing white balls from urns containing different numbers of white and black balls. Treatments vary by (1) whether or not some outcomes (white balls) are preferred because they lead to a side payment; and (2) whether or not individual performance (answering trivia questions correctly) affects the distribution of black and white balls in the urn. In the Baseline Treatment (upper-left box) there are no side payments and performance plays not role. In the Payment Treatment (upper-right box), there is a side payment, but performance plays no role. The horizontal arrow illustrates this modification to the Baseline Treatment. In the Performance Treatment (lower-left box), trivia answers affect the distribution. The vertical arrow illustrates that how treatment departs from the Baseline Treatment. In the Combined Treatment (lower-right box), some outcomes are preferred and performance plays a role. The diagonal arrow illustrates these two modifications to the Baseline Treatment.



(a)



(b)



(c)

Figure 2: Average Treatment Effects. We plot histograms of the Optimistic Shifts (2(a)), Overconfident Shifts (2(b)) and Combined Shifts (2(c)). These shifts are our main outcome variable and are measured as the average deviation between the treatment belief (Payment, Performance, Combined) and the Baseline belief at the individual-distribution level.

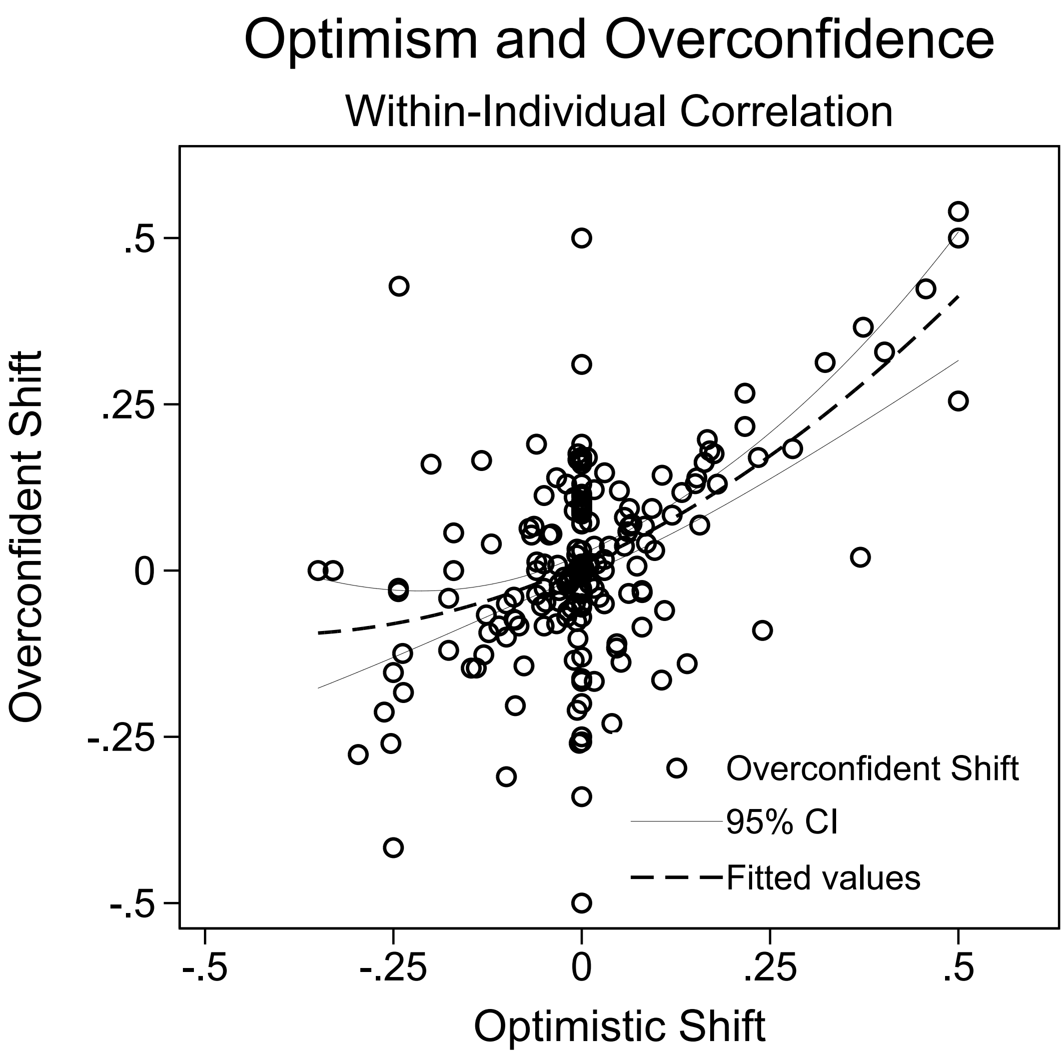


FIGURE 3: OPTIMISM AND OVERCONFIDENCE. This figure relates Optimistic Shifts ($\overline{shift}_{i,d,payment}$) and Overconfident Shifts ($\overline{shift}_{i,d,performance}$). Each point is an individual's deviation from Baseline Treatment beliefs in the Performance Treatment against deviations in the Payment Treatment. In addition, the plot includes a fitted second-order polynomial with confidence intervals at the 5%-level.

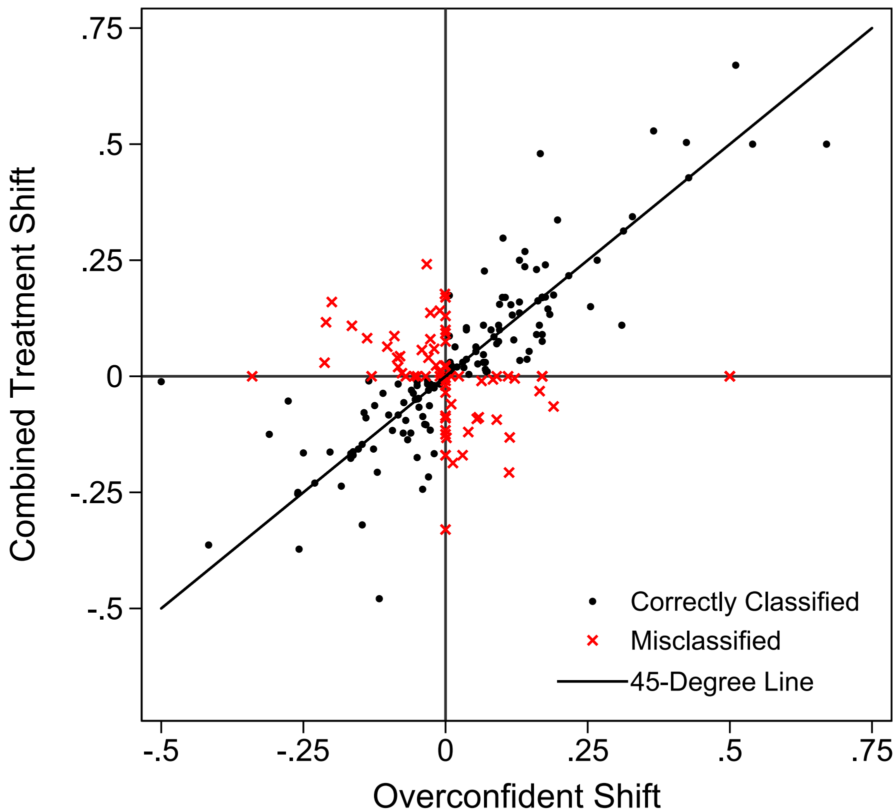


Figure 4: Misclassification of Overconfidence. This figure summarizes the misclassification that results from ignoring the role of optimism. The 45-degree line illustrates where Overconfident Shifts ($\overline{shift}_{i,d,performance}$) are equal to Combined Treatment Shifts ($\overline{shift}_{i,d,combined}$). 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. Further, observations in the upper-left quadrant represent under-confident individuals misclassified as overconfident, while observations in the lower-right quadrant represent overconfident individuals misclassified as under-confident.