Beliefs and actions: How a shift in confidence affects choices

Kai Barron and Christina Gravert

WZB, University of Gothenburg

12 February 2018

Beliefs and actions: How a shift in confidence affects choices

February 12, 2018

Abstract

Confidence is often seen as the key to success. Empirical evidence about whether such beliefs causally map into actions is, however, sparse. In this paper, we experimentally investigate the causal effect of an increase in confidence about one's own ability on two central choices made by workers in the labor market: choosing between jobs with different incentive schemes, and the subsequent choice of how much effort to exert within the job. An increase in confidence leads to an increase in self-selection into uncertain ability-contingent payment schemes. This is detrimental for low ability workers. Policy implications are discussed.

JEL Codes: C91, D03, M50, J24

Keywords: Overconfidence, experiment, beliefs, real-effort, grade inflation

1 Introduction

Confidence is often described as one of the most important ingredients for success. The beneficial effects of confidence on effort choices have been studied theoretically by Bénabou and Tirole (2002). They discuss how a higher level of self-confidence in one's abilities motivates people to work hard, overcome obstacles, and take beneficial risks. The connection between confidence and success is supported by evidence suggesting that higher confidence individuals are better at persuading others that they are of high ability (Burks et al., 2013; Schwardmann and Van der Weele, 2016), work harder (Puri and Robinson, 2007; Pikulina et al., 2017), and that overconfidence is evolutionary adaptive (Bernardo and Welch, 2001; Heifetz et al., 2007; Johnson and Fowler, 2011).

In contrast to the beneficial aspects of confidence, there is plenty of evidence documenting that overconfident individuals are more likely to make mistakes, such as taking unnecessary risks in stock trading (Odean, 1998; Barber and Odean, 2001), poor managerial decisions (Malmendier and Tate, 2005), or overentry into competition (Niederle and Vesterlund, 2007).

Despite the enormous literature on overconfidence, causal evidence on the implications of increasing an individual's confidence is sparse. Yet, without causal evidence of the effect of confidence on decisions, we can neither claim that increasing someone's confidence is beneficial, as suggested by the first set of papers nor that increasing it is detrimental, as suggested by the second set. Higher confidence could just be strongly correlated with unobserved characteristics that in turn generate positive or negative outcomes.

In this paper, we develop a simple theoretical framework and show empirically how an upwards shift in confidence could causally influence decision-making regarding payment scheme decisions, effort provision, and resulting earnings.

The following example illustrates how an exogenous increase in confidence can affect decision making in a labor market setting. Imagine a computer programmer fresh out of college (let's call her "Thandi"). Thandi is well trained and has the choice between a job at a mid-sized company that will pay her a fixed wage and the opportunity to work at a start-up with which she will earn far more if she is better than competing programmers and far less if she is worse.

Assume that Thandi, like most people, believes that she is better than average (see, amongst others Kruger (1999); Burson et al. (2006); Healy and Moore (2007); Moore and Healy (2008); Benoît et al. (2015)). This belief will be influential for two decisions: She needs to choose a type of job and then, conditional on being on the job, the effort she will exert. When it comes to the choice of job, she will be inclined to choose the start-up work to maximize her earnings due to her belief of being better than average (although risk aversion might temper this inclination (Bonin et al., 2007)). Choosing the ability-contingent payment scheme of the start-up is the correct choice for her if her ability level is *actually* above average, but harmful, if she *wrongly believes* she is better than average.

With regards to effort once on the job, however, her high self-confidence might motivate her to work harder in a start-up because her perceived returns from effort seem higher than in the corporate job. A high level of confidence will only be beneficial if she chooses the right job for her relative ability level. If she overconfidently chooses a job type she does not have the ability for, she faces the risk that her efforts do not pay off and she will work hard for a lower wage than what she would have earned at the corporate job.

Ideally, one would study these questions with actual job and effort choices. However, this approach poses several problems. The main challenge is that it is non-trivial to gain access to accurate measurements of the beliefs of job seekers about their abilities relative to their direct competitors, let alone identifying the influence of an exogenous shock to their beliefs. We circumvent these problems with a laboratory experiment. In the controlled environment of a laboratory study, we generate exogenous variation in beliefs to measure the causal effect of a shift in beliefs on (i) the selection into fixed or ability-contingent payment schemes, and (ii) effort exerted. We derive our hypotheses for the experiment from a simple theoretical framework.

In our experiment, a group of subjects takes a test measuring their relative cognitive ability. We then ask them to estimate their probability of being in the top half of their group, but provide no feedback. Next, they work on ten rounds of a real-effort task for which they can repeatedly choose one of two payment schemes. Subjects can either choose to work for an ability-contingent piece rate or a fixed piece rate regardless of their ability. The ability-contingent piece rate pays a high wage if the subject is in the top half of her group and nothing if she is in the bottom half. The fixed piece rate increases in each round, but always lies below the high piece rate of the ability-contingent piece rate. Thus, if a subject is certain of being in the top half of her group, choosing the ability-contingent piece rate maximizes her earnings.

The exogenous variation in beliefs about relative ability is generated in the experiment by exposing the subjects to either a harder or an easier version of the same test. Subjects randomly confronted with an easy test assess their position in a ranking of peers to be higher than subjects confronted with a harder test. [1] This is known as the hard-easy effect (Kruger, 1999; Moore and Kim, 2003; Moore and Small, 2007; Healy and Moore, 2007; Dargnies et al., 2016). Individuals fail to consider that the test is easier or harder for all participants, not just for themselves. This type

[1] Note, subjects are always compared only to others who have completed exactly the same test as they did. Therefore, subjects who face the hard test assess their relative position in relation to others who completed the hard test.

of overconfidence, believing that one is ranked higher than one actually is, is commonly referred to as "overplacement"[2]. We find strong evidence of the hard-easy effect for our subjects. Especially, the subjects in the bottom half of the group report on average higher beliefs in the easy treatment than in the hard treatment. The beliefs of those in the top half are less affected on average.

We find that higher average confidence about relative placement due to exposure to the easier test leads subjects to more often choose the ability-contingent piece rate. If randomly confronted with the hard test, subjects are more likely to choose the fixed piece rate. With regard to effort, we find that the motivation of all individuals is high, regardless of their beliefs and the incentives. The shift in beliefs has consequences for earnings of the bottom half of the group. The bottom group earns only about a quarter of what the top group earns on average, but their average earnings are reduced even further, by about 40 percent, when their confidence has been exogenously increased.

Overconfidence in relative ability is costly for below-average-ability individuals, as it increases their probability of choosing an ability-contingent incentive scheme. In a world in which ability and effort are perfect complements, a possible offsetting motivational effect of beliefs on effort is thus in vain since the payoff from each unit of effort is zero (holding ability constant). In our setting, high-ability individuals are almost never made underconfident enough for their beliefs to have a negative effect on their decisions.

When we split our results by gender, we find, unsurprisingly, that women in both conditions report lower average beliefs than men (see Bertrand (2011) and Niederle (2017) for an overview of the gender-gap in confidence). This difference is driven predominantly by men being more confident in their placement, on average, rather than by women underestimating their performance, although men and women are of equal average ability. We find no evidence of men and women reacting differently to the treatment. However, because of their higher beliefs leading to more mistakes in incentive scheme choice, in our setting men on average need to exert more effort to earn the same amount as women. This is due to a larger proportion of men earning nothing, despite their on average higher effort levels. A number of papers argue that differences in payment scheme choices are mainly driven by gender differences in risk preferences (see Eckel and Grossman (2008) and Croson and Gneezy (2009) for surveys). We do not find evidence of risk preferences explaining a significant part of the gender incentive-choice gap.

The paper additionally addresses one of the main challenges in the beliefs literature: understanding the causal mapping of beliefs to economic decisions. Many papers take as given that beliefs are causally related to decisions, and yet in the few papers that have investigated this rela-

[2] As opposed to "overprecision", overconfidence in the accuracy of one's beliefs, and "overestimation", the overconfidence in one's absolute ability.

tionship, the results are not so straight forward. Costa-Gomes and Weizsäcker (2008), for example, show that subjects in their games fail to best respond to their stated beliefs almost half of the time, providing evidence for the fact that beliefs do not directly map into actions. However, in a later follow-up paper, the authors use an exogenous variation in beliefs to show that beliefs can have a causal impact on choices (Costa-Gomes et al., 2014). A related paper by Smith (2013) also provides evidence of a causal effect of beliefs on actions, but to a smaller extent than one would estimate with an OLS regression. Both papers use an IV approach to reduce the endogeneity of beliefs.

The exogenous variation in beliefs through our experimental design distinguishes our paper from the ones looking at the relationship between beliefs about relative ability and choices (such as Bruhin et al. (2016); Murad et al. (2016); Cheung and Johnstone (2017); Pikulina et al. (2016, 2017)). For example, Bruhin et al. (2016) show that individuals are more risk taking when the probability of winning depends on their relative ability rather than on an exogenously imposed probability. However, overconfident individuals might be different to correctly calibrated individuals in other ways than just their beliefs. Further, our design rules out the possibility of substituting lack of ability with effort to keep earnings stable. In experiments where ability and effort are partial substitutes, such as that of Cheung and Johnstone (2017), less confident individuals could exert more effort to overcome their lack of ability.

Our paper provides evidence that a shift in beliefs leads to a shift in behavior and that this shift has meaningful consequences. The causal effect of beliefs on actions has implications for both career decisions as discussed in this paper, but is also meaningful for policy interventions in which beliefs are targeted in order to bring about a change in behavior.

2 Theoretical framework

Consider an individual i, who can earn money by performing a task that requires costly effort, e. She is either a high ability or low ability individual, $a \in \{a_L, a_H\}$. Prior to performing the task, the individual chooses between two incentive schemes: (i) one that pays a high wage to high-ability individuals, $w(a_H) = w_H$, and a low wage to to low-ability individuals, $w(a_L) = w_L$, or (ii) one that pays a fixed wage to everyone, $\bar{w}$, where $w_H > \bar{w} > w_L$. After choosing her incentive scheme, she chooses the level of effort, e, she would like to exert. A risk-neutral individual would choose her incentive scheme, $w \in \{\bar{w}, w(a)\}$, and effort level, e, by solving the following:

$$\max_{w\in\{\bar{w},w(a)\}} \max_{e\geq 0} \mathbf{E}_a[(s+w)\cdot e - c(e)] \tag{1}$$

where $c(e)$ is the cost of exerting effort, and is assumed to satisfy $c'(e) > 0$ and $c''(e) > 0$, the expectation operator, $\mathbf{E}_a$, denotes expectations in respect to the individual's ability, and s represents the individual's intrinsic motivation for completing the task. Following the approach by DellaVigna and Pope (2016), we view this intrinsic motivation term as including, in reduced form, any non-monetary reward the workers derive from working on the task. In terms of the laboratory experiment described below, this is taken to include any sense of duty to, or gratitude towards the experimenter for the fixed show-up fee.[3] Since, the individual's ability only affects her wage, we can rewrite equation 1:

$$\max_{w\in\{\bar{w},w(a)\}} \max_{e\geq 0} (s+\mathbf{E}_a[w])\cdot e - c(e) \tag{2}$$

It is clear from equation 2 that the individual's subjective belief regarding the likelihood that she is high-ability, $\hat{\pi} = \hat{\mathbf{P}}(a = a_H)$, is important both for her decision of which incentive scheme to take, and how much effort to exert if she chooses the ability-contingent incentives. Essentially, the choice of an incentive scheme involves a choice between being paid a certain piece rate of $\mathbf{E}_a[\bar{w}] = \bar{w}$, or an expected piece rate of $\mathbf{E}_a[w(a)] = \hat{\pi}\cdot w_H$ for each unit of effort. We normalise $w_H = 1$.

Conditional on incentive scheme, w, the individual chooses effort according to $e^* = c'^{-1}(s + \mathbf{E}_a[w])$. Under the certain piece rate (PR) incentive scheme she chooses $e^*_{PR} = c'^{-1}(s+\bar{w})$. Under the ability-contingent (AC) incentive scheme, she chooses $e^*_{AC} = c'^{-1}(s+\hat{\pi}\cdot w_H)$. If $\hat{\pi}\cdot w_H < \bar{w}$, then the individual exerts more effort under the certain piece rate incentives. However, if she is sufficiently confident in her own ability, such that $\hat{\pi}\cdot w_H > \bar{w}$ she expects a higher piece rate under ability-contingent incentives, and would work harder under these incentives. We define a threshold value of $\hat{\pi}$, namely $\pi^e := \frac{\bar{w}}{w_H}$, such that:

- If $\hat{\pi} \in [0, \pi^e]$, the individual exerts more effort under certain piece rate incentives than under ability-contingent incentives.
- If $\hat{\pi} \in [\pi^e, 1]$, the individual exerts more effort under ability-contingent incentives than under certain piece rate incentives.

For high levels of intrinsic motivation, s, differences in effort due to the incentive schemes will

[3] DellaVigna and Pope (2016) argue that this non-monetary reward term is important for explaining the commonly observed non-zero effort in fixed-wage laboratory experiments.

be harder to detect. This is a frequent challenge of laboratory real effort tasks (DellaVigna and Pope, 2016; Erkal et al., 2016; de Araujo et al., 2015). The length of the task constrains effort to be below a fixed effort level, $e \leq \bar{e}$, which can be binding, if $\bar{e} \leq e^*_{PR}$ and $\bar{e} \leq e^*_{AC}$. If this is the case, then the observed effort level chosen under both sets of incentives will be equal.

The individual chooses the ability-contingent incentives whenever she expects to earn more per unit of effort under them than she would under the certain piece rate per unit of effort.

$$(s + \hat{\pi} \cdot w_H) \cdot e^*_{AC} - c(e^*_{AC}) \geq (s + \bar{w}) \cdot e^*_{PR} - c(e^*_{PR}) \tag{3}$$

This inequality holds whenever $\hat{\pi} \cdot w_H \geq \bar{w}$.[4] It holds even if the effort level chosen under both incentives schemes is the same (i.e. if $e^* = \bar{e}$).

Under risk neutrality, the threshold for the choice of incentives, and the threshold for effort choices are equal (i.e. $\pi^i = \pi^e = \frac{\bar{w}}{w_H}$).

This equality leads to:

- if $\hat{\pi} \in [0, \pi^i]$, the *low confidence* individual will (i) choose the certain piece rate incentives, and (ii) exert (weakly) lower effort under ability-contingent incentives than under certain piece rate incentives.

- if $\hat{\pi} \in (\pi^i, 1]$, the *high confidence* individual will (i) choose the ability-contingent incentives, and (ii) exert (weakly) higher effort under ability-contingent incentives than under certain piece rate incentives.

In the appendix, we relax the risk neutrality assumption and show that the threshold belief at which individuals will switch incentive choice, differs from the one at which effort is affected by incentives.

2.1 Hypotheses

Our treatment has the objective of inducing an exogenous shift in the subjective belief of our subjects about their abilities. Instead of shifting an individual's belief $\hat{\pi}$, we will shift the distribution

[4] To see this, notice that if $\hat{\pi} \cdot w_H > \bar{w}$, the individual could simply choose the ability-contingent incentives and set effort equal to the optimal effort level under certain piece rate incentives, $e = e^*_{PR}$, and receive a higher expected payoff than under the certain piece rate incentives, i.e.

$$(s + \hat{\pi} \cdot w_H) \cdot e^*_{PR} - c(e^*_{PR}) \geq (s + \bar{w}) \cdot e^*_{PR} - c(e^*_{PR})$$

Since $e = e^*_{AC}$ maximises the LHS of this inequality, equation 3 must also hold.

of beliefs of a continuum of individuals with beliefs distributed on the unit interval, $\hat{\pi} \sim F(\hat{\pi})$, such that $f(\hat{\pi})$ is everywhere positive on $\hat{\pi} \in [0, 1]$. Since shifting the distribution is a prerequisite to the rest of the analysis, we call this Hypothesis 0.

- **Hypothesis 0 (Shift in Beliefs):** The Hard-Easy treatment will shift the distribution of beliefs upwards.

An upward shift in $\hat{\pi}$ for all individuals (constrained at $\hat{\pi} = 1$) will imply an increase in the fraction of individuals choosing the ability-contingent incentives since $\hat{\pi} \geq \frac{\bar{w}}{w_H}$ will hold for a greater fraction of individuals.

If the time constraint on effort choices is binding, then an increase in confidence will not change effort. Otherwise, more confident individuals will exert more effort under ability-contingent incentives, as shown above.

- **Hypothesis 1 (Incentive Choices):** An exogenous increase in confidence will lead to a higher fraction of individuals choosing the ability-contingent incentives.

- **Hypothesis 2 (Effort Choices):** We will observe one of the following two patterns of behavior for effort choices. Either: (i) [*interior solutions*] For high confidence individuals, effort choices are higher under the ability-contingent incentives, than under the certain piece rate incentives. In this case, an upward shift in confidence will increase overall average effort. (ii) [*boundary solutions*] Effort choices are not influenced by the incentive scheme. In this case, an exogenous shift in confidence will not affect effort choices.

Ultimately, we want to understand the effect of an upwards shift in confidence on earnings. By definition, half of our individuals are classified as below average and half of them as above average ability. Depending on their actual relative ability, a shift in beliefs will lead to different outcomes for both groups.

We define $F_L(\hat{\pi})$ to be the distribution of subjective beliefs of low ability individuals (a_L), and $F_H(\hat{\pi})$ to be the distribution of subjective beliefs of high ability individuals (a_H). This implies that $F_m(\pi^i)$, where $m \in L, H$, denotes the fraction of individuals with ability a_m that choose the certain piece rate incentives. Similarly, if we consider an upward shift in confidence, $\triangle\hat{\pi}$, then after the shift in beliefs, $F_m(\pi^i - \triangle\hat{\pi})$ is the fraction of individuals with ability a_m that choose the certain piece rate incentives.

Expressions 4 and 5 reflect the intuitive idea that an upward shift in confidence will be harmful to low ability individuals who switch to inappropriate ability-contingent incentives, and beneficial to high ability types who, absent the treatment, would not have chosen the ability-contingent scheme.[5].

Gain in Earnings for High Ability Individuals:

$$[F_H(\pi^i) - F_H(\pi^i - \triangle\hat{\pi})] \cdot (w_H \cdot e^*_{AC} - \bar{w} \cdot e^*_{PR}) \geq 0 \tag{4}$$

Loss of Earnings for Low Ability Individuals:

$$[F_L(\pi^i) - F_L(\pi^i - \triangle\hat{\pi})] \cdot (0 - \bar{w} \cdot e^*_{PR}) \leq 0 \tag{5}$$

The term $F_m^{SWITCH} = [F_m(\pi^i) - F_m(\pi^i - \triangle\hat{\pi})] \geq 0$ denotes the fraction of individuals of ability type a_m who switch from certain piece rate to ability-contingent incentives when there is an upward shift in confidence by $\triangle\hat{\pi}$. The magnitude of the change in earnings will depend on several factors: (i) the number of individuals who switch their incentive scheme choice, F_m^{SWITCH}, in each group, (ii) the change in effort between the two incentive schemes, and (iii) the incremental size of the gaps between the wages. Thus we cannot predict the average change in earnings.

However, if an upwards shift in confidence leads to an increase in earnings for high ability individuals and a decrease for low ability individuals then earnings inequality will increase.

- **Hypothesis 3 (Earnings):** While an increase in confidence has an ambiguous effect on average earnings, the framework suggests that: (i) it will lead to weakly lower average earnings for low-ability individuals, (ii) it will lead to weakly higher average earnings for high-ability individuals, and (iii) it will increase earnings inequality overall.

Considering the evidence on differences in beliefs between genders we derive additional predictions for testing gender effect. Assume that conditional on ability type, men have a higher confidence level than women, i.e. $\forall\hat{\pi} : F_m^M(\hat{\pi}) \leq F_m^W(\hat{\pi})$, where $F_m^G(\cdot)$ denotes the subjective

[5] The expressions in the main text assume an interior solution for effort choices. However, if there is a binding constraint on effort choices then expressions 4 and 5 simplify to:

Gain in Earnings for High Ability Individuals

$$[F_H(\pi^i) - F_H(\pi^i - \triangle\hat{\pi})] \cdot (w_H - \bar{w}) \cdot \bar{e}$$

Loss of Earnings for Low Ability Individuals

$$[F_L(\pi^i) - F_L(\pi^i - \triangle\hat{\pi})] \cdot (-\bar{w}) \cdot \bar{e}$$

belief CDF for individuals of gender, $G \in \{M, W\}$, and ability type a_m, with $m \in L, H$. This leads to the following hypothesis:

- **Hypothesis 4 (Gender Differences):** Conditional on ability level, the average man will: (i) hold higher beliefs about his ability, (ii) is more likely to choose the ability-contingent incentives, (iii) will choose higher effort (if non-binding), in comparison to the average woman.

3 Experimental Design

Our main objective is to assess how an agent's confidence in her relative ability causally affects her choice of incentive scheme for a real effort task in which she can choose between fixed or ability-contingent incentives. This mirrors the labour market decision of whether to pursue employment that is highly dependent on one's ability or not. Secondly, we evaluate the relationship between the agent's confidence and her effort provision under the chosen incentives.

An experiment with these objectives should have the following features:

i. **An exogenous shift in subjects' beliefs about their relative ability**,

ii. **a separation of the role of ability and effort in the production function**,

iii. **a fine-grained measurement of a participant's willingness to switch from a fixed incentive scheme to an ability-contingent scheme**, and

iv. **minimisation of the influence of social preferences, risk preferences, and competitive preferences**.

3.1 The ability and the effort measurement tasks

Our experiment has all four features. Figure 1 outlines the timeline of the experiment. The main components are the "Ability Task" to measure a and the "Effort Main Task" to measure e. While it is impossible to obtain a clean measurement of ability[6], completely free of the influence of effort[7], and vice versa, we argue that the choice of tasks comes as close as possible in an experimental setting.

[6] We view *ability*, a, as being a fixed characteristic of the individual that she cannot change during the time frame of the experiment.

[7] We view *effort*, e, as a being a malleable object that the participant has full control over.

The “Ability Task” consists of a test that is often used to measure IQ. Subjects have four minutes to solve as many Raven Progressive Matrices[8] as they can. Subjects can go back and forth between the 12 matrices and can change their answers until the time is up. Every correct answer yields one point, and there are no negative points for wrong answers. The task is not directly incentivised, however performing better in this task implies the possibility to earn more in the “Effort Main Task”. Furthermore, IQ tests tend to induce an intrinsic motive to perform well, and the lack of financial incentives has the advantage that it should reduce the hedging motive faced by subjects when we elicit their belief about their performance in the “Ability Task”. At this point, subjects did not know that they would be incentivized for accurate beliefs at a later time to prevent them from performing poorly.

The “Effort Main Task”, on the other hand, was chosen to be a task where participants had a lot of control over their performance (i.e. it depended predominantly on how much effort they chose to exert). We chose the slider task by Gill and Prowse (2011). Using the mouse, participants move sliders on the screen from position zero to position 50. Sliders are shown in a set of 20. When all 20 sliders are set to 50 the subject can click the submit button and the sliders will be reset to zero for a new round. In Section 3.4 below, we will discuss the incentive choice that subjects make, and elaborate on how the ten sets of two minutes allow us to obtain a fine-grained measure of subjects' valuation of the ability contingent incentives.

Since the treatment is introduced during the “Ability Task” phase of the experiment, we also measure each individual's baseline effort level prior to treatment under fixed piece rate incentives. This serves two purposes. Firstly, it allows us to check for balance of effort in the slider task between treatment groups, prior to the treatment manipulation. Secondly, it allows us to control for baseline effort levels when assessing the impact of the treatment, thereby reducing unobserved individual level heterogeneity.

We discuss the remaining phases of the experiment in more detail below.

3.2 The treatment variation

Our aim was to shift subject's beliefs, keeping the treatment manipulation minimal. Therefore, the two treatment conditions are completely identical except for a slight difference in the difficulty of the ability task. Within each session, subjects are randomly assigned to one of two groups. One

[8] A matrix consists of nine related patterns of which one is missing. Below the matrix, there are eight possible patterns to complete the set of nine. Subjects have to find the one piece that best completes the set. There is only one correct answer. Due to the short time frame for solving up to 12 matrices, performance depends mostly on cognitive ability. Subjects are either skilled at finding the missing pattern or not. Every matrix is shown on its own screen.

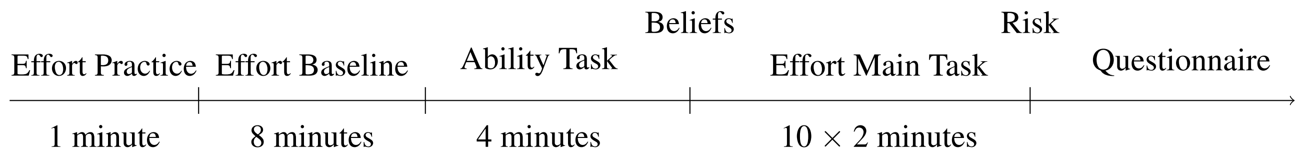


Figure 1: Sequence of experimental parts

group was exposed to a harder version, and the other to an easier version of the Raven Progressive Matrices. Eight of the twelve puzzles are identical across both treatments. The remaining four are either slightly easier or slightly harder than the rest.

This approach draws on the finding in the psychology literature that suggests that when individuals face a harder task, this shifts them towards lower confidence regarding their beliefs of their relative position in the ability distribution. When individuals face an easier task, this shifts them towards being more confident regarding their position in the distribution (Burson et al., 2006; Healy and Moore, 2007; Larrick et al., 2007; Moore and Healy, 2008; Bordley et al., 2014; Benoît et al., 2015). Importantly, it is assumed that the composition of the group stays constant, so there is no reason for the individual's actual rank to change when the difficulty of the test is shifted. However, when determining their ranking within a group, individuals anchor on their assessment of their own performance and don't adjust their belief about the distribution of others' scores sufficiently. Kruger (1999) shows that this miscalibration can lead to the majority of subjects evaluating themselves as worse-than-average in difficult tasks and better-than-average in easy tasks.

In our experiment, we therefore name the treatment in which subjects face the hard test, the "Low confidence" treatment, and the treatment in which subjects face the easier test, the "High confidence" treatment.

3.3 The belief elicitation

After the "Ability Task" we elicit subjects' beliefs about their relative performance in comparison to a group of 9 other participants in the same room who faced the same level of difficulty in the task[9]. The main question we are interested in is: "What do you think is the probability that you scored among the top 5 participants in the IQ picture task?". We provide the participants with a scale of possible answers ranging from "0 - I am certain that I scored in the bottom half" to

[9] Subjects knew that they were randomly assigned to one of two groups of ten within the session, and that all the other members of their group faced the same "Ability Task", however they didn't know that the other group faced a different "Ability Task". It was made clear to them that only their own group's performance was relevant for them, and there was no interaction at all between groups.

"100 - I am certain that I scored in the top half". They were free to state any number from 0 to 100. Their guess is incentivized using the quadratic scoring rule (Selten, 1998). The quadratic scoring rule is explained in detail to them in the provided instructions, both on screen and on paper. The scoring rule is designed to provide the highest expected payoff when subjects state their true beliefs. Maximum earnings are €2 for the belief elicitation task. Further, we ask them to report their best guess of how many points they scored in the task and what they believe the 5th highest score in their group is (unincentivized, in order to avoid hedging). The belief elicitation came as a surprise at this point in the experiment to prevent hedging.

3.4 The payment scheme choice

One of the objectives of the experimental design was to obtain a fine grained measurement of subject's valuation of the ability contingent payment scheme. Therefore, we constructed a payment scheme that aims to achieve this. Instead of having a single choice between a fixed piece rate and an ability contingent payment scheme, subjects face ten rounds of two-minute real effort tasks. In each of these rounds (except the first), subjects can choose whether they would prefer a fixed piece rate, or whether they want to work under the ability-contingent piece rate. Importantly, in order to assess the value of the piece rate that makes the subject indifferent between the two incentive schemes, the fixed piece rate is incrementally increased in each round (see Figure 2 below). The round in which the subject switches allows the analyst to assess the subject's valuation of the ability contingent incentives.

A second important feature of the "Effort Main Task" design is that the ability component of the production function is fixed through the "Ability Task", before participants report their belief, and before they reach this "Effort Main Task" in which they make choices between incentive schemes. The ability-contingent piece rate has a high payoff of €1 per 20 sliders, if the subject was in the top half of her group in the "ability task" and a low payoff of €0 per 20 sliders if she was in the bottom half. This implies that the optimal incentive scheme choice for the participant has already been determined when she makes her incentive scheme choices. Furthermore, we as the analysts know what her optimal choice is.

This rules out potential issues that can arise if one were to allow participants to first make an incentive scheme choice, and then produce output that depends on both ability and effort. Firstly, participants would then be able to endogenously adjust their effort to their choice, and try to "make" their choice optimal (for example, by choosing ability contingent incentives and then compensating for low ability by exerting very high effort). Secondly, it would make it very difficult for us to separately study the influence of a shift in beliefs on incentive scheme choices and effort choices. Our design aims to overcome these challenges.

Subjects receive no feedback about either their relative ability score or their performance in any components of the “Effort Task” tasks until the very end of the experiment. With the exception of Round 1, before the start of each round, participants choose whether they want to work for the fixed piece rate or whether they want to work for the ability-contingent piece rate. The fixed piece rate increases in each period from €0.15 per 20 sliders in the second period to €0.80 per 20 sliders in the last period. Incentives for the fixed piece rate work like a multiple price list for safe vs. risky choices. Once the expected earnings from the ability-contingent piece rate are lower than the fixed piece rate in that period, individuals should switch to the fixed piece rate and choose it for the remainder of the experiment, assuming risk neutral preferences. Expected earnings from the ability-contingent piece rate depend on the subject’s belief. For example, if a subject is 60% sure to be in the top half then she should switch to the fixed piece rate in period 7 when the fixed piece rate of €0.65 is larger than her expected ability-contingent piece rate of €0.60. Figure 2 gives an overview of the payment scheme.

Period Number (2 min each)	Option A Payment (per 20 sliders)	Option B Payment (per 20 sliders)
1	**Can't Choose Option A**	**€1** if TOP HALF; **€0** if BOTTOM HALF
2	**€0.1**	**€1** if TOP HALF; **€0** if BOTTOM HALF
3	**€0.3**	**€1** if TOP HALF; **€0** if BOTTOM HALF
4	**€0.4**	**€1** if TOP HALF; **€0** if BOTTOM HALF
5	**€0.5**	**€1** if TOP HALF; **€0** if BOTTOM HALF
6	**€0.55**	**€1** if TOP HALF; **€0** if BOTTOM HALF
7	**€0.6**	**€1** if TOP HALF; **€0** if BOTTOM HALF
8	**€0.65**	**€1** if TOP HALF; **€0** if BOTTOM HALF
9	**€0.7**	**€1** if TOP HALF; **€0** if BOTTOM HALF
10	**€0.8**	**€1** if TOP HALF; **€0** if BOTTOM HALF

Figure 2: Payment Scheme in Main Effort Task

In the first period, all participants have to work for the ability-contingent piece rate. This feature allows us to assess how the shift in confidence affects effort provision when all subjects are forced to work under the ability-contingent piece rate - in all other rounds the incentive scheme that a subject faces is endogenous. In the end, five of the ten rounds are randomly chosen for payment. Subjects had to answer four control questions before starting the task to ensure comprehension of the payment scheme. To elicit the baseline motivation of moving sliders, subjects complete 9 minutes of the "Effort Task" at the start of the experiment, the first minute being an unincentivized practice round. In the baseline round, we pay €0.30 per 20 sliders, and all completed sets are paid out.

3.5 The risk elicitation

Finally, we elicited risk preferences by adapting the preferences module on risk taking by Falk et al. (2016) to our setting. The staircase procedure is essentially equivalent to a traditional multiple price list, presenting multiple choices between a sure payoff and a gamble, but simply requires fewer decisions on the part of the subject in comparison to a traditional price list. It achieves this as follows: Depending on whether the subject chose the sure payment or the gamble, the algorithm generates a new choice, which makes the option that was not chosen slightly more attractive by increasing or decreasing the sure payoff. Therefore, it avoids asking subjects to make redundant choices.

The staircase has four choices between a sure payment and a gamble. The outcome of the gamble was always €0 or €1 and the probability of winning was 50 percent. The sure payment varied in its amount. One of the decisions was randomly chosen for payment.[10]

In the end, we administered a comprehensive questionnaire.

3.6 The procedure

The experiment was programmed in zTree (Fischbacher, 2007) and conducted at a laboratory at a large university in 2017. Participants were solicited through an online database using ORSEE (Greiner, 2015) from a subject pool of mostly undergraduate students from all faculties. In total 100 subjects participated in 5 sessions, 20 in each. 47 of them were female, 49 male and four chose not to self-report their gender. Subjects received a show-up fee of €5 plus their earnings from the tasks. Mean earnings for the 60 minutes sessions amounted to €13.30. The relevant instructions were handed out to participants at the beginning of each stage and read out loud.

[10] We have two more staircases for which we use the subjects' own reported beliefs as the probability for the gamble. We do this as it allows us to compare their choices in the risk task with their incentive scheme choices. While the risk task involves a choice between a pure lottery and a fixed payment, the main task involves the choices of which payment scheme to work under. Both tasks share the same values of outcomes and subjective probabilities. Therefore, it allows us to assess whether there are any differences in how probabilities affect choice tasks and effort tasks. As a caveat, the risk elicitation always comes after the incentive scheme choices, which could motivate subjects to make consistent choices in the risk task. The third staircase is a mirror of the second and only used for keeping incentives fair to not favor individuals that stated high beliefs, as high beliefs increase the chance of winning the gamble.

4 Results

Section 4 shows the effects of an exogenous shift in beliefs on the choice of the incentive scheme, effort, and earnings. In section 5 we draw a parallel between the choices induced by our treatment and observed gender differences in choice.

4.1 Exogenously shifting beliefs

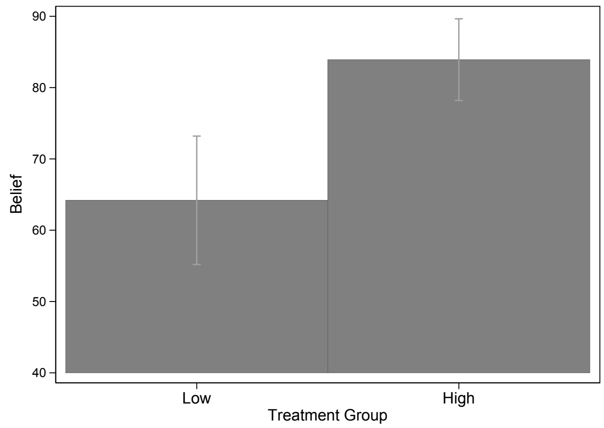
The main objective of our treatment was to achieve an exogenous shift in the participant's beliefs about their relative performance in an IQ test. It is therefore of central importance to confirm that we did indeed shift participants' beliefs. We hypothesized (*Hypothesis 0*) that average beliefs about relative performance in the HIGH confidence treatment would be higher than average beliefs about relative performance in the LOW confidence treatment.

We find a significant difference in the participants' level of confidence in the two treatment groups. We measure an individual's level of confidence as the stated probability of being in the top half of their randomly assigned group of ten subjects[11]. Participants exposed to the easy test state on average higher levels of confidence than participants exposed to the more difficult test.

Figure 3 shows that the average participant exposed to the easy ability task reported that there was an 84 percent chance that she was in the top half of her group of 10. Using a t-test, we find this to be significantly higher than the belief of 63 percent reported by participants exposed to the more difficult test (p-value < 0.01).

[11] There were twenty participants in each session. Within each session, half were randomly assigned to each treatment condition. However, subjects only knew there were two groups of ten and that they would only interact with their group.

Figure 3: Average stated beliefs by treatment



Note: (i) Vertical lines denote 95% CI around the mean.

This significant shift in observed beliefs justifies our use of the names, HIGH and LOW for the high confidence [easy IQ test] and low confidence [difficult IQ test] treatment groups.

To confirm that we are working with a balanced sample (which should be the case due to within-session randomization), we show that there are no significant differences in the characteristics of the subjects between the treatments. Table 5, reported in the Appendices, presents summary statistics for the participants, showing that the average of the gender, age, pre-treatment baseline effort level and 50-50 risk variables are not significantly different between treatments.

This table also reports the differences in absolute scores under the easy and difficult IQ test, which generates the treatment difference in beliefs about relative performance. In the HIGH confidence treatment subjects solved on average 10.9 Raven matrices with a standard deviation of 1.20. In the LOW confidence treatment subjects solved on average 6.9 Raven matrices with a standard deviation of 2.80. However, by construction, in each treatment exactly half the subjects are in the TOP HALF of the within-treatment ability distribution, and half are in the BOTTOM HALF.

Result 0 *A reduction in the level of difficulty of the ability task increases the average confidence of the participants about their relative performance.*

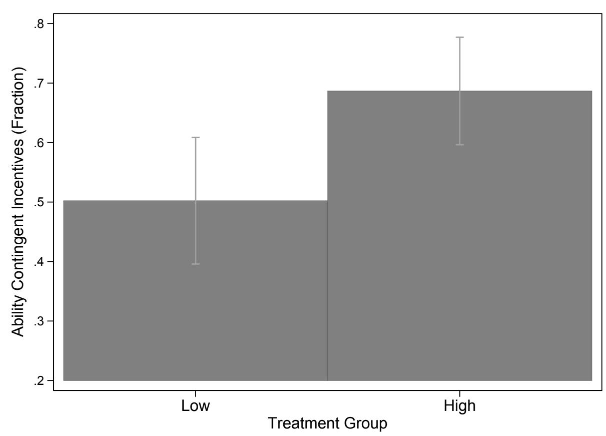
4.2 Influence of beliefs on incentive choice

Next, we show whether the exogenous increase in confidence leads to a higher fraction of individuals choosing the ability-contingent (AC) incentives (Hypothesis 1). This is important because an individual who chooses the incentive scheme that is most appropriate for her ability type will earn substantially more than the individual who chooses the inappropriate incentive scheme. In Section 4.4 below, we will discuss the magnitude of the influence of the incentive choice on an individual's earnings.

In the experiment, participants chose between a fixed piece rate and an ability-contingent piece rate nine times (i.e., in Rounds 2 to 10). In each successive round the value of the fixed piece rate was incrementally increased, as discussed in Section 3 above (i.e. let $\bar{w}^r$ denotes the fixed wage in round $r \in \{2, 3, ..., 10\}$). In order to maximize their expected earnings a risk-neutral individual should switch to the fixed piece rate once the level of the fixed piece rate exceeds the expected piece rate under ability-contingent incentives, according to the individual's belief about her type (i.e. for $\forall r : \hat{\pi} > \pi^{i,r}$ where $\pi^{i,r} := \frac{\bar{w}^r}{w_H}$).

This logic implies that an upward shift in participants' beliefs should lead to a shift towards choosing the ability-contingent incentives more often. Indeed we find that in the HIGH treatment the fraction of ability-contingent piece rate chosen is significantly higher than in the LOW treatment (ttest p < 0.01). See Figure 4 (and Table 5).

Figure 4: Propensity to Choose Ability-Contingent Incentives



Note: (i) Vertical lines denote 95% CI around the mean.

The first two columns of Table 1 below simply reiterate the same point, by showing that there is a significant difference in the propensity to choose the ability-contingent incentives between the two treatment groups. Columns (3) to (5) show that the subjects reported beliefs about their likelihood of being in the TOP HALF are highly predictive of their incentive scheme choices. More specifically, a 1 percentage point increase in a participants' belief is associated with her choosing the ability-contingent incentives 0.86 to 0.94 percentage points more often. A nice feature of our experiment is that the exogenous shift in beliefs due to our treatment, allows us to instrument for the belief variable in columns (4) and (5), showing that this result is not driven by other unobserved differences between individuals who hold high beliefs and low beliefs.

Table 1: Propensity to Choose the AC Incentives

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)
Treatment (HIGH=1)	0.18** (0.07)	0.18** (0.07)			
Subj Belief			0.86*** (0.10)	0.94*** (0.26)	0.93*** (0.26)
Risk (CE p=50)					0.21 (0.19)
Constant	0.50*** (0.05)	0.55*** (0.08)	-0.01 (0.09)	-0.07 (0.20)	-0.17 (0.22)
Session Fixed Effects		✓	✓	✓	✓
Observations	100	100	100	100	100
Adjusted R^2	0.058	0.068	0.462		
First-Stage F				13.91	13.88

Notes: (i) In the IV Regressions, Subjective Beliefs are instrumented using the treatment dummy. (ii) Standard errors in parentheses. (iii) Dependent variable: fraction of AC choices in rounds 2 to 10.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Result 1 *An increase in the participants' level of confidence led to a higher propensity to choose the ability-contingent incentives.*

While the discussion above has focused on the average propensity to choose the ability-contingent incentives across all rounds, we can also examine the incentive choices within each round. Since the fixed piece rate increases incrementally in value over the rounds, we would expect that the fraction choosing the ability-contingent incentives would decrease over the rounds in both treatment groups. However, the main question of interest is whether there is a significant difference in the fraction choosing ability-contingent incentives between the treatment groups. Figure 5 provides evidence towards addressing this question.

The figure plots the fraction choosing the ability-contingent incentives in each round for each treatment group. In addition, we plot a 95% confidence intervals around the means, clustering the standard errors at the individual level.[12]

Figure 5: Propensity to Choose Ability-Contingent Incentives (by Round)

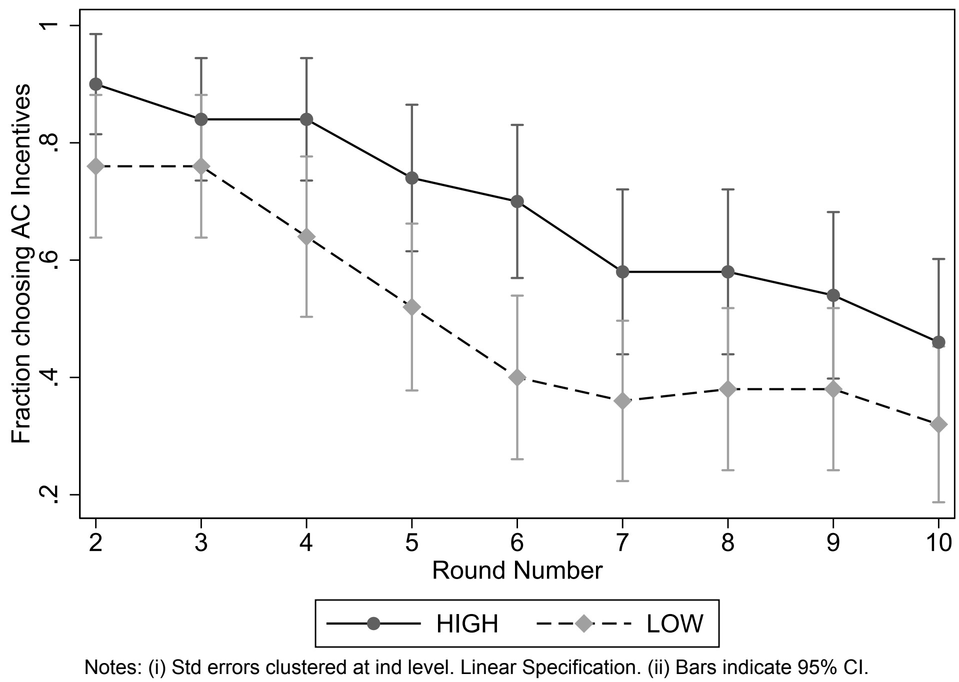


Figure 5 shows that in both treatments, the fraction of subjects choosing the ability-contingent incentives is decreasing from round 2 to round 10, with much of this decrease taking place between round 4 and 7. The fraction choosing ability-contingent incentives is higher in the HIGH treatment in all rounds. However, this difference is not significant for the low and high rounds. Table 7 in the Appendix shows that while there is significant difference in the propensity to choose the ability-contingent incentives when all rounds are pooled together for each individual (t-test $p < 0.01$), if we consider each round separately, only the middle rounds (ie. rounds 4 to 8) yield a significant treatment difference at the 5% level[13]. One reason for this is that there are some individuals who are highly certain that they are in the TOP HALF, and therefore always choose the ability-contingent incentives, and similarly, there are individuals who always choose the certain piece rate incentives.

[12] Essentially, we estimate a linear regression, and calculate 95% confidence intervals around the coefficients, using a typical marginal effects plot.

[13] Round 6 is significant at the 1% level, and round 2 is significant at the 10% level.

Therefore, in Rounds 2 and 3 both treatments are close to the ceiling of ability-contingent choices, while in round 9 and 10, both treatment groups are close to the floor of ability-contingent choices.

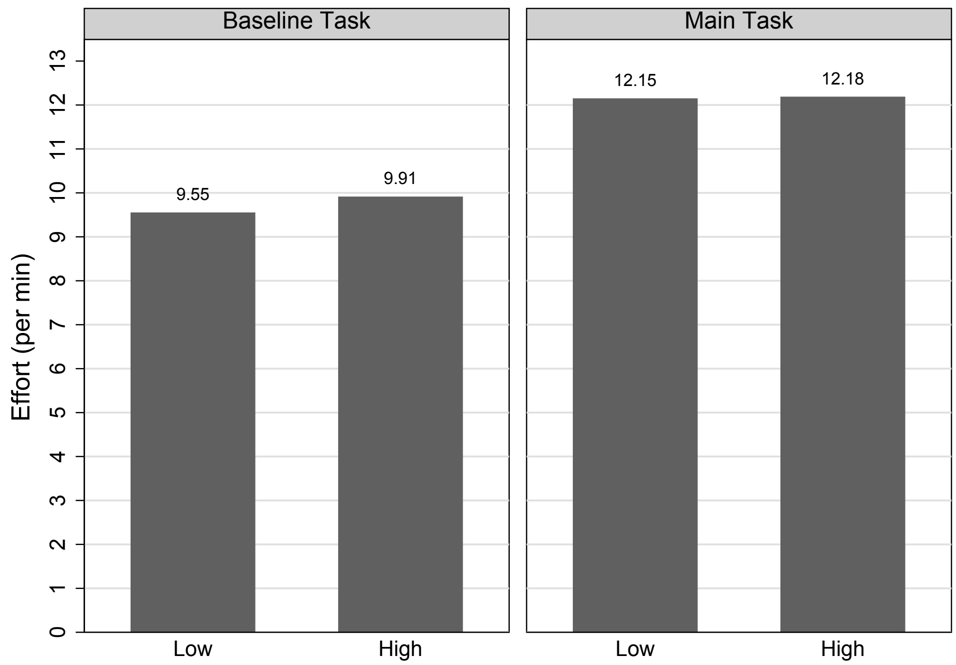
Overall, the evidence suggests that there was a strong behavioral response to the shift in confidence due to the treatment, implying a large difference in participants' likelihood of taking up the riskier ability contingent incentives. As we will discuss below, this shift in incentive scheme choices is harmful to some and beneficial for others.

4.3 Influence of beliefs and incentive choice on effort

Once an individual has chosen her incentive scheme, the second choice she has to make within a job is the choice of how much effort to exert. With regards to this effort choice, our simple theoretical framework yields two sets of predictions, depending on whether the effort choice is an interior solution to the optimization problem, or whether the conditions in the lab imply that effort is constrained (e.g., by the time available), leading to a boundary solution. As mentioned above, this latter scenario may be more relevant if the cost of effort in the lab depends mostly on the *duration* that effort is expended, and the time is insufficient for an interior solution. A high intrinsic motivation, s, might then mask any differences in effort due to incentives. The former may be the case if the cost of effort depends on substantially more on the *intensity* of effort exerted implying an interior solution is more likely in a short time-frame. This difference is only possible to observe if s, is low enough.

Taken together, the data collected in our experiment is more consistent with the second scenario (constrained effort choices) than the first. We measure effort using the variable "effort per minute", which reflects the number of sliders completed during each minute within a particular round. Mainly, we find no significant difference in average effort exerted between treatment groups (see Figure 6 and Table 7 in the Appendix).

Figure 6: Per minute effort in baseline and main task by treatment

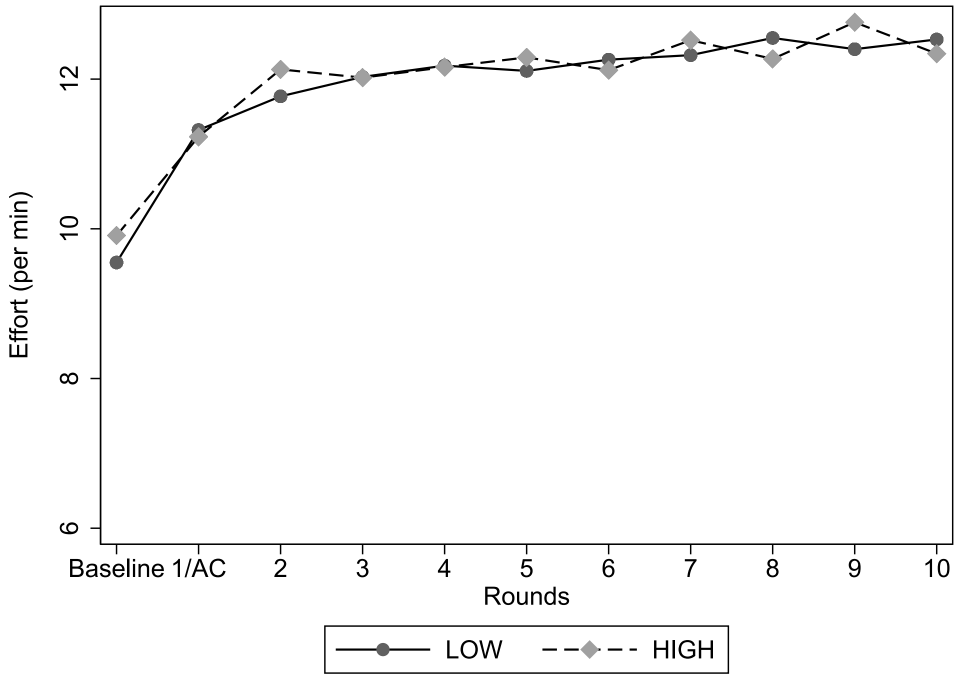


Additionally, we present two pieces of evidence that suggest that effort is not responding to the participant's incentive choice, nor to the participant's beliefs when she does face the ability-contingent incentives [14].

Figure 7 plots the average per minute effort exerted in the baseline round, as well as in every subsequent round. While we do see some initial learning, after the baseline round, there is very little change in effort exerted even though the value of the piece rate under the certain piece rate incentives increases from €0.1 to €0.8, and the fraction of individuals choosing these certain piece rate incentives increases substantially in both treatment groups.

[14] Recall, that in the discussion above we saw a strong response of incentive choice to beliefs above, so we can rule out the hypothesis that beliefs are not meaningful.

Figure 7: Effort choices across rounds, by treatment



Furthermore, Figure 8 focuses on the first round in which all participants were forced to face the ability-contingent incentive scheme. This feature avoids the endogeneity issue that we face in subsequent rounds in which participants choose both their effort level and their incentive scheme. Considering the full sample, this figure shows that in both treatments, effort is highly unresponsive to beliefs. In particular, even individuals who reported a belief of zero, still exert effort that is similar to the mean effort exerted by those who stated a belief of one hundred. While our treatment successfully shifted the beliefs of participants in the two treatments, it did not affect the relationship between beliefs and effort, which is rather flat.

This finding is not surprising considering the recent literature on the unresponsiveness of effort to incentives in real-effort lab experiments (de Araujo et al., 2015; Corgnet et al., 2015; DellaVigna and Pope, 2016; Erkal et al., 2016). Most papers find that there is significant intrinsic motivation to work on the task, regardless of any incentives. There is some evidence that this motivation can be reduced by giving the subjects the opportunity for leisure activities or the chance to leave the experiment or when each unit of effort itself is made costly (see for example the ball catching task by Gächter et al. (2016).) We did allow subjects to use their mobile phones in silent mode during the main task if they did not want to work.

Figure 8: Per minute effort in first round under ability-contingent incentives

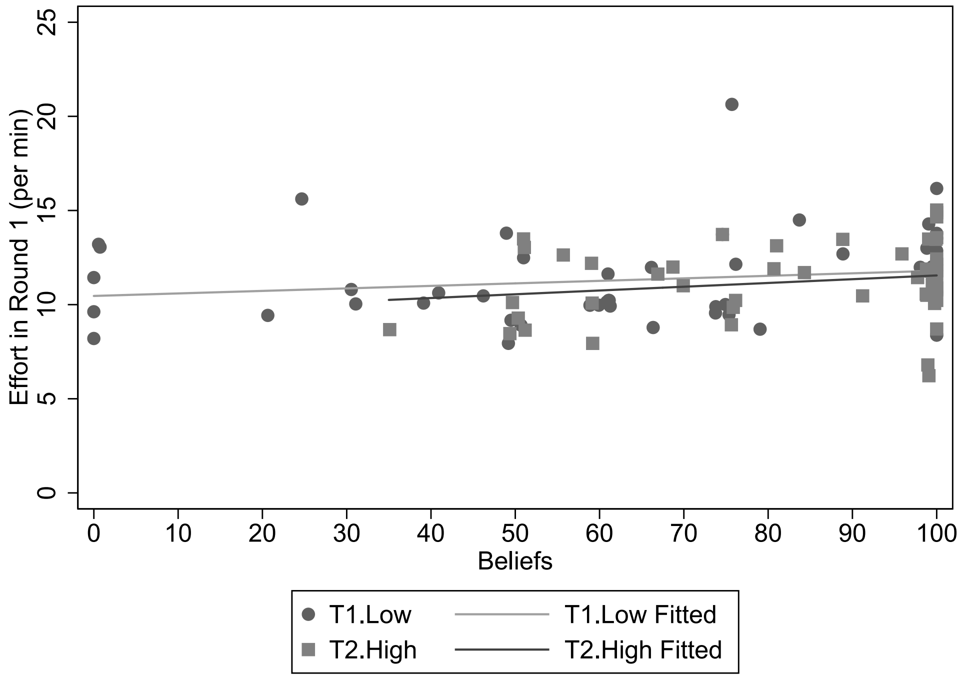


Table 2 reiterates these results by examining the correlates of effort in Round 1 (i.e., under ability-contingent incentives). In particular, columns (1) and (2) confirm that there is no treatment difference in effort choices, columns (3) and (4) provide further evidence that there is no significant relationship between an individuals' belief and her effort choices. In columns (5) and (6), we include baseline effort, which was chosen before participants received treatment. These regressions show a strong correlation between baseline effort, and effort in round 1, suggesting that intrinsic motivation, s, is an important component in the effort function. Furthermore, we again observe no significant relationship between effort and beliefs (although, the negative sign on the point estimates is slightly surprising). Taken together, we view the evidence as being highly consistent with the conclusion that effort choices are constrained within the context of our experimental design. The unresponsiveness of effort has the positive feature of allowing us to study the implications of shifting beliefs on incentives choices, and earnings, while effort choices are essentially constant.

Result 2 *Effort choices are largely unresponsive to shifts in beliefs, and to the participant's choice of incentive scheme. This is likely due to the time constraint of the experiment and a high intrinsic motivation.*

Table 2: Effort Choice (per minute) Under AC Incentives (Round 1)

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)
Treatment (HIGH=1)	-0.09	-0.09				
	(0.45)	(0.44)				
Subj Belief			1.11	-0.46	-2.29	-2.30
			(0.79)	(2.19)	(1.79)	(1.79)
Baseline Effort					1.00***	1.00***
					(0.16)	(0.16)
Risk (CE p=50)						0.50
						(1.37)
Constant	11.32***	11.97***	11.09***	12.27***	3.91***	3.68**
	(0.31)	(0.54)	(0.77)	(1.76)	(1.10)	(1.27)
Session Fixed Effects		✓	✓	✓	✓	✓
Observations	100	100	100	100	100	100
Adjusted R^2	-0.010	0.013	0.033			
First Stage F				13.92	13.09	13.03

Notes: (i) Dependent variable: Round 1 effort per minute. (ii) Higher values of risk variable (i.e. certainty equivalent for 50-50 gamble) imply risk loving. (iii) Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.4 Earnings

Following our theoretical framework, we now turn to the effect of increased confidence in earnings. Hypothesis 3 stated that increased average confidence will lead to i) weakly lower earnings for low-ability individuals, ii) weakly higher earnings for high ability individuals and iii) result in a higher earnings inequality overall. The discussion below provides evidence towards evaluating these hypotheses.

Our predictions for the influence of a shift in confidence on earnings of the average individual are ambiguous. Since effort choices are fairly inelastic, the effect of a shift in confidence on average earnings depends on: (i) the fraction of incentive choice switchers of each ability type, namely TOP HALF (α_H), and BOTTOM HALF (α_L), and (ii) the change in wage for each switcher ($\sum_r [0 - \bar{w}^r]$ for α_L, and ($\sum_r [w_H - \bar{w}^r]$ for α_h, summing over all rounds r for which the individual switches).

Table 3 shows the change in an individual's main task earnings if she switches from certain piece rate incentives to ability-contingent incentives in all rounds, conditional on her ability type.

Since there is some heterogeneity in effort choices between the TOP HALF and BOTTOM HALF individuals[15] the table uses the average effort choices of each ability group. Comparing these values to the average main task earnings of €7 shows that choosing the optimal incentive choice can have a substantial influence on earnings, within the context of the experiment. Furthermore, the benefit to TOP HALF individuals of switching to ability-contingent incentives is of a similar magnitude to the loss to BOTTOM HALF individuals. The overall effect of an increase in confidence on earnings is therefore likely to depend predominantly on the number of switchers of each ability type.

Table 3: Potential gains / losses from shifting incentive choice from PR-always to AC-always

	Piece Rate Inc (PR)	Ability-Contingent Inc (AC)	Net Change in Earnings
Bottom Half	€5.4	€0	**– €5.4**
Top Half	€6.8	€12.6	**€5.8**

Notes: (i) This table uses observed average effort by all TOP (12.58 per min) and BOTTOM (11.75 per min) individuals. (ii) The Net values are the change in earnings when switching from PR to AC. (iii) The difference between TOP and BOTTOM under PR incentives is due to Round 1, where all participants face AC incentives, and due to the slight difference in ave. effort mentioned in (i).

Figure 9: Average earnings by treatment

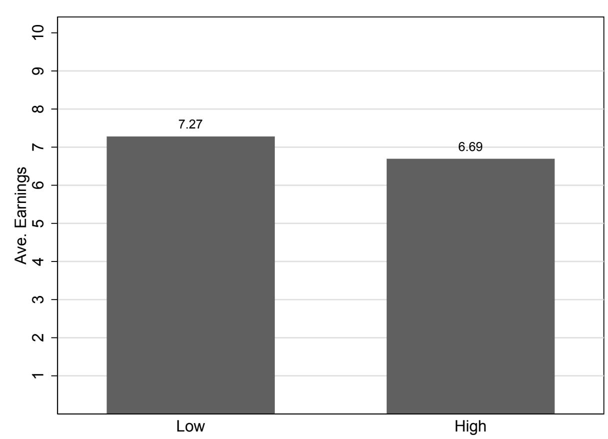


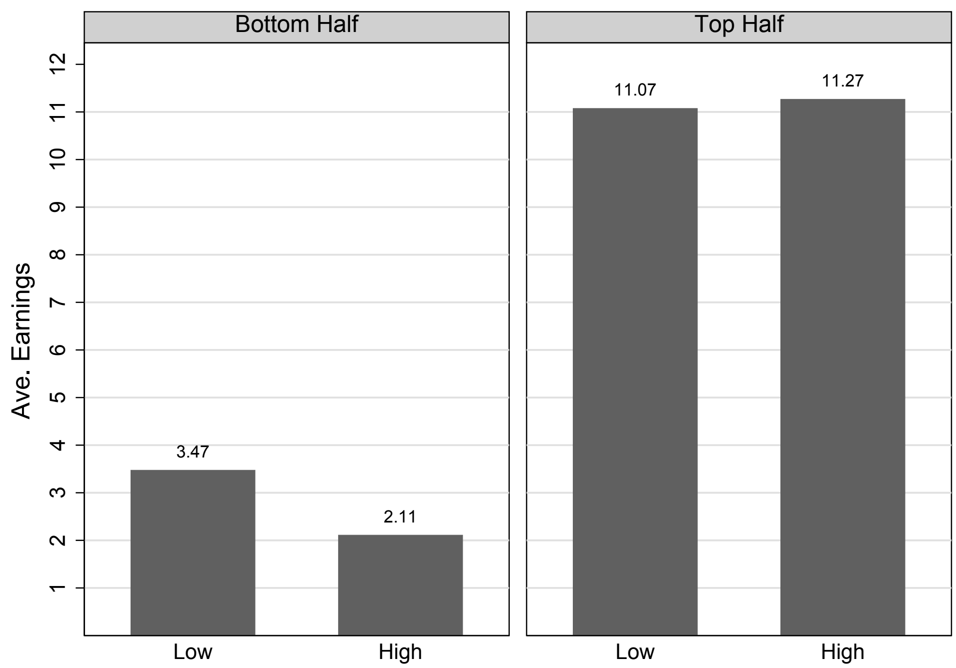
Figure 9 presents evidence on the average effect of the shift in confidence in our experiment. It

[15] TOP HALF individuals complete slightly more sliders (12.58 per minute) in comparison to BOTTOM HALF individuals (11.75 per minute). This difference is significant at the 10% level.

shows that average earnings decrease from €7.27 to €6.69 with an increase in confidence. However this difference is not significant (see Table 7 in the Appendix).

When we split the sample by ability, we find that this reduction in average earnings comes entirely from the low-ability individuals. We see this in Figure 10, with BOTTOM HALF individuals' earnings reduced by 40% from €3.47 to €2.11, and TOP HALF earnings almost unchanged by the treatment, at just above €11. Table 4 confirms the results displayed in these figures, showing that there is a significant drop in the earnings of the BOTTOM HALF group ($p < 0.05$). The TOP HALF experienced a small, insignificant increase in earnings. The average effect when pooling ability types has a negative sign but is not significant.

Figure 10: Average earnings by ability and treatment



This evidence illustrates the predictions discussed in the theoretical framework section, showing that an increase in confidence leads to a drop in earnings for the low ability individuals who are already earning far less, and thereby moving in the direction of higher overall earnings inequality[16].

[16] While the GINI coefficient increases from 0.275 in LOW to 0.293 in HIGH, and Figure 17 provides suggestive evidence of higher inequality in the HIGH treatment by plotting the earnings histograms of both treatments, a Mann-Whitney ranksum test indicates that there is no significant difference between the earnings distributions in the two treatments. Furthermore, a difference-in-difference estimate of the change in the earnings between the TOP HALF and BOTTOM HALF within each group has a negative point estimate of -1.56 but is not significant at the 10% level. We are underpowered to detect the effect of treatment on inequality.

Table 4: Change in Earnings due to Exogenous Belief Shift

	All		Bottom		Top	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (HIGH=1)	-0.59	-0.57	-1.36**	-1.35**	0.19	0.20
	(0.99)	(1.01)	(0.63)	(0.64)	(0.82)	(0.81)
Constant	7.27***	10.39***	3.47***	5.81***	11.07***	11.69***
	(0.70)	(2.32)	(0.45)	(1.55)	(0.58)	(1.87)
Baseline Effort		✓		✓		✓
Risk CE (p=0.5)		✓		✓		✓
Session Fixed Effects		✓		✓		✓
Observations	100	100	50	50	50	50
Adjusted R^2	-0.007	-0.029	0.070	0.062	-0.020	-0.001

Notes: (i) Dependent variable: Main Task Earnings, (ii) Std Errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

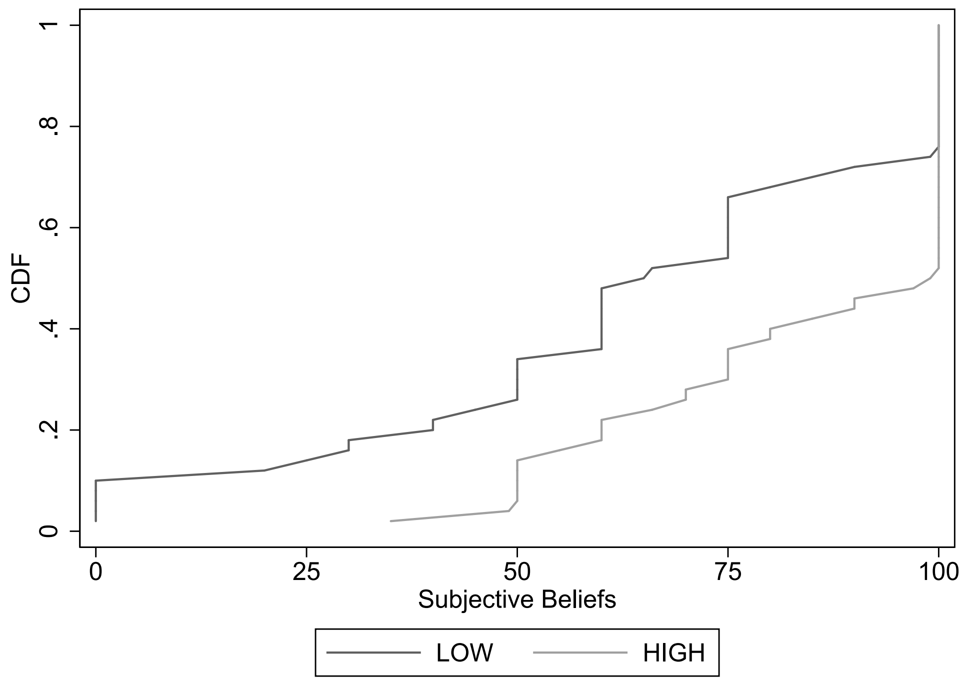
Result 3 *An increase in confidence leads to low ability individuals earning even less than their already low earnings, while high ability individuals are unaffected. This is suggestive of an increase in inequality with higher confidence, but our data does not permit us to estimate a significant change in inequality.*

One outstanding question is why we observe a relatively large decrease in earnings for the low ability individuals, but hardly any change in earnings for the high ability individuals. We consider this question in the following subsection.

4.4.1 Why is there a larger impact on the Bottom Half individuals?

The discussion above showed that while the increase in confidence reduced the earnings of the BOTTOM HALF individuals, it had no impact on the TOP HALF individuals. Since effort levels, within ability type, are relatively unresponsive to treatment, the lack of an impact on earnings of the TOP HALF individuals must be due to fewer of them switching their incentive scheme. In this section, we look at how the beliefs of the different ability types are affected by the treatment.

Figure 11: CDF of Beliefs, by Treatment



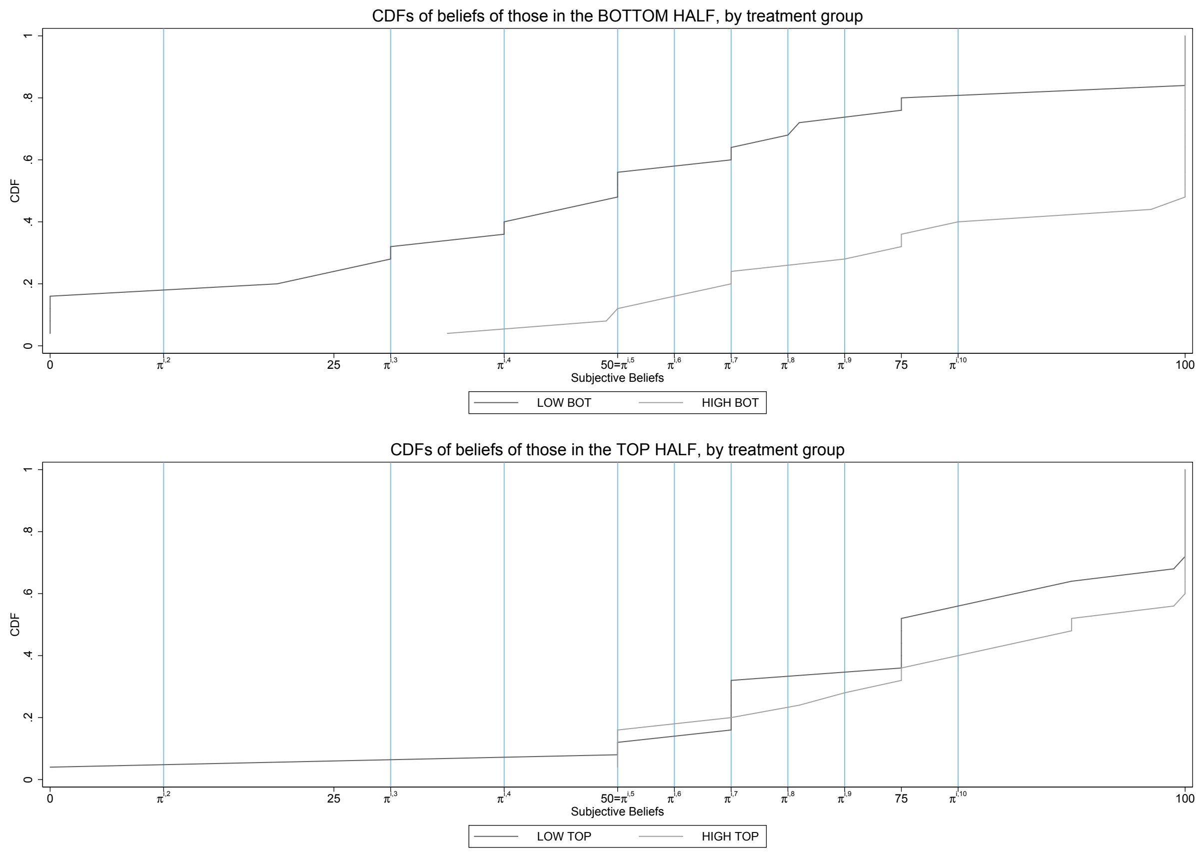
Firstly, Figure 11 plots the CDFs of beliefs in each of the treatment groups. The figure shows that the entire belief distribution is shifted to the right between the LOW and the HIGH treatment groups[17]. However, in order to understand why we only observe a shift in earnings for the BOTTOM HALF ability individuals, we need to consider the belief distributions of each ability type separately (as indicated in equations 4 and 5 by the $F_H(\cdot)$ and $F_L(\cdot)$ functions).

Figure 12 displays these belief CDFs for each ability type separately, comparing treatments. It is immediately apparent from these figures that the majority of the shift in beliefs between treatments is due to the shift in beliefs among individuals in the BOTTOM HALF of the distribution. One reason for this is that on average there is relative overconfidence even in the LOW treatment, with the TOP HALF individuals holding very high beliefs, leaving little room for their beliefs to increase. Essentially, the TOP HALF individuals are always confident that they are in the top half, and the treatment does little to shift this. In contrast, the BOTTOM HALF individuals appear to hold more malleable beliefs about their ability. When faced with an easier test, they adjust their

[17] A Mann-Whitney ranksum test indicates that the beliefs in the two treatments are drawn from different distributions ($p < 0.01$)

level of confidence upwards which leads to costly mistakes in incentive choices.[18]

Figure 12: CDF of Beliefs of TOP HALF and BOTTOM HALF, by Treatment



The vertical blue lines in Figure 12 refer to the beliefs thresholds, $\pi^{i,r}$, that indicate the optimal incentive choice for the risk-neutral individual in each round, r. For example, $\pi^{i,5} = 0.5$ is the threshold for round 5. The risk neutral individual should choose ability-contingent incentives in round 5 if her belief is higher than this threshold. Therefore, we can directly read off the fraction

[18] At first glance, this finding might remind the reader of the Dunning-Kruger effect (Kruger and Dunning, 1999). The Dunning-Kruger effect claims that low-ability individuals do not have the means to understand that they are low ability and thus grossly overestimate their relative ability, while high ability individuals can correctly assess their position or are even a bit underconfident in their relative abilities. If the Dunning-Kruger effect would be dominant in our experiment, we should have seen no effect of the treatment or potentially even the opposite. If more difficult tasks make it harder for low ability individuals to estimate their position in a relative ability ranking because they lack the knowledge to evaluate how well they did, then we should have seen *higher* average beliefs of the BOTTOM HALF individuals in the HARD test than in the EASY test treatment. And yet we see the opposite. It seems more likely that the perceived level of difficulty indicates how well they think they performed. A task that *feels* easy creates the belief that one is good at it and especially better than others.

of individuals who should choose ability-contingent incentives in each round, under risk neutrality, given their beliefs. This serves to illustrate equations visually 4 and 5, and to demonstrate how the shift in beliefs among the BOTTOM HALF translates into differences in incentive choices, which are less pronounced among TOP HALF individuals.

In summary, the impact of treatment on the earnings of the BOTTOM HALF, but not the TOP half is driven by the fact that the TOP HALF are already highly confident in their ability, and choosing ability-contingent incentives, while the BOTTOM HALF hold more malleable beliefs and are willing to be convinced that they are in the TOP HALF when taking an easier test.

5 Gender differences

Since there is the large body of evidence documenting that there tends to be a gender gap in confidence in one's own ability (see, e.g., Niederle and Vesterlund (2007), van Veldhuizen (2017) and Niederle (2017)), it is of interest to ask whether we observe this gender gap in our context. Based on this literature, we have hypothesized that the average man will i) hold higher beliefs about his ability and ii) is thus more likely to choose ability-contingent incentives[19].

Below, we test these hypotheses and describe how gender is correlated with our outcomes variables. Our subjects are balanced on gender between treatments and sessions. We show being male is associated with a similar magnitude upward shift in confidence as our treatment effect, and the pattern of outcomes observed in our treatment-control comparison is similar to the pattern of outcomes observed in our male-female comparison (except effort choices).

5.1 Gender differences in ability

To eliminate ability differences as an explanation for potential gender differences in beliefs and payment scheme choices, we selected an ability task that was not gendered and with no evidence of gender effects in previous experiments. Table 6 in the Appendix shows that men and women are almost identical in their average scores (8.98 and 8.92), and their probability of being in the top half of their group (0.51 and 0.49). Neither difference is statistically significant. The similarity in the performance of men and women implies that ability should not explain any observed differences

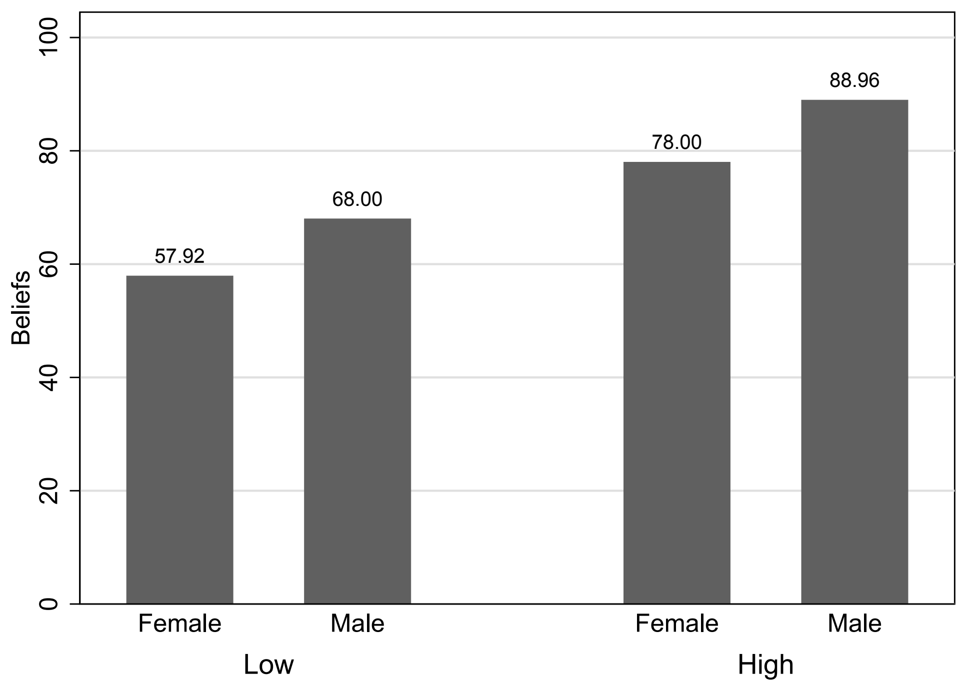
[19] We also hypothesized that the average man would choose a higher effort level under ability-contingent incentives if effort choices are unconstrained, due to his higher level of confidence. However, we have already established that effort choices are inelastic concerning beliefs in the context of our experiment. Table 6 shows that men do in fact exert significantly more effort than women ($p < 0.01$), but since this is also true for baseline effort ($p < 0.1$), it appears to be driven predominantly by something other than the gender difference in beliefs.

in beliefs between men and women.

5.2 Gender differences in beliefs and incentive choices

In line with the literature documenting the gender confidences gap, Figure 13 shows that in both our treatment conditions women state lower average beliefs about being in the top half. Table 6 shows that the gender-confidence gap is on average 11 percentage points and is significant at the 5% level[20].

Figure 13: Average beliefs about being in the top half by gender



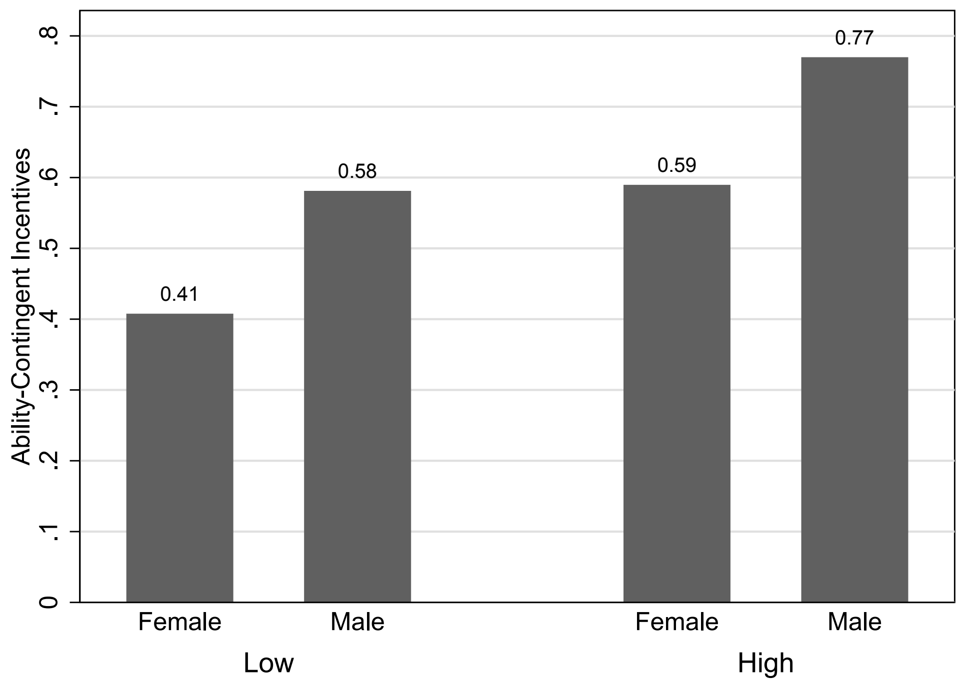
Regarding incentives choices, following their beliefs, men choose the ability-contingent incentives more often than women, despite not being more likely to be in the top half and benefit from these incentives. Figure 14 shows the choices of payment scheme separately by gender and treatment. The pattern is similar to the gender-treatment pattern observed for beliefs in 13. Table 6 shows that the average woman chooses ability-contingent incentives 50% of the time, while the average man chooses ability-contingent incentives 68% of the time ($p < 0.05$). Table 8 shows that

[20] If we test for the gender-confidence gap within each treatment separately, the ttest has a p-vale of 0.054 for the HIGH treatment and 0.29 for the LOW treatment, but our sample size is insufficient for robust analysis at this level of disaggregation.

men choose ability-contingent incentives more often in every individual round, with a significant difference in six of the nine rounds. The table is suggestive of a larger gender gap in choices for later rounds.

More specifically, there appear to be some men who are very unwilling to switch away from the ability-contingent incentives to the certain piece rate incentives. A striking illustration of this is that even in round 10 when the certain piece rate incentive piece rate is €0.8, approximately 50% of men prefer to gamble on being in the top half and getting a piece rate of €1, and earning €0 if they are wrong.

Figure 14: Propensity to Choose Ability-Contingent Incentives



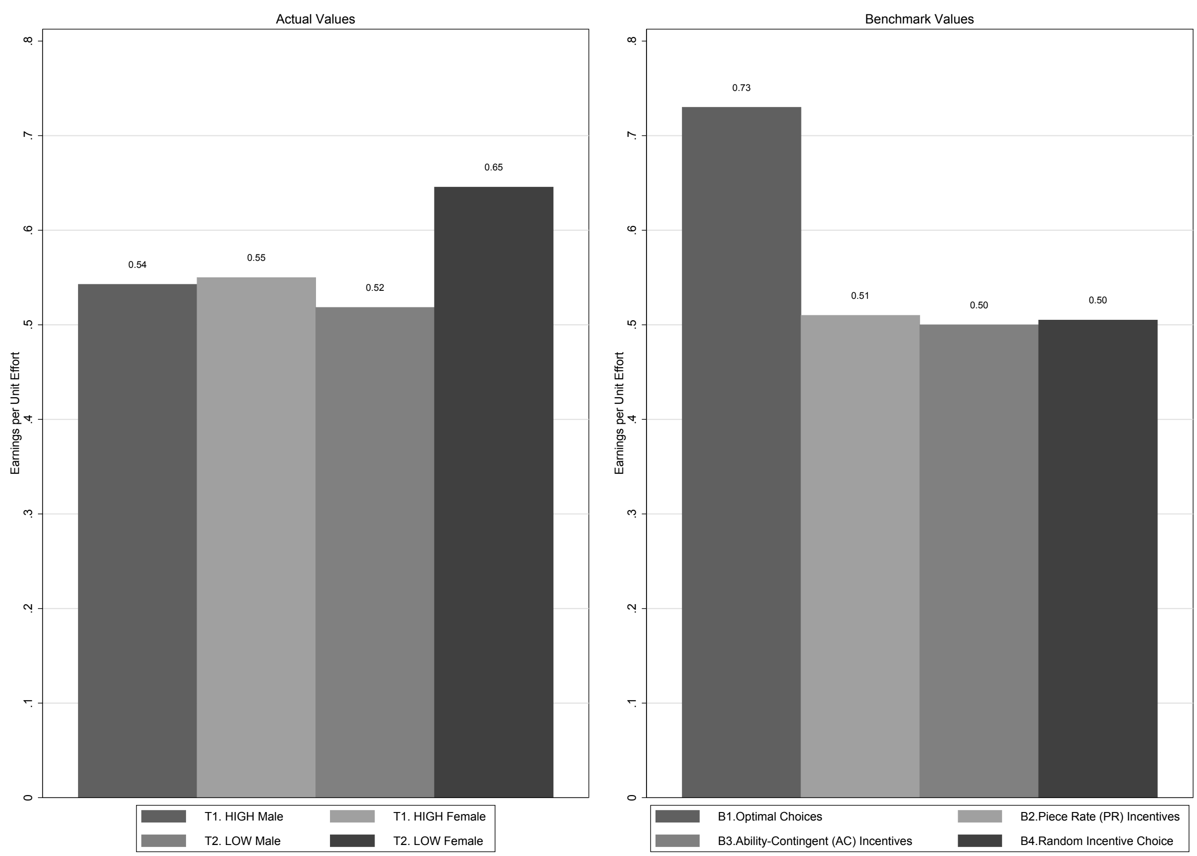
5.3 Gender differences in earnings

In section 4.4 we showed that an increase in average confidence had no significant effect on average earnings, but hurts the low-ability individuals. In this section, we consider whether the gender confidence gap translates into gender differences in earnings. Table 6 shows that on average, there is no significant difference between the earnings of men (€6.95) and women (€7.10)[21].

[21] Figure 19 in the Appendix reports earnings in each treatment-gender group separately. The differences are not significant.

However, Table 6 also shows that men are exerting significantly higher effort than women ($p < 0.01$). Since earnings are determined by both effort and incentive choices, we need to consider the contributions of effort and choices to earnings separately to understand the mechanisms. To do this, we remove the role played by effort choices by constructing a variable that stands for the *earnings per unit of effort*. Essentially, at the individual level, we calculate how much an individual earns for each set of 20 sliders she completes. This variable allows us to measure how optimal the incentive choice decisions are given an individual's level of effort.

Figure 15: Earnings per unit effort, by gender and treatment



The left panel of Figure 15 reports the average value of this "earnings per unit of effort" variable for each gender-treatment group. The interesting result here is that women in the LOW confidence treatment group are choosing the correct incentives far more often than any other group. The intrinsically lower beliefs of women together with the exogenously triggered lower beliefs by the difficult task lead to more efficient choices by women in this group. To provide a benchmark for how well these women do, in the right-hand panel we plot four benchmark earnings per unit effort possibilities: (1) average earnings for a group who all choose completely optimally, (2) average earnings for a group who all always choose the certain piece rate incentives, (3) average earnings for a group who all always choose the ability-contingent incentives, (4) average earnings for a group who always choose randomly between incentive schemes.

Except women in the LOW treatment, participants are on average performing rather poorly in terms of incentive choices, scoring between €0.52 and €0.55 per unit of effort. This amount is only a little more than they would earn if they chose completely randomly (€0.50). However, women in the LOW treatment earn €0.65 per unit effort, which is closer to the first best value of €0.73 than random choice.

5.4 Gender summary

While the starting level of confidence between men and women was different with women having a lower average confidence, the treatment had the same effect on both genders. In our results, there is no evidence that risk aversion had a mediating effect for either gender. This result adds to the evidence that gender differences in payment scheme choices are significantly affected by beliefs about relative ability rather than just competitive preferences or risk aversion (van Veldhuizen, 2017). Since it is not a competition, anxiety, fear or thrill should not play a role during the real effort task. There are no externalities imposed on the other participants when an individual exerts a high effort, so other regarding preferences are not relevant here. Women chose the ability-contingent piece rate when they believe that they are in the upper half of the group. In our design, the most successful subjects were the on average less confident women because they were better at selecting into the top and bottom group according to their actual ability.

6 Discussion and Policy Implications

One needs to be careful when extrapolating laboratory results to real world settings. Nevertheless, sometimes the stylized nature of the experiment allows for an analysis of mechanisms that cannot be disentangled in observational data and which can generate new ways of looking at economic phenomena.

So how do our findings translate to the world outside of the laboratory? While our illustrative model and experiment adopt a slightly extreme assumption of ability and effort being perfect complements, there are many professions in which low ability can only be compensated for by increased effort to a small extent. Especially in creative domains such as the arts, and writing or even research, quantity cannot compensate for quality. The programmer, Thandi, in our initial example might spend hours writing code without any of it being useful. Similarly, many entrepreneurs spend years working hard and do not succeed in their endeavors. In the case of start-ups, "ability" can be understood as the business idea and the quality of the product or the service. Entrepreneurs might have overconfident beliefs about their product compared to others. An analysis of failed start-ups showed that it was not the lack of passion or perseverance that caused the failure, but for 42 percent of analyzed start-ups there was no market for the offered product (Griffith, 2014). A lack of demand can usually not be solved by higher amounts of effort. Being overconfident in one's idea or own ability will then waste resources and motivation that could be better employed towards a different endeavor.

If beliefs about relative ability are so important for labor market choices, it is relevant to ask when and where these beliefs develop. For many, the first job out of college is a decisive step for their future career. Therefore, the beliefs about their relative ability created in college will be the basis on which graduates make their labor market entry decisions. One would think that universities should have an interest in generating accurate beliefs in their students. Nevertheless, the educational sector seems to favor promoting students' general confidence over accurate beliefs. The past two decades have seen an enormous grade inflation both in the US and many European countries (Rosovsky and Hartley, 2002). According to the Higher Education Statistics Agency (HESA, 2017), in 1994, only 7 percent of all students received a first class degree in the UK. In 2016, it is now more common to receive a first class degree than a lower second (24 percent vs. 21 percent). Some universities award first class degrees to more than one-third of their graduates.[22] There is evidence that grade inflation is even higher in the creative fields such as music, where ability and talent are even more important. 64 percent of students receive firsts at the Royal Academy of Music in the UK. This grading system might make students feel good about themselves as in the case of our HIGH confidence treatment, but it does not provide precise feedback of where they stand in the ability distribution. More precise feedback could benefit both students at the top because the signaling value of the grades increases, and the students further down, because it

[22] Interestingly, the high fraction of first class degrees is especially prevalent in the higher ranked schools such as Imperial College London (41.8 percent of first degrees) and University College London (35.6 percent). In 2013, the most frequently awarded grade in both Harvard and Princeton was a straight A with the median being an A- (Clarida and Fandos, 2013).

provides more useful feedback on a student's skill in a particular subject compared to her peers.[23]

In general, it is often difficult to obtain accurate and precise feedback about sensitive attributes in everyday life. One reason for this is perhaps that social stigma against providing negative feedback about sensitive attributes, implying a coarsening of the feedback actually provided by parents, friends and colleagues. In particular, Gneezy et al. (2017) show that individuals are usually unwilling to give negative open feedback. However, without honest feedback, there is a risk that those individuals, who are by definition below average in a particular ability, will hold inflated beliefs about themselves and waste precious time and effort. In our experiment, ability is one-dimensional, but in the real world, these individuals have other skills that they could focus on instead.

The feedback that individuals receive either through means of formal channels, such as grading and work evaluations or through informal channels, such as communication from friends and family, is very important in shaping an individual's confidence level. This experiment, as well as evidence from other studies discussed in the introduction, show that confidence is not a static trait, but can be affected by something as simple as the level of difficulty of a task. Remember that in our experiment the only difference between the two conditions were four slightly harder puzzles out of 12 in the LOW confidence treatment that resulted in the significant shift in beliefs for the below average ability individuals. This feature implies that real feedback is likely to have a far larger influence on an individual's confidence and decision making. Therefore, while there may certainly be substantial benefits to building up confidence, this paper then serves as a cautionary tale regarding some of the potential limits to the benefits of building confidence.

7 Conclusion

How important is confidence for success? How malleable are people's beliefs about their relative ability? How does a shift in beliefs affect people's choice of ability-contingent payment schemes and the level of effort they put in the task?

We present a novel experimental design in which we exogenously shift subjects' beliefs in

[23] It is worth pointing out that there are also cases where the reverse is true, namely that individuals make poor decisions due to underconfidence. For example, students sometimes select out of difficult subjects too quickly because they are underconfident. STEM subjects see very high drop-out rates early on, especially among women and minority students (Arcidiacono et al., 2016). If these students neglect to take into account that what they experience as difficult is also difficult for their peers, then those students will give up too quickly, assuming that they are worse than average. Therefore, in this case, the same lack of well-calibrated beliefs about one's abilities leads to suboptimal decision making - only here it is due to underconfidence. These students, too, would benefit from more precise feedback about their ability.

their relative ability by exposing them to a slightly easier or a slightly harder version of the same test. We then let them choose an incentive scheme under which they will work on a real effort task. They can either choose to work on an ability-contingent payment scheme which pays a high amount if they are in the top half of their group in the ability test and a low amount if they are in the bottom half; or they can choose to work at a fixed piece rate. We develop a theoretical framework that separates the ability and the effort component in the production function and use this as input for our experimental design. This feature allows us to analyze the effect of a shift in beliefs on incentive choices and effort separately.

We find that exposing subjects to an easy test, compared to the hard test increases the average beliefs about how likely they are in the top half of their group. This treatment leads to more subjects choosing the ability-contingent incentives in the easy test condition. The average level of effort is, however, the same in both conditions and is only slightly affected by incentives. This is most likely due to a time constraint imposed by the experimental set-up. The overconfidence is a detriment for the low ability individuals, who are most affected by the treatment and who wrongly self-select into the ability-contingent payment scheme, causing an even larger inequality in earnings between both types.

One implication is that an intervention meant to increase confidence might hurt exactly those people it is intended to motivate. This holds for tasks in which ability and effort have a high degree of complementarity in the production function. Our results show that women have lower pre-treatment levels of confidence than men, but both genders react the same to the treatment. In our experiment, men's higher beliefs lead them to make poorer decisions than women on average.

While we need to be careful with extrapolating our results to decision making in the real world, our results illustrate that it is relatively easy to affect people's confidence with a slight change in the level of difficulty of a task and that there are settings in which inflating people's confidence can have negative implications. Thus, a student facing a challenging natural science degree might benefit from a few easier tests here and there to keep her confidence up, but a music student in a program where more than half of the class graduated with the highest honors might be better off in the long-run with some stricter grading and more informative feedback about his relative abilities.

References

Arcidiacono, P., E. M. Aucejo, and V. J. Hotz (2016, March). University differences in the graduation of minorities in stem fields: Evidence from california. *American Economic Review 106*(3), 525–62.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Bénabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics 117*(3), 871–915.

Benoît, J.-P., J. Dubra, and D. A. Moore (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association 13*(2), 293–329.

Bernardo, A. E. and I. Welch (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy 10*(3), 301–330.

Bertrand, M. (2011). New perspectives on gender. *Handbook of Labor Economics 4*, 1543–1590.

Bonin, H., T. Dohmen, A. Falk, D. Huffman, and U. Sunde (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics 14*(6), 926–937.

Bordley, R., M. LiCalzi, and L. Tibiletti (2014). A target-based foundation for the hard-easy effect bias. Working Papers 23, Department of Management, UniversitÃ Ca' Foscari Venezia.

Bruhin, A., L. Santos-Pinto, D. Staubli, et al. (2016). How do beliefs about skill affect risky decisions? Technical report.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2013). Overconfidence and social signalling. *Review of Economic Studies 80*(3), 949–983.

Burson, K. A., R. P. Larrick, and J. Klayman (2006). Skilled or unskilled, but still unaware of it how perceptions of difficulty drive miscalibration in relative comparisons. *Journal of personality and social psychology 90*(1), 60.

Cheung, S. L. and L. Johnstone (2017, February). True Overconfidence, Revealed through Actions: An Experiment. IZA Discussion Papers 10545, Institute for the Study of Labor (IZA).

Clarida, M. and N. Fandos (2013). Substantiating fears of grade inflation, dean says median grade at harvard college is a-, most common grade is a. Technical report.

Corgnet, B., R. Hernán-González, and E. Schniter (2015, Jun). Why real leisure really matters: incentive effects on real effort in the laboratory. *Experimental Economics 18*(2), 284–301.

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

Costa-Gomes, M. A. and G. Weizsäcker (2008). Stated beliefs and play in normal-form games. *The Review of Economic Studies 75*(3), 729–762.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–74.

Dargnies, M.-P., R. Hakimov, and D. Kübler (2016). Self-confidence and unraveling in matching markets. *WZB Working Paper Series*.

de Araujo, F. A., E. Carbone, L. Conell-Price, M. W. Dunietz, A. Jaroszewicz, R. Landsman, D. Lamé, L. Vesterlund, S. Wang, and A. J. Wilson (2015). The effect of incentives of real effort: Evidence from the slider task.

DellaVigna, S. and D. Pope (2016). What motivates effort? evidence and expert forecasts. *The Review of Economic Studies*.

Eckel, C. C. and P. J. Grossman (2008). Chapter 113 men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results 1*, 1061 – 1073.

Erkal, N., L. Gangadharan, and B. Han Koh (2016). What motivates participants in real-effort tasks: Is it the journey or the destination? *Working Paper*.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2016, February). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. Working Papers 2016-003, Human Capital and Economic Opportunity Working Group.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Gächter, S., L. Huang, and M. Sefton (2016). Combining "real effort" with induced effort costs: the ball-catching task. *Experimental economics 19*(4), 687–712.

Gill, D. and V. Prowse (2011). A novel computerized real effort task based on sliders. IZA Discussion Papers 5801, Institute for the Study of Labor (IZA).

Gneezy, U., C. Gravert, S. Saccardo, and F. Tausch (2017). A must lie situation - avoiding giving negative feedback. *Games and Economic Behavior 102*, 445 – 454.

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

Griffith, E. (2014). Why startups fail, according to their founders. Technical report.

Healy, P. J. and D. A. Moore (2007). Bayesian overconfidence.

Heifetz, A., C. Shannon, and Y. Spiegel (2007). The dynamic evolution of preferences. *Economic Theory 32*(2), 251.

HESA, H. E. S. A. (2017). Qualifications obtained.

Johnson, D. D. and J. H. Fowler (2011). The evolution of overconfidence. *Nature 477*(7364), 317–320.

Kruger, J. (1999). Lake wobegon be gone! the below-average effec and the egocentric nature of comparative ability judgments. *Journal of personality and social psychology 77*(2), 221.

Kruger, J. and D. Dunning (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology 77*(6), 1121.

Larrick, R. P., K. A. Burson, and J. B. Soll (2007). Social comparison and confidence: When thinking you are better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes 102*(1), 76 – 94. Special Issue on Social Comparison Processes.

Malmendier, U. and G. Tate (2005). Ceo overconfidence and corporate investment. *The journal of finance 60*(6), 2661–2700.

Moore, D. A. and P. J. Healy (2008). The trouble with overconfidence. *Psychological review 115*(2), 502.

Moore, D. A. and T. G. Kim (2003). Myopic social prediction and the solo comparison effect. *Journal of personality and social psychology 85*(6), 1121.

Moore, D. A. and D. A. Small (2007). Error and bias in comparative judgment: on being both better and worse than we think we are. *Journal of personality and social psychology 92*(6), 972.

Murad, Z., M. Sefton, and C. Starmer (2016, Feb). How do risk attitudes affect measured confidence? *Journal of Risk and Uncertainty 52*(1), 21–46.

Niederle, M. (2017). A gender agenda: A progress report on competitiveness. *American Economic Review 107*(5), 115–19.

Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? do men compete too much? *The Quaterly Journal of Economics*, 1067–1101.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of finance 53*(5), 1775–1798.

Pikulina, E., L. Renneboog, and P. Tobler (2016). Do confident individuals generally work harder?

Pikulina, E., L. Renneboog, and P. N. Tobler (2017). Overconfidence and investment: An experimental approach. *Journal of Corporate Finance 43*, 175–192.

Puri, M. and D. T. Robinson (2007). Optimism and economic choice. *Journal of Financial Economics 86*(1), 71 – 99.

Rosovsky, H. and M. Hartley (2002). Evaluation and the academy: Are we doing the right thing.

Schwardmann, P. and J. J. Van der Weele (2016). Deception and self-deception. *Working Paper*.

Selten, R. (1998). Axiomatic characterization of the quadratic scoring rule. *Experimental Economics 1*(1), 43–62.

Smith, A. (2013, Sep). Estimating the causal effect of beliefs on contributions in repeated public good games. *Experimental Economics 16*(3), 414–425.

van Veldhuizen, R. (2017). Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness? Technical report, Discussion Paper.

A Appendices

A.1 Theoretical Framework with Risk Aversion

In the main text, we used a simple theoretical framework to organize thinking about what we might expect from the experimental data. In that discussion, we assumed that individuals are risk neutral. However most of the intuitions carry through when we relax the risk neutrality requirement. Below, we show that a similar logic applies if we allow for risk aversion over monetary payoffs. More specifically, consider an individual who chooses w and e to maximise:

$$\max_{w \in \{\bar{w}, w(a)\}} \max_{e \geq 0} \mathbf{U}(w, e) = s \cdot e + \mathbf{E}_a[u(w \cdot e)] - c(e)$$

where $u'(\cdot) > 0$, $u''(\cdot) < 0$ and $u'(0) + s > c'(0)$. First, we consider the individuals choice between the two incentive schemes.

Choice of Incentive Scheme

As in the main text, let the optimal effort levels under the ability-contingent and certain piece rate incentives be denotes by e^{**}_{AC} and e^{**}_{PR} (assuming an interior solution). Given these optimal effort choices under the two incentive schemes, we have the following condition for the individual choosing the ability-contingent incentives. The individual prefers ability-contingent incentives if:

$$s \cdot e^{**}_{AC} + \hat{\pi} \cdot u(w_H \cdot e^{**}_{AC}) - c(e^{**}_{AC}) \geq s \cdot e^{**}_{PR} + u(\bar{w} \cdot e^{**}_{PR}) - c(e^{**}_{PR})$$

This inequality implies that there is a threshold subjective belief, $\pi^{i,r}$, such that the individual prefers to choose the certain piece rate incentives for lower subjective beliefs, i.e. $\hat{\pi} \leq \pi^{i,r}$, and prefers to choose ability-contingent incentives for larger subjective beliefs, i.e. $\hat{\pi} \geq \pi^{i,r}$. In general, we can specify $\pi^{i,r}$ as follows:

$$\pi^{i,r} := \frac{u(\bar{w} \cdot e^{**}_{PR})}{u(w_H \cdot e^{**}_{AC})} + \frac{s \cdot (e^{**}_{PR} - e^{**}_{AC}) - (c(e^{**}_{PR}) - c(e^{**}_{AC}))}{u(w_H \cdot e^{**}_{AC})}$$

Notice that if there is a binding constraint on effort in the lab, and optimal effort under both incentive schemes is given by $\bar{e}^{**}$, this expressions simplifies to:

$$\pi^{i,r} = \frac{u(\bar{w} \cdot \bar{e}^{**})}{u(w_H \cdot \bar{e}^{**})}$$

The discussion above can be summarised with the following decision rules for incentive choices:

i. The individual chooses the certain piece rate incentives if she has low self-confidence (i.e. beliefs on the interval: $\hat{\pi} \in [0, \pi^{i,r}]$).

ii. The individual chooses the risky ability-contingent incentives if she has high self-confidence (i.e. beliefs on the interval: $\hat{\pi} \in [\pi^{i,r}, 1]$).

Choice of Effort

In the main text, under risk neutrality, the threshold belief at which the individual is indifferent between choosing the ability-contingent and certain piece rate incentives was the same as the threshold at which the individual would choose the same level of effort under both incentives schemes. With risk aversion, these two thresholds for (i) incentive scheme schoice, and (ii) effort choice, are no longer the same.

In this section, we first consider the case where there is an internal solution for the optimal effort level. Later, we return to the case where effort is constrained in the lab (e.g. due to time constraints). Effort choices under the two incentive schemes are obtained by solving the following two equations (assuming interior solutions):

$$\hat{\pi} w_H \cdot u'(w_H \cdot e) = c'(e) - s \text{ gives } e^{**}_{AC} \tag{6}$$

$$\bar{w} \cdot u'(\bar{w} \cdot e) = c'(e) - s \text{ gives } e^{**}_{PR} \tag{7}$$

It is clear that in general e^{**}_{AC} may be either larger or smaller than e^{**}_{PR}, depending on the individual's subjective belief about her own ability, $\hat{\pi}$. In particular, when $\hat{\pi} = 0$, the individual believes that she is low ability for sure, and she chooses more effort under the piece rate incentives ($e^{**}_{AC} < e^{**}_{PR}$). The reverse is true when she believes that she is the high ability type for sure ($\hat{\pi} = 1 \Rightarrow e^{**}_{AC} > e^{**}_{PR}$).

However, since individuals are given a choice of which incentive scheme to choose, it is of interest to know whether all individuals who choose the ability-contingent incentives choose a higher effort level than the individuals who choose the fixed piece rate.

In order to answer this question, firstly, notice that the RHS of equations 6 and 7 is identical. Since the LHS of each equation is decreasing in e, considering the ratio of the LHS of the two expressions, $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)}$, can help us to assess which optimal effort choice will be higher.

Secondly, since $w_H > \bar{w}$ and $u''(\cdot) < 0$, the ratio $\frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)}$ is increasing as effort increases. This implies that $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)}$ is also increasing in e. One can think of this term as the ratio of the marginal shift in utility due to money of an additional unit of effort, under the two incentive schemes. For ease of exposition, we will refer to the "monetary component" of the utility function as the part that excludes the intrinsic reward for effort, and the cost of effort (i.e. $\mathbf{E}_a[u(w \cdot e)]$).

Thirdly, the optimal effort level under thecertain piece rate incentives doesn't depend on the individual's subjective belief about her ability - all individuals choose the same optimal effort level under certain piece rate, e^{**}_{PR}, irrespective of their belief, $\hat{\pi}$. This means that irrespective of $\hat{\pi}$, we can evaluate the ratio of marginal utilities of money, $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)}$, at the point of optimal effort under the certain piece rate incentives, e^{**}_{PR} (which doesn't depend on $\hat{\pi}$) to determine the threshold value of $\hat{\pi}$ for which the optimal effort choice switches from being higher under certain piece rate incentives to being higher under ability-contingent incentives.

Define $\pi^{e,r}$ to be the value of $\hat{\pi}$ at which the ratio of marginal utilities of money, $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)}$, evaluated at e^{**}_{PR}, is equal to 1. In particular:

$$\pi^{e,r} := \frac{\bar{w}}{w_H} \cdot \frac{u'(\bar{w} \cdot e^{**}_{PR})}{u'(w_H \cdot e^{**}_{PR})} \tag{8}$$

With these definitions for $\pi^{i,r}$ and $\pi^{e,r}$ in hand, we can assess the relationship between the individual's choice of incentive scheme and her effort choice.

Proposition 1 *If the individual chooses the ability-contingent incentives, then she exerts more effort than she would have if she had faced the fixed piece rate incentives. In particular,*

$$\hat{\pi} \geq \pi^{i,r} \Rightarrow \hat{\pi} \geq \pi^{e,r}$$

If she chooses the piece rate incentives, then she chooses more effort under certain piece rate incentives if $\hat{\pi} \leq \pi^{e,r}$, and more effort under ability-contingent incentives if $\hat{\pi} > \pi^{e,r}$.

Essentially, we show that there are three possible intervals for $\hat{\pi}$ that specify incentive choice and effort choice behaviour:

- If $\hat{\pi} \in [0, \pi^{e,r}]$, the individual chooses the certain piece rate incentives and exerts more effort under certain piece rate incentives than ability-contingent incentives.
- If $\hat{\pi} \in (\pi^{e,r}, \pi^{i,r})$, the individual would prefer the certain piece rate incentives, but if she

faced the ability-contingent incentives, then she would exert more effort than under the certain piece rate incentives.

- If $\hat{\pi} \in [\pi^{i,r}, 1]$, the individual chooses the ability-contingent incentives and exerts more effort under ability-contingent incentives than certain piece rate incentives.

Case 1: First we consider $\hat{\pi} \in [0, \pi^{e,r})$, or equivalently, $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e^{**}_{PR})}{u'(\bar{w} \cdot e^{**}_{PR})} < 1$. This means that the monetary marginal utility of effort is lower under ability-contingent incentives than under certain piece rate incentives at $e = e^{**}_{PR}$. Therefore, since $\bar{w} \cdot u'(\bar{w} \cdot e^{**}_{PR}) = c'(e^{**}_{PR}) - s$, we have:

$$\hat{\pi} w_H \cdot u'(w_H \cdot e^{**}_{PR}) < \bar{w} \cdot u'(\bar{w} \cdot e^{**}_{PR}) = c'(e^{**}_{PR}) - s \tag{9}$$

But since $u''(\cdot) < 0$ and $c''(\cdot) > 0$, there cannot be an $e > e^{**}_{PR}$ such that $\hat{\pi} w_H \cdot u'(w_H \cdot e) = c'(e) - s$, since the left hand side of equation 6 is decreasing in e and the right hand side is increasing in e. This implies that when $\hat{\pi} \in [0, \pi^{e,r})$, we must have $e^{**}_{AC} < e^{**}_{PR}$. Essentially, this says that if the monetary marginal utility of effort is larger under PR, evaluated at $e = e^{**}_{PR}$, then the individual will choose a higher effort level under certain piece rate incentives than under ability-contingent incentives. Notice, also, that since this ratio of monetary MUs is increasing in e, $\forall e \leq e^{**}_{PR}$, we must have: $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e)}{u'(\bar{w} \cdot e)} < 1$. Therefore, the monetary marginal utility $\forall e \leq e^{**}_{PR}$ is larger under the certain piece rate incentives than ability-contingent incentives, and the non-monetary component of the utility function is identical. Therefore, integrating the marginal utilities over e implies that $\forall e \leq e^{**}_{PR}$:

$$s \cdot e + \hat{\pi} \cdot u(w_H \cdot e) - c(e) \leq s \cdot e + u(\bar{w} \cdot e) - c(e)$$

and since $e = e^{**}_{PR}$ maximises the RHS, and $e^{**}_{AC} < e^{**}_{PR}$ it must be the case that:

$$s \cdot e^{**}_{AC} + \hat{\pi} \cdot u(w_H \cdot e^{**}_{AC}) - c(e^{**}_{AC}) < s \cdot e^{**}_{AC} + u(\bar{w} \cdot e^{**}_{AC}) - c(e^{**}_{AC}) \leq s \cdot e^{**}_{PR} + u(\bar{w} \cdot e^{**}_{PR}) - c(e^{**}_{PR})$$

In summary, if $\hat{\pi} \in [0, \pi^{e,r})$, then: (i) $e^{**}_{AC} < e^{**}_{PR}$, and (ii) the individual chooses the certain piece rate incentives (i.e. $\hat{\pi} \in [0, \pi^{i,r})$).

Case 2: Second, we consider $\hat{\pi} \in [\pi^{e,r}, 1]$, or equivalently, $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e^{**}_{PR})}{u'(\bar{w} \cdot e^{**}_{PR})} \geq 1$. Now, the monetary marginal utility of effort is higher under ability-contingent incentives than under certain piece rate incentives at $e = e^{**}_{PR}$. This implies that $\hat{\pi} w_H \cdot u'(w_H \cdot e^{**}_{PR}) > \bar{w} \cdot u'(\bar{w} \cdot e^{**}_{PR}) = c'(e^{**}_{PR}) - s$, which means that the overall marginal utility of effort under the ability-contingent incentives is positive at $e = e^{**}_{PR}$ (i.e. $\frac{\partial \mathbf{U}_{AC}(w,e)}{\partial e}\Big|_{e=e^{**}_{PR}} > 0$. Therefore, the optimal effort level under the ability-contingent incentives is higher than under the piece rate incentives $e^{**}_{AC} > e^{**}_{PR}$. However, it is important to note that $\frac{\hat{\pi} w_H}{\bar{w}} \cdot \frac{u'(w_H \cdot e^{**}_{PR})}{u'(\bar{w} \cdot e^{**}_{PR})} \geq 1$ does not necessarily imply that the individual chooses the ability-contingent incentives. In Case 2, we can have either choice of incentives.

In summary, if $\hat{\pi} \in [\pi^{e,r}, 1]$, then: (i) $e^{**}_{AC} \geq e^{**}_{PR}$, and (ii) the individual chooses the ability-contingent incentives if $\hat{\pi} \in [\pi^{i,r}, 1]$.

Together, Case 1 and Case 2 show that for the interval of subjective beliefs $\hat{\pi} \in [0, \pi^{e,r}]$, the individual would choose certain piece rate incentives, and would choose higher effort under certain piece rate incentives than she would under ability-contingent incentives.

For the interval $\hat{\pi} \in [\pi^{i,r}, 1]$, the individual would choose the ability-contingent incentives, and would choose higher effort under ability-contingent incentives than she would under certain piece rate incentives. And for the interval $\hat{\pi} \in (\pi^{e,r}, \pi^{i,r})$, the individual prefers the certain piece rate incentives, but would choose higher effort if she were to face the ability-contingent incentives.

As a caveat, the discussion above refers to interior solutions for effort choices. Of course, if effort choices in the lab face a binding constraint (e.g. the time limit), then the discussion above does not apply, and effort choices are the same under the two incentives schemes.

Influence on Earnings

The discussion in the main text for the risk-neutral agent essentially maps directly to the case of the risk-averse agent, replacing: (i) π^i with $\pi^{i,r}$, (ii) e^*_{PR} with e^{**}_{PR}, and (iii) e^*_{AC} with e^{**}_{AC}. All the main intuitions remain the same. Therefore, we have:

<u>Gain in Earnings for High Ability Individuals Due to Upward Shift in Confidence</u>

$$[F_H(\pi^{i,r}) - F_H(\pi^{i,r} - \triangle\hat{\pi})] \cdot (w_H \cdot e^{**}_{AC} - \bar{w} \cdot e^{**}_{PR}) \geq 0 \tag{10}$$

<u>Loss of Earnings for Low Ability Individuals Due to Downward Shift in Confidence</u>

$$[F_L(\pi^{i,r}) - F_L(\pi^{i,r} - \triangle\hat{\pi})] \cdot (0 - \bar{w} \cdot e^{**}_{PR}) \leq 0 \tag{11}$$

.

A.2 Additional Tables

Table 5: Comparison of means by treatment

	LOW	HIGH
Male (=1)	0.48	0.54
	(0.51)	(0.50)
Age	25.20	24.62
	(6.55)	(3.88)
Effort (Baseline Task, per min)	9.55	9.91
	(1.80)	(1.70)
Risk (CE p=50)	0.48	0.49
	(0.16)	(0.14)
Treatment Variables		
Treatment (High =1)	0	1
	(0.00)	(0.00)
Ability Score	6.92	10.90***
	(2.79)	(1.20)
Ability Top Half (=1)	0.50	0.50
	(0.51)	(0.51)
Outcome Variables		
Subj Belief (%)	64.20	83.92***
	(31.67)	(20.16)
AC Incentive Choice (Frac)	0.50	0.69***
	(0.37)	(0.32)
Effort (Main Task, per min)	12.15	12.18
	(2.29)	(2.04)
Earnings (Main Task)	7.27	6.69
	(4.78)	(5.14)
Risk (CE p=Subj Belief)	0.56	0.71***
	(0.23)	(0.22)
N	50	50

Notes: (i) standard deviations in parentheses, t-tests:
*=10%, **=5%, ***=1%

Table 6: Comparison of means by gender

	Female	Male
Male (=1)	0	1
	(0.00)	(0.00)
Age	23.96	25.16
	(3.61)	(4.37)
Effort (Baseline Task, per min)	9.46	10.11*
	(1.62)	(1.79)
Risk (CE p=50)	0.48	0.48
	(0.16)	(0.13)
<u>Treatment Variables</u>		
Treatment (High =1)	0.49	0.55
	(0.51)	(0.50)
Ability Score	8.98	8.92
	(2.71)	(3.23)
Ability Top Half (=1)	0.51	0.49
	(0.51)	(0.51)
<u>Outcome Variables</u>		
Subj Belief (%)	67.74	79.55**
	(29.55)	(26.38)
AC Incentive Choice (Frac)	0.50	0.68**
	(0.36)	(0.34)
Effort (Main Task, per min)	11.65	12.79***
	(1.89)	(2.17)
Earnings (Main Task)	7.10	6.95
	(4.15)	(5.62)
Risk (CE p=Subj Belief)	0.56	0.70***
	(0.24)	(0.22)
N	47	49

Notes: (i) standard deviations in parentheses, t-tests: *=10%, **=5%, ***=1% (ii) The 3 missing women, and 1 missing man answered “Prefer not to answer” to the gender question in the survey.

Table 7: Comparison of mean incentive choices by treatment

	LOW	HIGH
Treatment (High =1)	0	1
	(0.00)	(0.00)
Subj Belief (%)	64.20	83.92***
	(31.67)	(20.16)
Ability Top Half (=1)	0.50	0.50
	(0.51)	(0.51)
<u>Incentive Choices</u>		
AC Incentive Choice (All Rounds)	0.50	0.69***
	(0.37)	(0.32)
AC Incentives Round 1 (=1)	1	1
	(0.00)	(0.00)
AC Incentives Round 2 (=1)	0.76	0.90*
	(0.43)	(0.30)
AC Incentives Round 3 (=1)	0.76	0.84
	(0.43)	(0.37)
AC Incentives Round 4 (=1)	0.64	0.84**
	(0.48)	(0.37)
AC Incentives Round 5 (=1)	0.52	0.74**
	(0.50)	(0.44)
AC Incentives Round 6 (=1)	0.40	0.70***
	(0.49)	(0.46)
AC Incentives Round 7 (=1)	0.36	0.58**
	(0.48)	(0.50)
AC Incentives Round 8 (=1)	0.38	0.58**
	(0.49)	(0.50)
AC Incentives Round 9 (=1)	0.38	0.54
	(0.49)	(0.50)
AC Incentives Round 10 (=1)	0.32	0.46
	(0.47)	(0.50)
N	50	50

Notes: (i) Ability Top Half: Reflects fraction of individuals in top half. Equals 0.5 by construction. (ii) AC Incentive Choice (All Rounds): individual level variable, averaged across all an individual's choices. (iii) standard deviations in parentheses, t-tests: *=10%, **=5%, ***=1%

Table 8: Comparison of mean incentive choices by gender

	Female	Male
Treatment (High =1)	0.49	0.55
	(0.51)	(0.50)
Subj Belief (%)	67.74	79.55**
	(29.55)	(26.38)
Ability Top Half (=1)	0.51	0.49
	(0.51)	(0.51)
Incentive Choices		
AC Incentive Choice (All Rounds)	0.50	0.68**
	(0.36)	(0.34)
AC Incentives Round 1 (=1)	1	1
	(0.00)	(0.00)
AC Incentives Round 2 (=1)	0.74	0.92**
	(0.44)	(0.28)
AC Incentives Round 3 (=1)	0.70	0.88**
	(0.46)	(0.33)
AC Incentives Round 4 (=1)	0.70	0.80
	(0.46)	(0.41)
AC Incentives Round 5 (=1)	0.60	0.67
	(0.50)	(0.47)
AC Incentives Round 6 (=1)	0.40	0.67***
	(0.50)	(0.47)
AC Incentives Round 7 (=1)	0.34	0.59**
	(0.48)	(0.50)
AC Incentives Round 8 (=1)	0.34	0.59**
	(0.48)	(0.50)
AC Incentives Round 9 (=1)	0.36	0.55*
	(0.49)	(0.50)
AC Incentives Round 10 (=1)	0.28	0.49**
	(0.45)	(0.51)
N	47	49

Notes: (i) Ability Top Half: Reflects fraction of individuals in top half. (ii) AC Incentive Choice (All Rounds): individual level variable, averaged across all an individual's choices. (iii) standard deviations in parentheses, t-tests: *=10%, **=5%, ***=1%

A.3 Additional Figures

Figure 16: Risk preferences

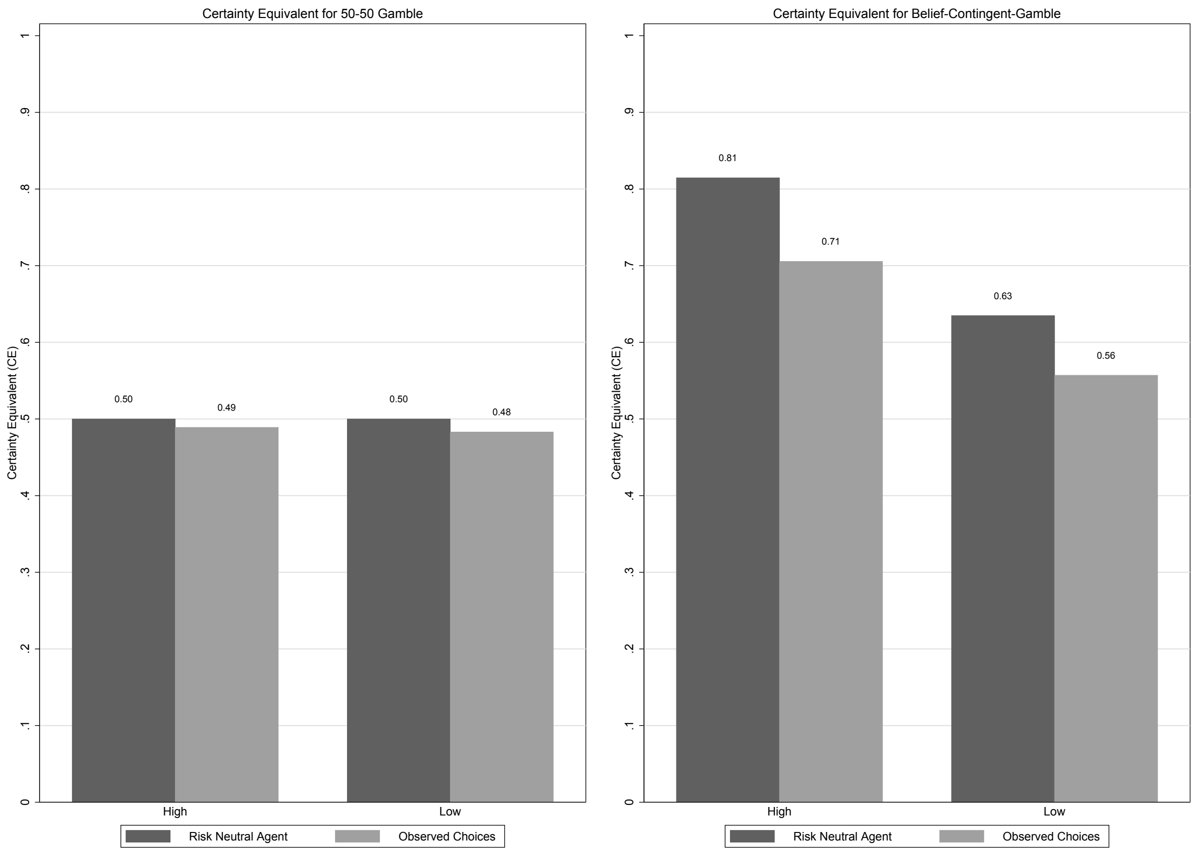


Figure 17: Distribution of Earnings between Treatments

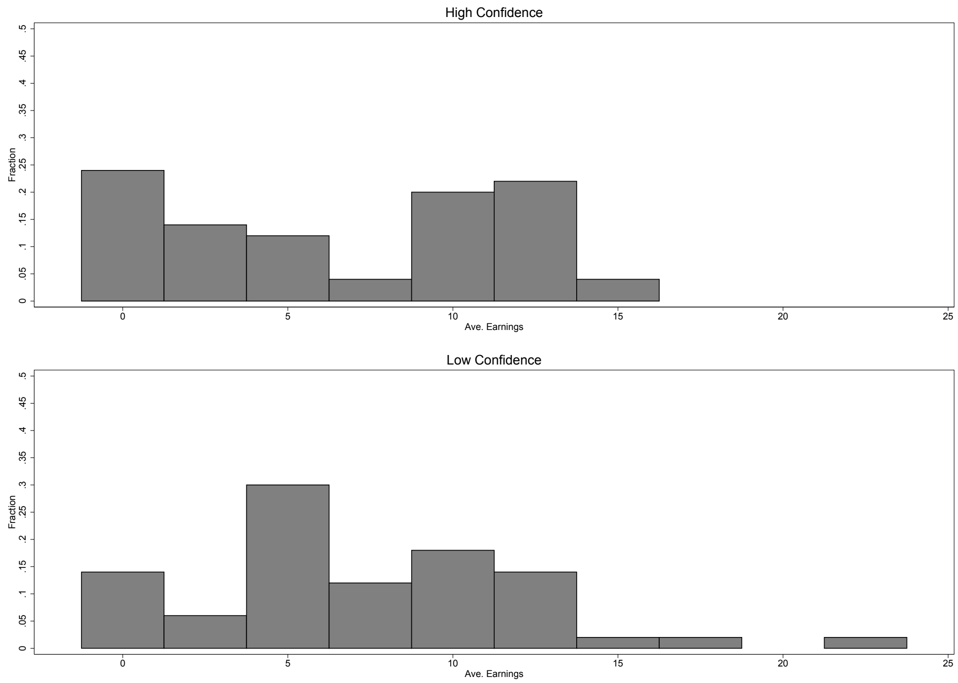


Figure 18: Effort per unit effort, by gender and treatment

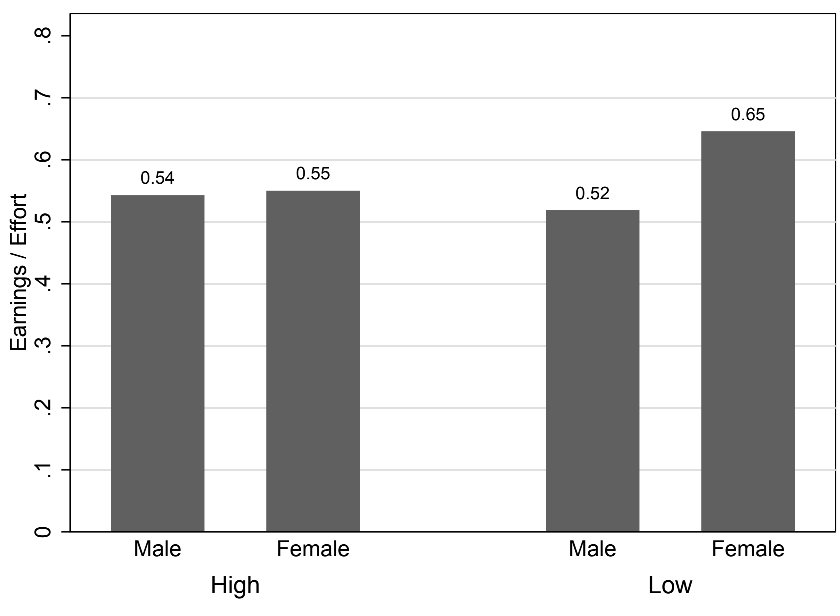


Figure 19: Average earnings by gender and treatment

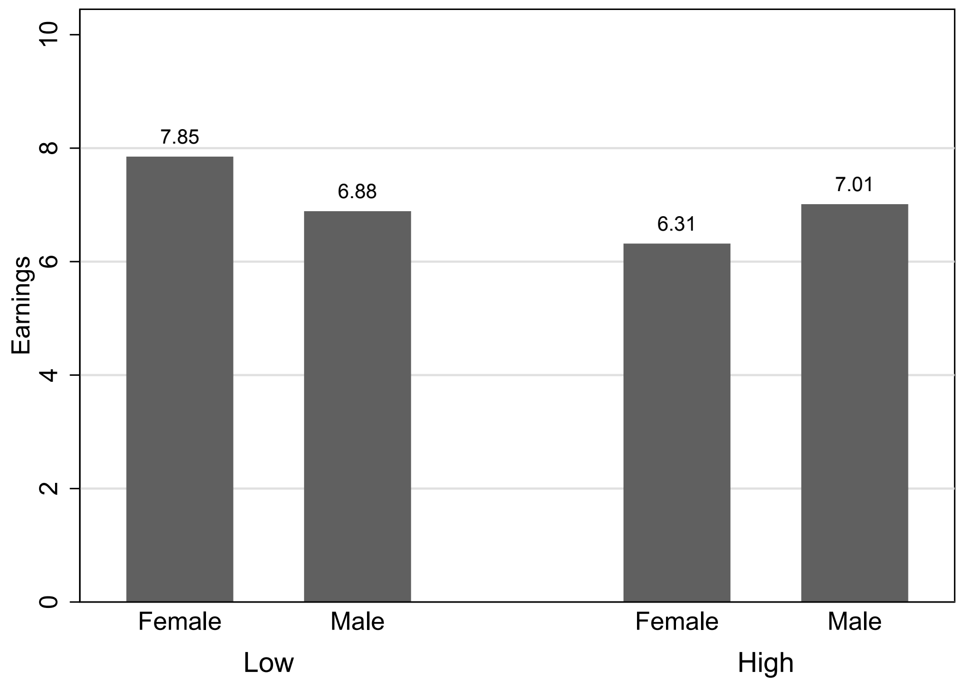


Figure 20: Propensity to Score in the Top Half

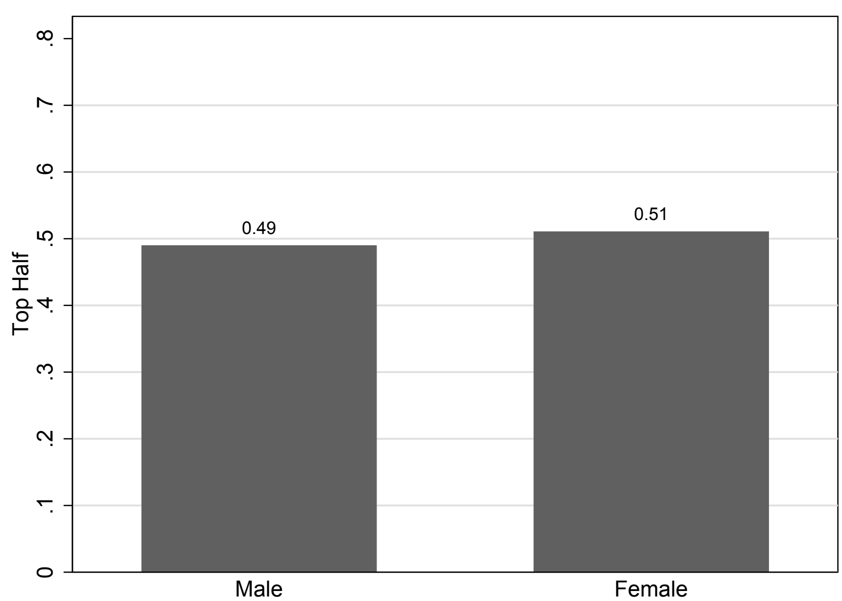


Figure 21: Average effort level by gender and treatment