5-9-2024

Essays in Behavioral and Experimental Economics

Yves-Paul Auffray
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Arts & Sciences
Department of Economics

Dissertation Examination Committee:
Brian W. Rogers, Chair
Mariagiovanna Baccara
Marcus Berliant
John Nachbar
Paulo Natenzon

Essays in Behavioral and Experimental Economics

by

Yves-Paul Auffray

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2024
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply grateful to my advisor, Brian Rogers, for his invaluable comments and support throughout the years. On multiple occasions, it was only as I reviewed my notes of our meetings that I truly understood the depth and insight of his comments. None of the projects in this thesis would have succeeded without his guidance. I greatly appreciate his enthusiasm for new projects and ideas, as well as how generous he was with his time.

I would also like to extend my gratitude to Mariagiovanna Baccara for her tremendous generosity with her time, invaluable support, and professional attitude towards work and research. She has taught me the importance of striving to clarify ideas and the significance of paying attention to details. I deeply value her honest feedback, as it has undoubtedly shaped me into a better researcher. Lastly, I want to express my thanks to Marcus Berliant for his support and willingness to keep his door open for discussions.

I also want to thank my friends in the Economics Department at Washington University, as well as the members of the research groups for cultivating a fruitful research environment. Last but not least, I want to thank my family and friends in France for their constant support.

I gratefully acknowledge that this dissertation was funded by the Weidenbaum Center on the Economy, Government, and Public Policy at Washington University in St. Louis and the Center for Research in Economics and Strategy (CRES), in the Olin Business School.

Yves-Paul Auffray

Washington University in St. Louis

May 2024

ABSTRACT OF THE DISSERTATION

Essays in Behavioral and Experimental Economics

by

Yves-Paul Auffray

Doctor of Philosophy in Economics

Department of Economics

Washington University in St. Louis, 2024

Professor Brian W. Rogers, Chair

This dissertation consists of three self-contained articles in the field of Behavioral and Experimental Economics. Chapter one and three investigate the implications of non-standard preferences, i.e., social preferences for economic decision making. Another common theme in Behavioral Economics is that of bounded rationality. In chapter two, I look at correlation neglect, which is one example of bounded rationality.

The first chapter ”Circumstances, effort choice and redistribution,” investigates the effects of differences in personal circumstances on individual effort choice and on income redistribution. I construct a model in which a worker's production depends on a combination of circumstances, effort and unknown luck. Theoretically, I find that workers with better circumstances exert less effort on average than those who fail, yet succeed more. This counter-intuitive result has implications for third parties with meritocratic fairness preferences, as they tend to reward effort and offset the effects of luck. This work also implies that heterogeneity in circumstances matters for redistribution.

The second chapter ”Correlation neglect and overconfidence: theory and experimental evidence,” explores how probabilistic reasoning affects the consumption of information, the formation of opinions, and belief updating. I investigate whether correlation neglect leads to overconfidence. I establish that individuals who neglect correlation tend to value and seek additional information more and become more overconfident. I conduct a novel experiment involving endogenous acquisition of highly correlated signals and use two incentivized methods to identify individual levels of correlation

neglect. Interestingly, ”full neglecters” are willing to pay more for additional information, and high correlation neglecters exhibit greater overprecision. These findings suggest that drawing decision-maker’s attention to the high degree of correlation in the information they consume - for instance, because their friend’s share similar news on social medias - can mitigate overconfidence. The third chapter ”Guilt aversion or social image concern? Evidence from a modified Dictator Game,” uses Game Theory to disentangle two motives for pro-social behavior: guilt aversion or social image concerns. I use a dictator game in which I vary the recipient’s payoff expectation, and whether they can observe the dictator’s action. Increasing expectations only affects guilt-averse dictators, while increased exposure prompts those concerned about their social image to behave more pro-socially. These psychological motivations may be subtle but have significant real-world implications, such as guiding government policies to reduce tax evasion or helping businesses increase employee tips. They will not implement the same policy depending on the population’s dominant motives.

Chapter 1: Circumstances, Effort Choice and Redistribution

This chapter investigates theoretically and experimentally how difference in individual circumstances affects effort choice and attitudes towards redistribution. I introduce a model in which a worker's production depends on effort, exogenous circumstances, and luck. I show that when workers aim to reach a target production, those who succeed on average are endowed with better circumstances and exert less effort. A third party who is informed only of the worker's outcome - success or failure - believe that workers who succeed on average exert less effort than workers who fail. I test the theory using a real-effort experiment in which circumstances, effort and the distribution of luck determines the likelihood of success. The experimental results suggest that effort choice responds to circumstances. Workers with better circumstances exert less effort yet succeed more. Eighty-seven percent of workers who provide no effort start with circumstances below the median. Third parties make the right inference about the role that circumstances plays in determining success or failure. Earnings redistribution within pairs of workers with asymmetric outcomes is higher when third parties are informed of worker's circumstances. The higher level of redistribution is suggestive of meritocratic behavior.

Keywords: Inequality; circumstances; opportunity; preference for fairness; redistribution; experiments

JEL Codes: C91, D83, D91

1.1 Introduction

”For faction is everywhere due to inequality, ... for some consider themselves wholly equal if they are equal in a certain respect, whereas others claim to merit an unequal share of everything if they are unequal in a certain respect.” Aristotle, *Politics*

Many authors have documented the rise in income inequality in the past fifty years in the United States, and in particular since the 1980s, as in [1]. In the 2021 General Social Survey, 73% of respondents considered that differences in income were too large in the United States[1]. Whether income inequalities are deemed fair or unfair has implications for which inequalities are accepted, and which policies citizens prefer: the level of income taxation, redistribution through social transfers in the form of unemployment benefits, health insurance, and so forth. The sources of income inequality critically affect how people redistribute income, as in [2]. Income inequality may be caused by factors that agents cannot control such as a physical disability, and factors that agents are responsible for, typically the choice of how hard to work.

The majority of people tend to redistribute earnings when inequalities are due to factors beyond agents’ control [3, 4], but are more inclined to tolerate income inequalities if they are caused by differences in effort [5]. This type of fairness preference is called meritocratic. However, uncertainty about the source of income inequality is common in real life: did a person succeed because they were hard-working or did they just got lucky?

A crucial point is that personal circumstances affect individual’s effort choice. When people know their circumstances and how likely they are to succeed, they will take this into account to determine how much effort to exert.[2] Consider the case of a high-school student who wishes obtain a score of 1500 on the SAT. They may attend a good private school which offers a rich curriculum, or they may grow up in a disadvantaged neighborhood whose public school only gives a basic

[1]Specifically, the question was: Do you agree or disagree? Differences in income in America are too large. See https://gssdataexplorer.norc.org/variables/4295/vshow. The General Social Survey is a project of the independent research organization NORC at the University of Chicago, with principal funding from the National Science Foundation.

[2]The same observation is made by [6] who note that ”*the choice to work hard is often shaped by circumstances*”.

educational program. Knowing their circumstances, the student who attend the private school can choose to study hard to maximize their chance of getting 1500, or reduce their effort and rely their acquired knowledge. The public-school student may study hard knowing that they start with unfavorable circumstances, or feel discouraged by the magnitude of the work required to reach the cutoff score.

I build a theoretical model in which production is generated by a combination of costly effort, exogenous circumstances, and luck. The agent's production is compared to a fixed and exogenous threshold. If the production reaches or exceeds the threshold, the agent succeeds and earns a fixed payoff. If the production is lower than the threshold, they fail and receive nothing.[3] Hence, the outcome is binary: success or failure. In the model, *circumstances* are observed by the agent before they choose their effort level. *Luck* is a multiplier of effort and is unknown to the agent before they choose how much effort to exert. Success is uncertain even if the agent makes infinite effort.[4] The agent thus faces a trade-off between their cost of effort and their chance of success. Exerting more effort is costlier, but it increases their chance of reaching the target. A third party who observes the agent's outcome but not their effort, circumstances, or luck, is always uncertain about their respective contribution to the agent's production.

The model has two main results. First, I show that the agent's optimal effort is decreasing in circumstances. Intuitively, being endowed with favorable circumstances increases the agent's chance of success, which in turn decreases effort. Turning to inferences, I show that the expected circumstances conditional on success is higher than conditional on failure. Therefore, if a third party observes only the agent's outcome, they expect those who succeed to have received, on average, better circumstances than those who fail. Because effort is decreasing in circumstances, it follows that the expected effort of those who succeed is lower than those who fail. This is a counterintuitive result which has implications for redistribution. A third party with meritocrat preference might

[3]In this paper, I use interchangeably the words threshold and target to describe the target production required to succeed.

[4]The agent's circumstances can never be greater than the threshold, and because there is always a chance to receive a luck of zero, applying infinite effort does not guarantee success.

be tempted to reward success as it is normally associated with higher effort. However, in this environment success is a signal of lower effort. I run a real-effort experiment to replicate the theoretical set up with the goal of answering three main questions:

(i) How do agents adjust their level of effort to various circumstances?

(ii) Are third parties able to draw the right inferences about why people succeed or fail? Do third parties attribute success (failure) to better (worse) circumstances or hard work (lack of effort)?

(iii) How do those inferences and heterogeneity in circumstances affect third parties' inclination to redistribute income?

In the experiment, participants complete multiple rounds of a modified version of the slider task introduced by [7, 8].[5] In each round, before doing the task, participants are told their circumstances value and the distribution of luck. Then, they perform the slider task. When they decide to stop working on the task, they learn their luck draw and their total production is computed as follows: we multiply their effort - the number of sliders correctly positioned - by their luck, then we add their circumstances. If their production is greater than a threshold, they succeed and earn $1. Otherwise, they earn nothing. In the baseline *Identical-Low* treatment, all participants start with the same circumstances and have the same expected luck. The *Identical-High* treatment is similar to the baseline treatment, expect that the expected luck twice higher than in the baseline. In those two treatments, participants have the same ex-ante chance of success.[6] The main treatment of interest is the *Heterogeneous* treatment, in which I vary the circumstances that people start with (the expected luck is equal to that of the baseline). In total, participants perform four rounds of the *Heterogeneous* treatment, one round of *Identical-Low* and one round of *Identical-High* treatment, in random order. Hence, each participant is exposed to potentially five different circumstances values. For each treatment, I elicit their expectations conditional on the outcome. After the real-effort tasks, pairs comprised of one successful and one unsuccessful member are randomly formed. A bonus of $1 is

[5] In the original slider task, people can do a maximum of 60 tasks in 2 minutes. In the experiment, I use a longer version in which people can do up to 300 tasks in essentially as much time as they wish.

[6] Some participants of course end up being luckier if they draw a higher luck. But in each of those treatments participants face the same decision problem since they have the same circumstances and only learn their luck after the effort task.

given by default to the successful member. Participants then act as a third party and are allowed to redistribute the bonus within the pair.

The main results of the experiment can be summarized as follows. The experiment is successful at making effort endogenous. *Within subject*, people adjust their effort level to the circumstances they receive. We rule out that effort changes are driven by time constraints or exhaustion. In the *Heterogeneous* treatment, the success rate is increasing in circumstances. Around 10% of people in the lowest decile of the circumstances distribution reach the target, whereas 96% of people in the top decile do. As predicted by the theory, people who have circumstances values near the top of the distribution exert on average less effort, yet succeed more. However, I also observe lower levels of effort when people have unfavorable circumstances. This is partly driven by a significant fraction of subject who choose to provide no effort when they have unfavorable circumstances, which I call the *demotivation effect*. Eighty-seven percent of people who provide zero effort have circumstances below the median.

Inferences are accurate. Overall, participants understand that difference in circumstances, effort and luck matter in determining outcomes. Within each treatment, they properly infer that those who succeed on average exert more effort and are luckier than those who fail. In the *Heterogeneous* treatment, 91% of people believe that the expected circumstances conditional on success is higher than conditional on failure. This is the rational inference predicted by the model. Yet 70% of people also believe that those who succeed exert more effort on average than those who fail. This is what our data indeed shows, but it is different from what the model predicts in the case where all workers optimal effort is positive. In the other treatments with identical circumstances, people believe that those who succeed are luckier (i.e., draw higher luck) in *Identical-High* than in *Identical-Low* treatment, but that they exert more effort in the latter. Again, this prediction is in line with our data.

When third parties observe only the outcome and not effort or luck directly, beliefs seem to have very little effect on redistribution. We do not find that third parties redistribute earnings differently across treatments. However, giving information about the circumstances in addition to the outcome changes redistribution in an important way. First, third parties care about difference in circumstances. The larger the circumstances difference between the members in the pair, the higher the level of redistribution. Second, the share of third parties who do not redistribute increases to 38% when the successful member in the pair is disadvantaged and starts with worse circumstances. In this case, third parties think that the successful member provides significantly more effort than the other member in the pair. This behavior, which compensates for luck and rewards effort, is consistent with meritocratic preferences.

1.1.1 Related Literature

This paper relates to the literature on fairness preference and redistribution. [2] distinguish three types of fairness preferences. Libertarians never redistribute income so that they consider any income difference as fair. Egalitarians always redistribute income equally and therefore view any income inequality as unfair. Finally, meritocrats care about the source of income inequality. When income inequality is due to luck, meritocrats tend to redistribute earnings [3, 4]. They tolerate more income inequality - hence, redistribute less - if it is caused by differences in performance, as in [5]. In studies, people who hold meritocratic fairness views are often the most prevalent group. [9] estimate their share at around 38% in the United States and 43% in Norway.[7] In most experimental and field studies, redistributive behavior has been investigated when the outcome is either due to effort or luck, and the cause of inequality is known by participants. This work has been essential to construct a typology of fairness preferences, and understand how these preferences relate to the sources of inequality.

The effect of uncertainty on redistribution has only been recently considered in the literature. [5] use an experiment to investigate how people redistribute earnings when income inequality is caused by effort *or* luck, in varying proportions. [10] study an environment where both luck

[7] [9] estimate that the US population is made of 29% of libertarians and 15% of egalitarians. The term "meritocrat preference" encompasses a wide range of behaviors. Some meritocrats may fully compensate for inequalities caused by luck, while others would only do so partially. Some may not redistribute a cent if inequalities are due to performance, while others would redistribute partially.

and performance determine earnings, but in proportions that are fixed and known by those who redistribute.[8] In practice, determining the relative contribution of effort and luck to someone's outcome is complicated as luck and effort are often intertwined, and people may define effort and luck differently.[9] Throughout this paper, we retain the distinction made in the literature and call effort all the factors that the agent is responsible for, and luck all the factors that are beyond the agent's control. Beliefs and inferences about the source of income inequality will affect how meritocrats redistribute earnings. A meritocrat who is convinced that someone's earnings are solely due to effort may oppose redistribution. In contrast, another meritocrat who believe that the earnings are entirely due to luck might support equalizing ex-post earnings.[10] Because people with meritocratic preferences represent the largest share in Western societies, one of the goals of this paper is to understand which inferences people make about the sources of income inequality in an uncertain environment, and how their beliefs affect the way they redistribute income.

The experimental methodology is well suited to circumvent two issues. In real life, the relative contributions of effort and luck to someone's outcome is uncertain and unobserved. In addition, the effort choice is shaped by individual circumstances, which are mainly fixed and time-invariant. Consider a person who grows up in a wealthy family; we cannot observe the choices they would have made if they had been exposed to different circumstances, such as being born into a less affluent environment.[11] Using experiments, we can quantify precisely the contribution of effort and luck to someone's success or failure, and the beliefs that people have about those factors. Measuring those beliefs is important because they affect meritocrat redistribution. In our experiment, we control how

[8]For example, in one treatment 10% of earnings were allocated by luck and the 90% remaining by effort. In another treatment, it was the other way around with 10% of earnings determined by effort and the rest by luck. By contrast, in [5] design, earnings are determined by a coin flip with p% chance, or by performance with $(1\ p)$% chance.

[9]A similar point is made in [11] who note that *"What is luck and what is effort is, in practice, an issue on which people may strongly disagree. Is being born smart purely luck? If so, how do I disentangle success in life that results from some combination of effort and intelligence? Being born in a wealthy family is luck, but what if the wealth accumulated by our parents (...) is the result of great effort?"*

[10]In the first case, the meritocrat will behave like a libertarian, whereas in the second case they will behave like an egalitarian. As this example shows, meritocrats do not necessarily favor more redistribution. An unequal income distribution is considered fair if it is caused only by difference in effort levels. Beliefs about source of inequality are irrelevant for other fairness views. Any income inequality is deemed fair for libertarians and unfair for egalitarians, regardless of the cause.

[11]Vice versa, we cannot observe the choice a child born in a poor family would make if they were born wealthy.

much luck affects the outcome: equalizing circumstances reduces the role of luck since everyone starts with the same chance of success, while making circumstances more unequal means some will be more likely to succeed. Next, experiments also allow us to examine how effort responds to circumstances, by exposing people to different levels of circumstances. The closer set up to mine is found in [12]. They introduce differences in opportunities by using effort multipliers only. In their experiment, altering the multiplier values neither affects the effort exerted by workers nor changes the third parties' effort expectations. The study by [13] is also related to this work, and shows that many people favor equalizing the both ex-ante chance of success and ex-post outcomes.

This paper also contributes to the literature on inequality of ex-ante opportunity, endogenous effort, and real-effort experiments.

1.2 Set-up

1.2.1 Model

A worker i generates production P_i which is a function of effort, luck and circumstances.

$$P_i = \underbrace{\kappa_i}_{circumstances} + \underbrace{\lambda_i \, e_i}_{luck * effort}$$

Circumstances κ_i are drawn from a uniform distribution with support $[\kappa - \delta, \kappa{+}\delta]$, and are known to the agent before they choose how much effort e_i to exert. Unlike circumstances, the agent does not observe their luck, but they know the distribution of luck, which is uniformly drawn from 0 to $\overline{\lambda}$.[12] The agent's outcome is binary. If the agent's production is greater than or equal to an exogenous target T, the agent succeeds and earns a fixed reward x. If $P_i < T$ the agent earns nothing.[13] In the rest of the paper, I normalize the target to one. Circumstances can never be greater than the target production $\kappa + \delta < T$. The lower bound of the luck distribution is chosen to be 0, so

[12] The assumption that people are unaware of their own ability is found in other strands of the literature such as models of career concerns (see for example [14]).

[13] The fixed reward can be interpreted as the labor market return received by an agent for entering a prestigious university or obtaining a certification.

that the agent is never certain to reach the target even if they exert infinite effort.[14] It is possible to interpret luck as an agent's talent or ability. Increasing $\overline{\lambda}$ increases the variance in abilities among agents.

If an agent has unfavorable circumstances with a low k_i, they start far away from the target production. On the other hand, an agent who has favorable circumstances starts closer to the target. Circumstances capture all the factors that can either give an advantage (if good) or make it harder (if bad) for the agent to reach the target: family background, parent's wealth or education, social connection, inherited wealth, discrimination faced etc.[15] The higher the δ, the higher the heterogeneity in circumstances. If $\delta = 0$, everyone has identical circumstances κ. The mean of the distribution κ can be shifted up or down to adjust the expected distance to the target in the population.

1.2.2 Redistribution and Third Party Fairness Preference

A group is comprised of two agents i and j, and one third party. Each agent who reaches the target earns a fixed amount x. The third party observes each agent's outcome and decides how to redistribute the total earnings between the agents in case the agents outcomes are asymmetric, i.e., one of them succeeds and the other fails. Let s_i be the share given to the agent i who succeeds in the pair. The agent j who fails then receives the share $1 - s_i$. If no agent succeeds, there is no earnings to redistribute. If both agents succeed, there is no redistribution and each agent keeps their earnings x. Note that agent j may fail but still receives a payoff if the other agent in the pair reaches the target, and the third party decides that s_i does not equal 1.[16] Hence, after redistribution agent i receives:

[14] By modifying the support of the luck distribution such that $\lambda_i \sim \mathcal{U}\left[\underline{\lambda}, \overline{\lambda}\right]$ with $\underline{\lambda} > 0$, effort might become deterministic if the agent is sure to reach the target at $\underline{\lambda}$.

[15] Some people may be born with favorable social or economic circumstances which makes them more likely to attain a higher level of income. [15] shows for example that the elasticity between a father and son's earning is around .5 in the US. Intergenerational earnings mobility measures gives the flavor that circumstances outside of the child's control - his father's earnings - is positively correlated with their actual earnings.

[16] We do not allow the third party to destroy earnings. [16] conducts an experiment in which subjects are allowed to burn income to reduce inequality.

$$\pi_i = x\,\mathbb{1}_{(P_i \geq 1 \times P_j \geq 1)} + s_i\,x\,\mathbb{1}_{(P_i \geq 1 \times P_j < 1)} + (1 - s_i)\,x\,\mathbb{1}_{(P_i < 1 \times P_j \geq 1)}$$

I follow the literature, e.g. [5], and assume that the third party considers it fair that the high earner H in the pair (i.e., the worker who succeeds) receives a share f_H of the earnings after redistribution. The third party chooses to allocate the share s_H of earnings to the high earner. It is costly for the third party to deviate from what they think is the fair share f_H for the worker who succeeds:

$$U(s_H, f_H) = -(s_H - f_H)^2$$

The third party is only informed whether each agent reaches the target, not their effort level. In other words, the third party cannot condition his choice of s_H on observed effort, but only on the agent's outcome. Following the literature, we distinguish three types of fairness preference. A third party with *egalitarian* fairness preference thinks that $f_H = .5$ and splits equally the total amount to be shared regardless of whether agents succeed or not. Hence, they choose $s_H = .5$. With *libertarian* preference, a third party believes that it is fair that the player who succeeds keeps their earnings $f_H = 1$. Therefore they do not redistribute and chooses $s_H = 1$. The redistribution in the previous two cases is outcome based. It does not require the third party to draw inferences about the agent's effort and luck based on their outcome.

With *meritocrat* preference, the third party redistribution is a function of their belief about the agent's effort, circumstances, and luck. Given that meritocrats reward effort, a conjecture is that s_H is an increasing function of the expected effort conditional on success $\mathbb{E}[e_i \mid P_i \geq 1]$. Meritocrats also offset the effect of factors that are not in the agent's control. Hence, a meritocrat third party is more likely to redistribute earnings if they believe that the low earner L has unfavorable circumstances. That is, s_H is decreasing in $\mathbb{E}[\kappa_i \mid P_i < 1]$.

1.2.3 Optimal Effort

Effort is costly for the agent. I assume that the cost of effort is increasing convex in e_i, so that $c'(e_i) > 0$, and $c''(e_i) > 0$. In the rest of the paper, I assume that $c(e_i) = b\frac{e_i^2}{2}$. This functional form is used in [17] to estimate individual preferences and cost of effort using data from a real-effort experiment.[17] Let p be the probability that agent i succeeds and q the probability that the agent j succeeds.[18] The agent i maximizes his expected payoff:

$$\max_{e_i \geq 0} \; x\, pq + s_i\, x\, p(1-q) + (1-s_i)\, x\, (1-p)q - c(e_i) \tag{1.1}$$

Proposition 1 *The optimal effort e_i^* is equal to*

$$e_i^* = \left(\frac{x\,(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}\right)^{\frac{1}{3}} \quad \text{if} \quad \kappa_i > 1 - \overline{\lambda}\left(\frac{x\,s_i}{b}\right)^{\frac{1}{2}} \tag{1.2}$$

$$e_i^* = \quad 0 \quad \text{otherwise}$$

Proof. See Appendix A1.

The agent trades costly effort off for an increase in probability of success. If, given the effort prescribed by solving problem (1.1), it is impossible for the agent to succeed even if they draw the highest possible luck $\overline{\lambda}$, then they are better off exerting no effort. I characterize the the lower bound on the agent's circumstances for which the agent exerts positive effort, which is given by:

$$\kappa_i > 1 - \overline{\lambda}\left(\frac{x\,s_i}{b}\right)^{\frac{1}{2}} \tag{1.3}$$

[17] [17] use the power cost function $c(e_i) = \frac{b\,e_i^{1+\gamma}}{1+\gamma}$ to estimate their model. In my paper, I make the additional assumption that $\gamma = 1$. The authors also consider an exponential cost function $c(e_i) = \frac{b\,exp(\gamma e_i)}{\gamma}$.

[18] $p(e_i, \kappa_i) = p(\lambda_i \geq \frac{1-\kappa_i}{e_i}) = \frac{e_i\overline{\lambda} + \kappa_i - 1}{e_i\overline{\lambda}}$

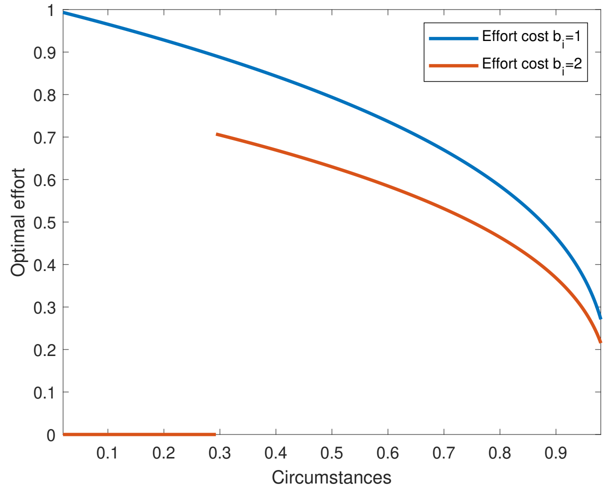


Figure 1.1: Optimal effort e_i^*

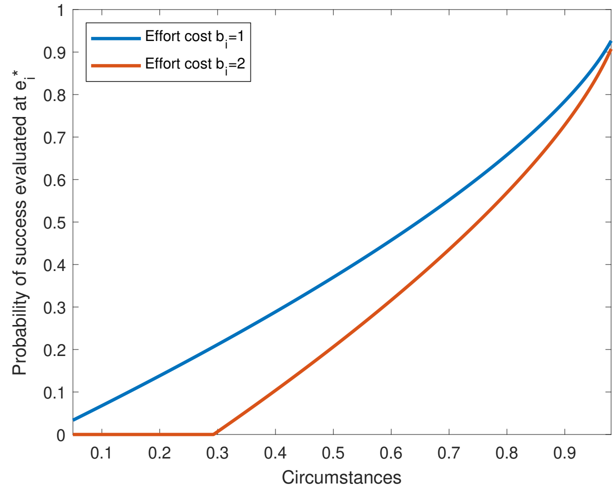


Figure 1.2: Probability of success given e_i^*

If (1.3) does not hold, the circumstances is too small and the agent can never reach the target given the optimal effort induced by $c(e_i)$, hence chooses $e_i^* = 0$. This occurs for instance when agents have a high effort cost b which decreases the optimal effort. Figure 1.1 displays the optimal effort using two different effort cost ($b = 1$ and $b = 2$), and the following parameters: $\kappa = .5, \delta = .49, \overline{\lambda} = x = s_i = 1$. Figure 1.2 plots the probability of success evaluated at the optimal effort for the same parameters. The red curve for $b = 2$ illustrates the corner solution. The agents are better off exerting no effort if they have circumstances below .29. Then there is a discontinuous jump in optimal effort above those circumstances.[19]

Proposition 2 *When effort is positive, effort is increasing in x and s_i, and decreasing in b, κ_i and $\overline{\lambda}$.*

Proof. See Appendix A2.

Proposition 2 summarizes the comparative statics when effort is positive. Increasing the benefit in case of success has a positive effect on effort. So, the larger the reward x or the share given when agents succeed, the higher is the effort. On the other hand, increasing the worker's cost of effort

[19]One has to be careful with the interpretation of figure 1.2. The probability of success is computed for e_i^*. So this probability is zero below .29 because the optimal effort is zero below that threshold. But the optimal effort is zero because the agent's initial effort (possibly positive) resulting from the maximization problem gives them 0% chance to succeed.

b reduces their effort. Finally, better circumstances κ_i or a higher expected luck $\bar{\lambda}$ increases the probability that the workers succeed, which reduces effort.

The agent's effort also depends on how much they expect to receive from the redistribution, and so depends on the third party's fairness preference. If the third party has egalitarian fairness preferences, they choose $s_i = .5$. If the agent knows this, or expects the third party to have such preference, the optimal effort is:

$$e_i^* = \left(\frac{1}{2}\frac{x\,(1-\kappa_i)}{b\,\bar{\lambda}}\right)^{\frac{1}{3}}$$

A libertarian third party chooses $s_i = 1$. If the agent knows this, the optimal effort is:

$$e_i^* = \left(\frac{x\,(1-\kappa_i)}{b\,\bar{\lambda}}\right)^{\frac{1}{3}}$$

Hence, the optimal effort induced by the libertarian rule is always strictly greater than under an egalitarian rule.

Observation 1 *The optimal effort is higher when a worker knows or expect the third party to have libertarian preferences.*

The intuition is that under the egalitarian rule, the agent receives a lower share if he succeeds alone, and a higher share if he fails alone. So, the incentive to work is lower in both cases.

1.2.4 Third Party's Conditional Beliefs

Proposition 3 *The expected circumstances conditional on success and failure are respectively given by:*

$$\mathbb{E}[\kappa_i \mid P_i \geq 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i \, \frac{p(P_i \geq 1 \mid \kappa_i, e_i^*)\, p(\kappa_i)}{\int_{\kappa-\delta}^{\kappa+\delta} p(P_i \geq 1 \mid \kappa_i, e_i^*)\, p(\kappa_i)\, dk} \, dk \tag{1.4}$$

$$\mathbb{E}[\kappa_i \mid P_i < 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i \, \frac{p(P_i < 1 \mid \kappa_i, e_i^*) \, p(\kappa_i)}{\int_{\kappa-\delta}^{\kappa+\delta} p(P_i < 1 \mid \kappa_i, e_i^*) \, p(\kappa_i) \, dk} \, dk \tag{1.5}$$

Proposition 4 *The expected effort conditional on success and failure are respectively given by:*

$$\mathbb{E}[e_i \mid P_i \geq 1] = \left(\frac{x\,(1 - \mathbb{E}[\kappa_i \mid P_i \geq 1])\, s_i}{b\,\bar{\lambda}}\right)^{\frac{1}{3}} \tag{1.6}$$

$$\mathbb{E}[e_i \mid P_i < 1] = \left(\frac{x\,(1 - \mathbb{E}[\kappa_i \mid P_i < 1])\, s_i}{b\,\bar{\lambda}}\right)^{\frac{1}{3}} \tag{1.7}$$

Proofs. See Appendix A3.

The only information available to the third party is the agents' outcome, not their circumstances. Conditional on observing success or failure, the third party forms a belief about the circumstances of the agent. When doing so, the third party estimate the probability of success (or failure) for each level of circumstances assuming that worker exert the optimal effort $e_i(\kappa_i)$.[20] Finally, the expected (optimal) effort conditional on the outcome is just found by plugging the expected circumstances in the optimal effort function.

I simulate the model with the following parameters: $\lambda_i \sim \mathcal{U}\,[0, 1]$; $\kappa = .5$; $\delta \in (0, .45)$; $s_i = 1$ and $b = 1$. Because $s_i = 1$, this simulates the case where the third party is libertarian and the agents know this. Figure 1.3 below presents the result of the simulation. The higher the heterogeneity in circumstances, the higher is the expected circumstances conditional on success. This implies that the expected effort conditional on success is lower than conditional on failure. In addition, the difference between the expected effort conditional on success and the expected effort conditional on failure increases as the heterogeneity in circumstances grows.[21]

[20]For example, the third party forms his circumstances expectation conditional on success using: $p(P_i \geq 1 \mid$

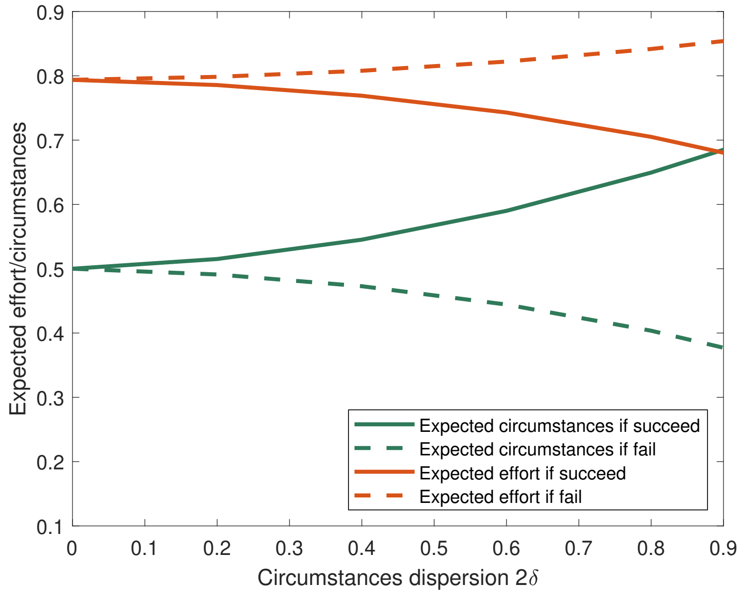


Figure 1.3: Expected circumstances and effort conditional on outcome

1.3 Experimental Design

1.3.1 The Slider Task

The slider task by [7] involves a screen with many sliders (see figure 3.2). Each slider is initially positioned randomly, and participants have to use the computer mouse to move each slider. Their performance is determined by the number of sliders that are positioned exactly in the middle when they decide to stop working on the task. In the original [7] task, people have two minutes to position as many sliders as they can out of a maximum possible of 60. I extend this to allow people to work for as long as they want on a task and position up to 300 sliders. The slider task is chosen because it requires costly effort, yet that ability plays a limited role in the participant's production and I expect little learning or exhaustion. Therefore, the observed differences in production that are not due to variations in circumstances should only be caused by differences in effort.[22] As I describe in the

$\kappa_i, e_i^*) = \frac{e_i^*\bar{\lambda}+\kappa_i-1}{e_i^*\bar{\lambda}}$.

[21]In See Appendix A3. I also discuss how third party's inference are influenced by changes in the individual cost of effort b and the distribution of luck.

[22]Hearing many participants sigh while performing the task during the experiment confirmed our prior belief that the task was costly enough.

result section, I find limited evidence of ability, learning or exhaustion affecting effort in the data.

Figure 1.4: The slider task in the experiment.

Notes: The green dot indicates that the seven sliders on the top-left have been positioned correctly. Participants use their mouse to scroll through the rest of the screen, or use the scrolling bar on the right-side. Participants click on a button "Next" located at the bottom of the page once they decide to stop working on the task.

1.3.2 Treatments

There are two treatment variables: circumstances and luck. First, I vary whether people start with the same or different circumstances (Identical vs Heterogeneous circumstances). Next, I vary the extent to which luck plays a role (Low vs High luck). Recall that the participant only knows the distribution of luck and not his luck draw before performing the task. Otherwise, effort is deterministic and workers exert the minimum effort required to reach the target, and no more (provided that their optimal effort is positive).

Table 1.1: Experimental treatments

		Expected luck	
		Low $\bar{\lambda}$	High $\bar{\lambda}$
Circumstances	Identical ($\delta = 0$)	*IL*	*IH*
	Heterogeneous ($\delta > 0$)	*Heterogeneous*	-

The Identical-Low (*IL*) treatment is our baseline treatment. All participants have the same circumstances, and the same expected luck. This treatment has little variance in luck, so that effort matters more to reach the target. The Identical-High (*IH*) treatment is similar to the *IL* treatment, except that the expected luck is twice as high as in *IL*. There is more luck variance. In that case, the luck draw matters more to determine who succeeds. In both treatments with identical circumstances, participants face the same decision problem because they do not know their luck draw before performing the task. Hence, any difference in effort between participants can be attributed to difference in individual cost of effort. Therefore, those treatments will be a good benchmark to estimate individual effort costs. The Heterogeneous treatment is our main treatment of interest. In this treatment, the expected luck is the same as in *IL*, but most importantly there is heterogeneity in circumstances. So, people may start with very different circumstances. A participant with favorable circumstances starts closer to the target, while a participant who receive unfavorable circumstances may find himself very far from the target, with little chance to succeed even when doing the maximum effort. Participants perform multiple rounds of the *Heterogeneous* treatment, so that they are exposed to different circumstances.

1.3.3 Hypotheses

I summarize here the hypotheses based on the model prediction. I make three broad categories of hypotheses. They are about the worker's effort level, the rational belief of a third party observing only the outcome, and the redistribution behavior. The first three hypotheses are about the chance of success and the optimal effort.

HYPOTHESIS 0: *The probability of success is increasing in circumstances.*

HYPOTHESIS 1: *Effort is decreasing in circumstances.*

HYPOTHESIS 2: *Effort is decreasing in expected luck.*

Hypotheses 1 and 2 follow from the fact that better circumstances or expecting higher luck increase the probability of success. Both hypotheses should hold within and between subjects. The next three hypotheses are about third party's beliefs about what causes success or failure when they observe only the workers outcome.

HYPOTHESIS 3: *Comparing between the identical treatments, third parties expect workers who succeed to exert on average more effort in the baseline than in Identical-High. Third parties expect workers who succeed to have on average higher luck in Identical-High than in the baseline.*

$$\mathbb{E}[e_i \mid P_i \geq 1]^{IL} > \mathbb{E}[e_i \mid P_i \geq 1]^{IH} \text{ and } \mathbb{E}[\lambda_i \mid P_i \geq 1]^{IL} < \mathbb{E}[\lambda_i \mid P_i \geq 1]^{IH}$$

In addition, *within treatment*, third party expects workers who succeed to have on average higher luck than workers who fail.

$$\mathbb{E}[\lambda_i \mid P_i \geq 1] \; > \; \mathbb{E}[\lambda_i \mid P_i < 1] \quad \text{in both } IL \text{ and } IH \text{ treatments}$$

Note that within each treatment I do not make any prediction about difference in effort conditional on the outcome. The reason is simple: participants have the same circumstances and so face the same problem. Thus, they should choose the same effort level, assuming identical cost of effort.

HYPOTHESIS 4: *In the Heterogeneous treatment, third parties expect workers who succeed to have on average better circumstances than workers who fail.*

$$\mathbb{E}[\kappa_i \mid P_i \geq 1]^H > \mathbb{E}[\kappa_i \mid P_i < 1]^H$$

Hence, hypothesis 4 implies:

HYPOTHESIS 5: *In the Heterogeneous treatment, third parties expect workers who succeed to exert on average lower effort than workers who fail.*[23]

$$\mathbb{E}[e_i \mid P_i \geq 1]^H < \mathbb{E}[e_i \mid P_i < 1]^H$$

The last hypothesis is about the redistribution behavior of meritocrats. Libertarians and egalitarian's beliefs about the source of inequality do not affect their redistribution.[24]

HYPOTHESIS 6: *Third parties with meritocrat preferences redistribute more when they attribute success to circumstances or luck, and redistribute less when they attribute success to effort.*

[23] Assuming that people choose effort rationally.

[24] Libertarians never redistribute ($s_i = 0$), while egalitarians divide earnings equally ($s_i = .5$).

1.3.4 Procedure

The experiment has three stages. In stage 1, participants act as workers and have up to 45 minutes to complete five rounds of the slider task.[25] Each worker performs three rounds of *Heterogeneous*, one round of *IL* and one round of *IH*, in random order determined at the individual level. In each round, a new circumstances and a new luck are drawn for each worker. Before doing the slider task, workers are told their circumstances, and the distribution of luck. Once they decide to stop working on the task, their production is computed as follows: the number of sliders they have correctly positioned is multiplied by their luck, to which we add their circumstances.[26] A worker receives $1 if she reaches a fixed production target, and nothing otherwise. Therefore, the first stage replicates the case in which workers face a libertarian third party who never redistribute income. The within-subject design has two advantages. It exposes each worker to different circumstances, so I can observe how effort respond within subject to changes in circumstances. In addition, the design allows me to control for heterogeneity in effort cost across the circumstances distribution.

Table 1.2: Experiment Parameters

		Parameters				
		κ	δ	Luck distribution	Target	Possible effort range*
Treatments	Identical-Low	1250	0	$\mathcal{U}[0, 10]$	2500	0-300
	Identical-High	1250	0	$\mathcal{U}[0, 20]$	2500	0-300
	Heterogeneous	1250	1250	$\mathcal{U}[0, 10]$	2500	0-300

*Effort is participant's choice variable, within the range

The parameters chosen for each treatment are presented in table 1.2. In the identical treatments, each worker circumstances are set at 1250 ($\kappa = 1250$, $\delta = 0$). In the heterogeneous treatment,

[25]Forty-five minutes give participants ample time to exert as much effort as they want in each round, so that time is not binding in the experiment. If they finish early, they can use their phone while others finish the stage.

[26]Subjects are explicitly told that they are not expected to correctly position all the 300 sliders in each round. They are shown multiple examples of how the production is computed. Before starting the first stage, they also have to answer a comprehension question asking them to compute a production. The details of the instructions is available in Appendix C.

the circumstances are randomly draw from a discrete uniform distribution between 0 and 2500 ($\kappa = 1250$, $\delta = 1250$).[27] The parameters are calibrated so that no matter the circumstances, workers always have some chance to reach the target if they exert the maximal effort. The first time they complete each treatment, I elicit participant's belief about the average effort, luck and circumstances of workers who succeed and fail in their session. People are paid for their guesses using a quadratic scoring rule.[28]

Stage 2 involves hypothetical redistribution decisions. First, I elicit participants' fairness preference in case the source of inequality is known to be either effort or luck. Participants can redistribute income within a pair of workers. In the *luck only* scenario, a bonus of $1 is randomly allocated by the computer to one worker in the pair. In the *effort only* scenario, the bonus is given to the worker in the pair who exerted more effort. Those scenarios are commonly used in the literature for example in [9] or [5]. Next, using the results from stage 1, I form for each treatment "one-winner" pairs in which one worker reaches the target and the other does not. A bonus of $1 is by default awarded to the worker who succeeds. In those cases, participants are there is uncertainty about what causes income inequality: a worker's success is due to a combination of effort, luck, and - in the Heterogeneous treatment - circumstances. Participants are reminded of their conditional beliefs from the first stage and can redistribute the bonus within the pair. People are told that their choices are hypothetical but that they should try to answer as if they had real consequences. I use hypothetical questions because I don't inform workers in stage 1 that a subsequent redistribution phase occurs. So, implementing the redistribution decisions would be deceiving.[29] Hence, hypothetical decisions

[27]In order not to confuse participants more than necessary, I chose a round number for the upper bound of the circumstances distribution (2500 and not 2499), so it equals the target in my experiment. In practice, this has little consequence because drawing a circumstances of 2500 and hence succeeding without providing effort occurs with only 0.04% chance (1/2501). In fact, no worker has circumstances of 2500 in my experiment.

[28]I only elicit participant's belief about the average circumstances for the *Heterogeneous* treatment since in the other two treatments participants start with the same circumstances. For example, subjects are asked: "*Think of participants who reach the target in this round, excluding yourself: what do you think is their average effort?*". The same question is asked for people who miss the target. I ask people to consider everyone but themselves to avoid biasing their reports. A participant who exerts the maximal effort but fail because of a luck draw of 0 may overestimate the effort level of people who fail. In addition, this prevents people from strategically choosing their effort to maximize their gains from the belief elicitation, e.g. by choosing to exert no effort and hoping to be the only worker who fail for that treatment.

[29]In stage 1, I am interested in how workers' effort respond to changes in circumstances and the distribution of luck. Telling workers that they will be ex-post paired based on their performance and that a third party will redistribute

seem to offer a good compromise in this specific case.[30]

In stage 3, I combine effort and redistribution. In this stage, each participant has two distinct roles: that of a worker and a third party. As workers, participants have 10 minutes to perform one round of the *Heterogeneous* treatment. As third party, they redistribute earnings between pairs of workers. Before the task, workers are told that ”one-winner” pairs will be formed as in stage 2, and that a third party in their session will be told each worker’s starting circumstances and may redistribute the bonus within the pair. Each third party makes a redistribution decision for two pairs of workers, and their choice directly affect other participants earnings.[31] Finally, I conclude the experiment with a survey which includes demographic questions (gender, socioeconomic class), opinion about income inequalities, fairness views, social mobility and political identification. Questions are borrowed from [19] and the General Social Survey.

1.4 Results

A total of 87 subjects participated in the experiment which was conducted at the Missouri Social Science Experimental Laboratory at Washington University in St. Louis.[32] I ran five sessions of the experiment in March and April 2023. The experiment was coded using oTree [20]. Each session lasted approximately 90 minutes. The average earnings including a $5 show-up fee was $21.82, which was paid in cash at the end of the experiment.

earnings introduces a new potential confound: worker’s effort level might change depending on how much they expect to receive from the redistribution.

[30] To be clear, incentivized decisions are key to experimental economics. In some specific domains such as time preference, studies compared hypothetical to incentivized decisions and showed that they did not differ significantly from each other (see for example [18]). Fairness preferences or more generally other-regarding preferences might be domains in which people’s hypothetical decision may not significantly differ from real ones.

[31] Everything is anonymous: workers do not know who is the other worker in their pair or the identity of the third party who makes the redistribution decision for their pair. Third parties do not know the identity of the workers in the pair when they redistribute. If a worker’s pair is shown to more than one third party, one of the decisions made for the pair is randomly selected to determine the amount earned by the worker. See the experiment instructions in Appendix C for more details.

[32] A participant asked and was granted the permission to leave the experiment after realizing in the first round that they could not perform the slider task due to a medical condition. I excluded this participant from all the analyses.

1.4.1 Effort distribution

In all the analysis that follows, my measure of effort is the number of sliders that are correctly positioned in a given round.[33] Across all treatments, workers correctly position 194 sliders on average. The effort distribution is slightly right censored, as shown in figure 1.5. Workers correctly position more than 290 sliders in 19.4% of the observations, and exert the maximum effort of positioning 300 sliders in 16.7% of the observations across all treatments. The right censoring is partly driven by the first round, in which 36.8% of workers exert the maximum effort. Excluding round one, as shown by the red histogram, only 12.6% of workers exert the maximum effort across all treatments, and 15.3% place correctly more than 290 sliders. Presumably participants are learning about the task during the first round and solving all the sliders might be the natural choice for workers who are still unsure of how the production is calculated. After each round, workers learn their production and whether they reach the target or not, so the task should be well understood after the first-round feedback. Finally, workers chose to provide zero effort in 7.1% of the observations. Figure 1.6 displays the effort distribution for each treatment in our experiment. Comparing between treatments, we also observe that a higher share of participants exerts the maximum effort if the baseline *Identical-Low* treatment. In this treatment, 25.3% of workers correctly place 300 sliders (29.9% correctly place more than 290 sliders), whereas 14.9% do so in the other two treatments.[34]

As mentioned in section 1.3.1, difference in production should reflect differences in effort and not between subject difference in abilities. Next, within subject, effort should not be affected by exhaustion or time constraints, so that any effort variation can be attributed to changes in the experiment parameters i.e., circumstances and luck. In appendix A.2.1, I provide evidence that none of these factors affect worker's effort level in a significant way. In particular, worker's productivity do not decrease over time, and around 93% of participants enter round 5 with more than four minutes

[33]In practice we can define effort in different ways. It can be the total time spent by people in a given round, or worker's efficiency: how many seconds a worker needs to make one unit of effort, i.e., properly position one slider.

[34]Since the order of the treatments is randomized, the higher level of effort in the baseline treatment is not due to the learning effect of round 1. That is, there is as many participants who start with the *Identical-Low* treatment as participants who start with the *Identical-High* treatment. And since participants have to complete three rounds of *Heterogeneous*, they are actually more likely to face it in the first round.

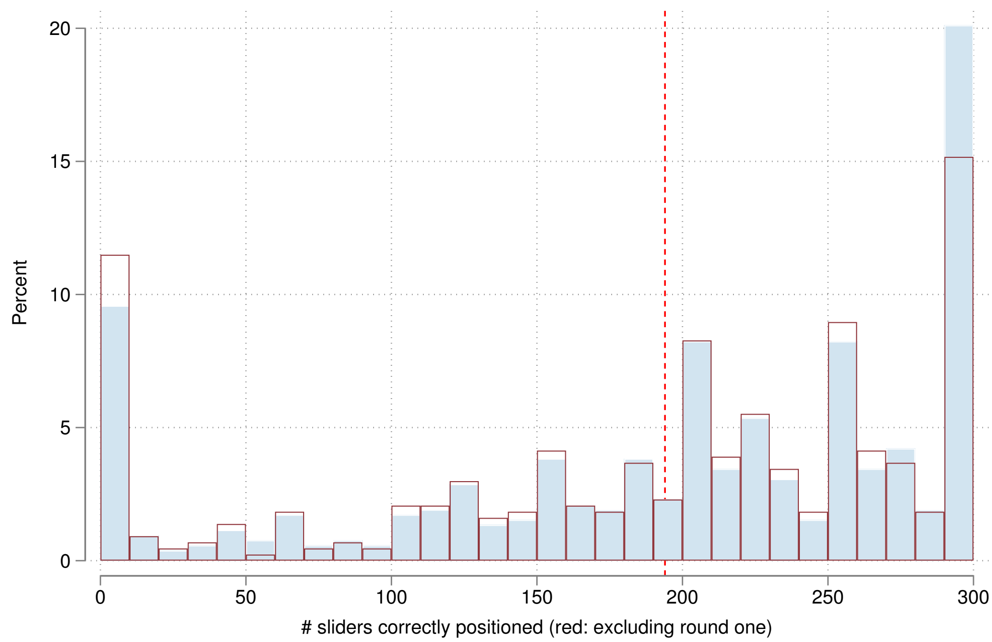


Figure 1.5: Effort distribution with bins width of 10

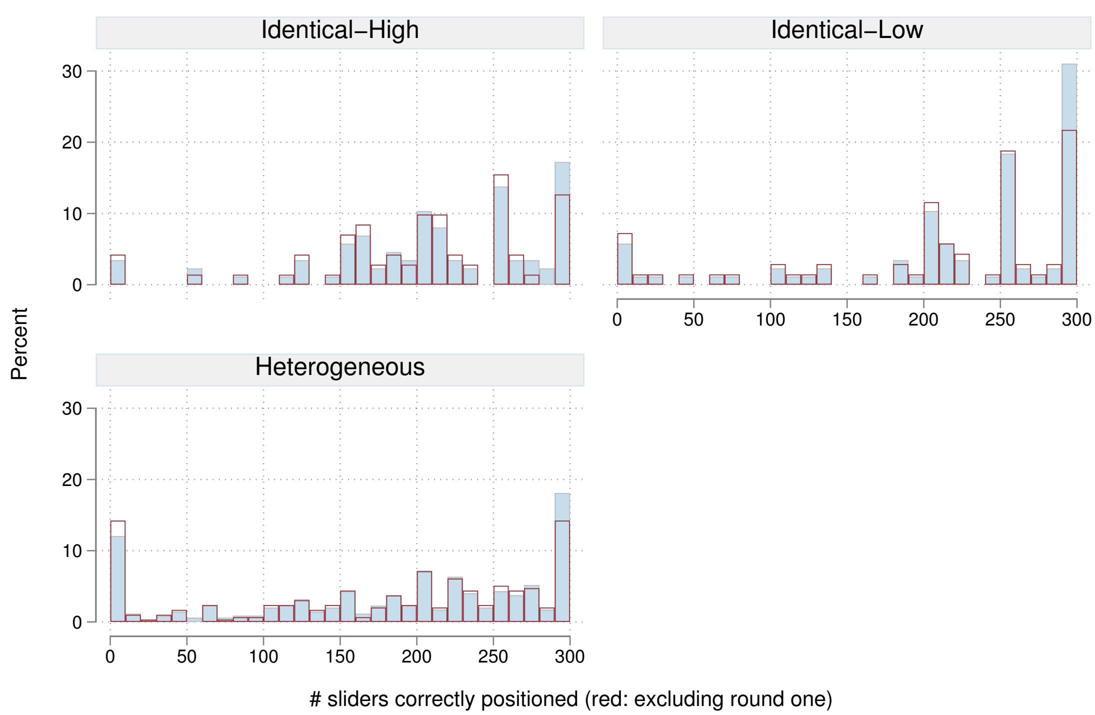


Figure 1.6: Effort distribution by treatment with bins width of 10

left to complete the round.[35]

1.4.2 Effort Choices

In this section I focus on how changes in circumstances and distribution of luck affect individual effort choice.

Effect of Circumstances on Effort

Figure 1.7 show the success rate and the average effort level conditional on participant's circumstances, for the *Heterogeneous* treatment. Because each subject performs the heterogeneous treatment four times, they receive four different circumstances. Hence, a participant may have circumstances at the bottom of the distribution in a given round, and in the next round be at the top of the distribution. This enables to control for individual difference in effort cost. With our 87 participants, I have a total of 348 observations for the *Heterogeneous* treatment, around 35 for each decile. The average success rate is 46.8%. The success rate is almost monotonically increasing in circumstances: while only 10.3% of people who have circumstances in the lowest decile of the distribution (that is, less than 251 out of possible 2500) reach the production target, 95.8% of those who have circumstances in the top decile (above 2250) reach the target.[36] So, people who have better circumstances are more likely to succeed.

We now turn to how effort responds to changes in circumstances. The orange line in figure 1.7 displays the average effort by decile with the 90% confidence interval bounds. Within subject, participants adjust their effort level to the circumstances they receive. The effort level looks like an inverted U-shape. On average, lower effort levels are observed in the tails. People who have circumstances in the top 10% of the distribution on average properly position 108 sliders, those in

[35]The average number of sliders solved per round does not decreases over time, and workers do not spent more time per slider as the rounds progress.

[36]Success depends on people's effort choice, circumstances and luck. Upon closer inspection of the data, one explanation for the fact that the success rate is not monotonically increasing is because participants in the third decile of my experiment were unlucky compared to those of the fourth decile: they receive on average a luck of 4.7 out of 10, whereas the average luck draw of the fourth decile is 5.3.

the bottom 10% position 133 sliders. The highest average effort levels (221 and 230) are observed when people have circumstances between 1250 and 1750. There is a negative correlation between circumstances and effort above the median of the circumstances distribution ($\rho = -0.43, p < .01$), which is even stronger in the top 20% ($\rho = -0.58$). Our experiment succeeds in making effort endogenous, as participants *choose* to exert different effort levels depending on the circumstances. At the top, the conjecture is that receiving a favorable circumstances makes success very likely and hence reduce the need for effort. Consider the following case of a worker who has circumstances of 2300. The expected luck is 5 in the *Heterogeneous* treatment is 5. Hence, by placing properly 50 sliders the expected production is 2550 which is higher than the production target. If the worker places properly 100 sliders, the only way for them to miss the target is if they are unlucky and draw a luck of 0 or 1, which occurs with around 18% probability.

RESULT 1: *In the heterogeneous treatment, the success rate is increasing in circumstances. The levels of effort are lower when workers have circumstances in the bottom or the top decile of the circumstances distribution. In the top decile, workers exert less effort yet succeed more. Effort is negatively correlated with circumstances in the top half of the distribution ($\rho = -.43, p < .01$).*

What explains the lower effort levels at the bottom of the distribution? In the experiment, worker can choose to provide no effort in a round. The behavior of exerting zero effort is much more common when participants have unfavorable circumstances. We call this the *demotivation effect*. The gray line in figure 1.8 display the average effort level by decile when workers who provide zero effort are excluded. Removing them boost the average effort in the bottom half of the circumstances distribution. Among people who exert zero effort, 87% have circumstances below the median. Twenty-eight percent of participants who have circumstances in the lowest decile choose to provide zero effort. In addition, workers provide zero effort more frequently in the *Heterogeneous* treatment (in 8.4% of the observations) than in the other two treatments (in 2.3% of the observations). All this

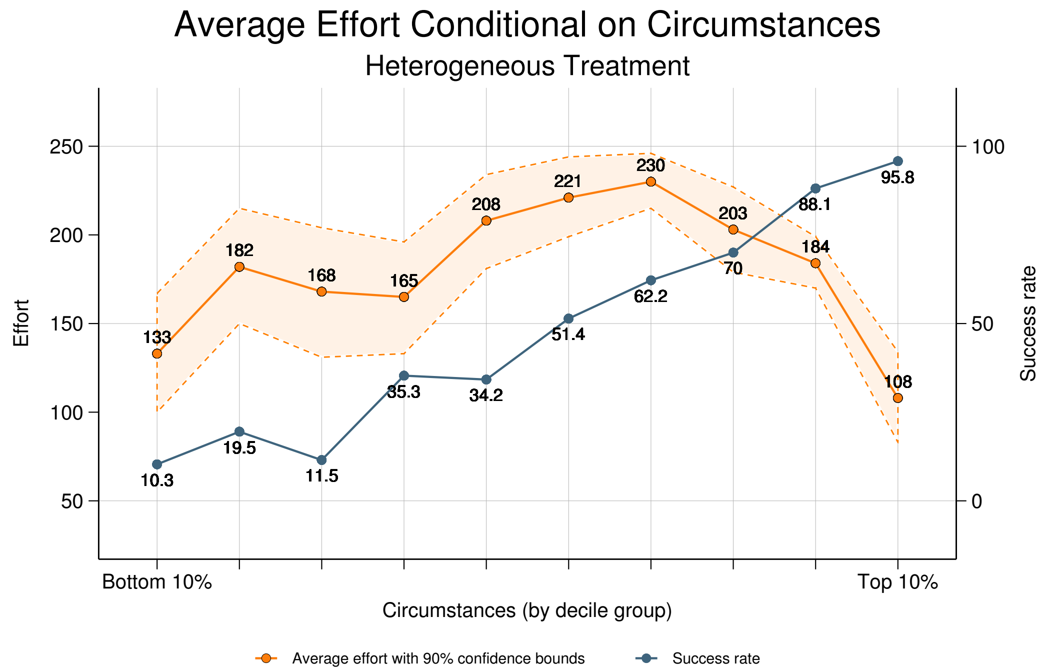


Figure 1.7: Success rate and average effort conditional on circumstances in the heterogeneous treatment.

suggests that thedemotivation effect is driven by unfavorable circumstances.[37]

There are several possible explanations for why workers choose to exert no effort.[38]

1. *High cost of effort:* when workers have high effort cost, it is rational to exert no effort if the optimal effort given their circumstances is too low to give them a positive chance of success. This is the corner case described in section 1.2.3. I estimate each participant's effort cost $\hat{b_i}$ based on the effort level they provide in the baseline treatment.[39] Using $\hat{b_i}$, I then estimate worker's optimal effort $\hat{e_i}(\hat{b_i}, \kappa_i)$ given the circumstances that they receive for all the *Heterogeneous* rounds. Finally, I compute the probability of success given this estimated

[37]This excludes the four people who provide zero effort in the last round of stage 1 because they run out of time before.

[38]Recall that running out of time is not one of those explanation because time is rarely binding, as discussed in 1.4.1. In addition, out of the four participants who run out of time before the last round of stage 1, only one of them should have performed the *Heterogeneous* treatment.

[39]There are two participants who should have performed the baseline *Identical-Low* treatment in round 5 but who reach the time limit before entering the round. For those participants, I estimate the effort cost based on the effort level in the *Identical-High* treatment.

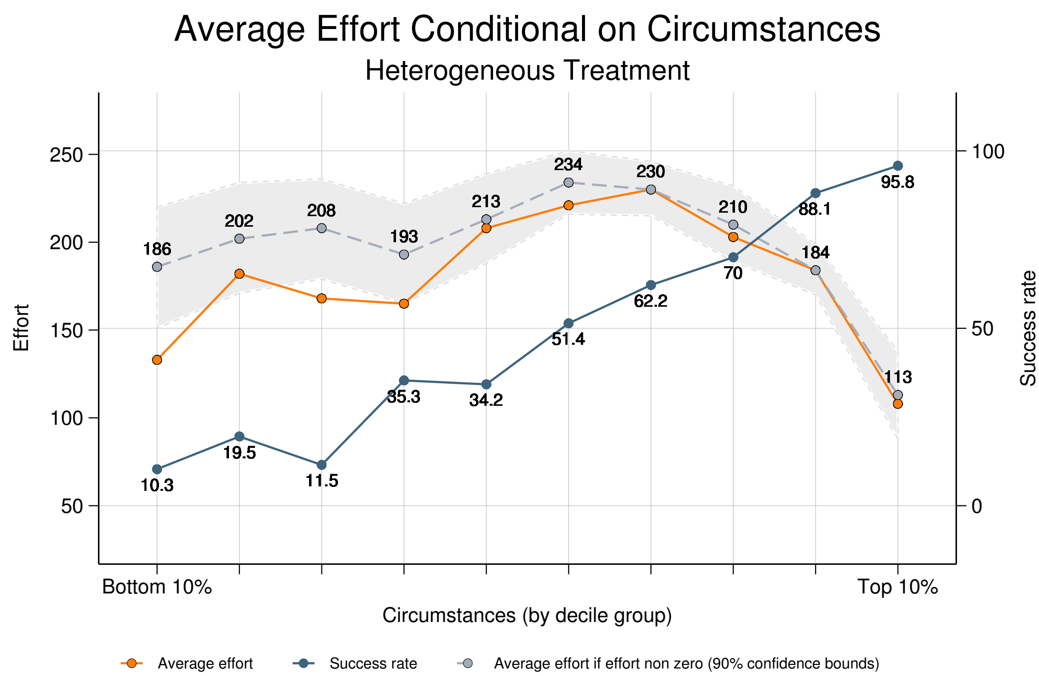


Figure 1.8: Average effort conditional on circumstances in the heterogeneous treatment when effort is non zero

optimal effort. If the estimated probability of success is not positive, then it is rational for the worker to exert zero effort. Using this methodology, providing no effort is rational for only 2 out of the 30 observations in which I observe zero effort. In other words, the demotivation effect does not seem to be due to high effort costs.[40]

2. *Hedging:* workers may be hedging between rounds. When they have unfavorable circumstances, they decide to exert no effort and instead save their energy for future rounds in which they hope to have better circumstances. If this is the case, I should observe no demotivation effect in worker's last round. In the last round of stage 3, 8% of participants exert zero effort, even though they know this is the last effort task. This confirms the result of the regression in table 1.3, which showed no round effect.

3. *Previous round failure:* people may be more likely to provide no effort if they fail in the

[40] I perform the same analysis using the *IH* treatment as benchmark to estimate effort costs and arrive at the same conclusion. In this case, in only 1 out of the 30 observations doing no effort is rational as workers have no positive probability of success given their effort cost.

previous round and end up being disappointed by their previous outcome. Overall, the failure rate in the *Heterogeneous* treatment is 48.3%. Among people who provide zero effort, 58.6% of people fail in the previous round. Therefore, this effect appears limited.

I confirm these with logit regression (see table 1.3) where the choice to exert zero effort, i.e., being demotivated, is the dependent variable. Moving up in the circumstances distribution by one decile reduces probability of being demotivated by 2.8%. Apart from the circumstances, no other variable significantly affect the probability of exerting zero effort. In particular there is no effect of the round number, no treatment effect and no effect of the previous round outcome. The overall picture that emerges from this analysis is that the demotivation effect is a behavioral phenomenon. Only in rare cases it is a rational response, in particular to high effort costs.

Table 1.3: Logit With Dependent Variable: Choose No Effort (Demotivated)

	(1)
Circumstances (decile)	−0.431
	(0.104)
Treatment: Identical-Low	0.218
	(.808)
Treatment: Heterogeneous	0.633
	(.689)
Fail in round n−1	−0.465
	(.394)
Constant	−1.214
	(0.873)
Round controls	Yes
Subject FE	Yes
Observations	435

Standard errors in parentheses, clustered at participant level

$p < 0.10$, $p < 0.05$, $p < 0.01$

RESULT 2: *In the heterogeneous treatment, the lower effort levels at the bottom of the circumstances distribution are partly caused by workers exerting no effort (demotivation effect). Workers' circumstances is below the median in 87% of the observations in which the demotivation effect occurs. Moving up in the circumstances distribution by one decile reduces probability of exerting no effort by 2.8%. The demotivation effect does not seem to be caused by high effort cost.*

Effort Choice and Rationality

Some workers may fail to realize that given their circumstances and the luck distribution, they choose an effort level that is too low to ever reach the target. For instance, in the baseline treatment a worker with circumstances of 500 who chooses an effort level of 100 is sure to fail, because even if they draw the highest possible luck of 10, their production can at most be 1500, which is below the target production. The worker should choose an effort level of at least 200 sliders to have a small chance to succeed (in that case, the probability of reaching the threshold is 1/11%, as this happens only if the worker draws a luck of 10). Across all treatments, workers choose a positive effort level that yields no chance of success in 14.4% of the observations. This occurs very rarely in the *IH* treatment (2.4% of the observations) because with a maximal luck draw of 20, even low levels of effort bring some probability of success. Non rational effort levels occurs more frequently (18.6% of the observations) in the *Heterogeneous* treatment due to unfavorable circumstances.[41]

I documented two types of irrational behavior. First, with the exception of a few rare cases, workers providing zero effort (i.e., the demotivation effect) can not be explained by high effort costs. The second type of irrational behavior is workers who exert a level of effort which, given their circumstances, is not enough to have a positive probability of success. Figure 1.9 displays the average effort and success rate by decile of the circumstances distribution for the observations in

[41] Irrational effort levels may be either one-time mistakes, or the result of dynamic inconsistency. In the latter case, workers start a round planning to exert a predetermined effort level, but fail to carry out this plan. Before doing the task they might underestimate their effort cost which leads them to overestimate their optimal effort. In the data, 62.5% of workers choose an irrational effort level once, and 37.5% of workers choose an irrational effort more than once (the vast majority twice). So, the data seems to point in the direction of the single mistake explanation.

which workers behave rationally.[42] This represents 75% of the observations in the *Heterogeneous* treatment. When we consider only those observations, the average effort is more in line with the theory prediction of a monotonic decrease in effort as circumstances increase. So, do workers who succeed exert more or less effort in the experiment? In the *Heterogeneous* treatment, workers who succeed have better circumstances (1649 vs 842, $p < 0.01$) and exert more effort (216 vs 154, $p < 0.01$). But in observations for which effort is rational, workers who succeed have better circumstances (1649 vs 1081, $p < 0.01$) and exert less effort (216 vs 234, $p < 0.05$). In the Identical treatments, workers who succeed have better circumstances and exert more effort. This is also true in observations with rational effort, even though the difference in effort between the two groups is no longer significant.

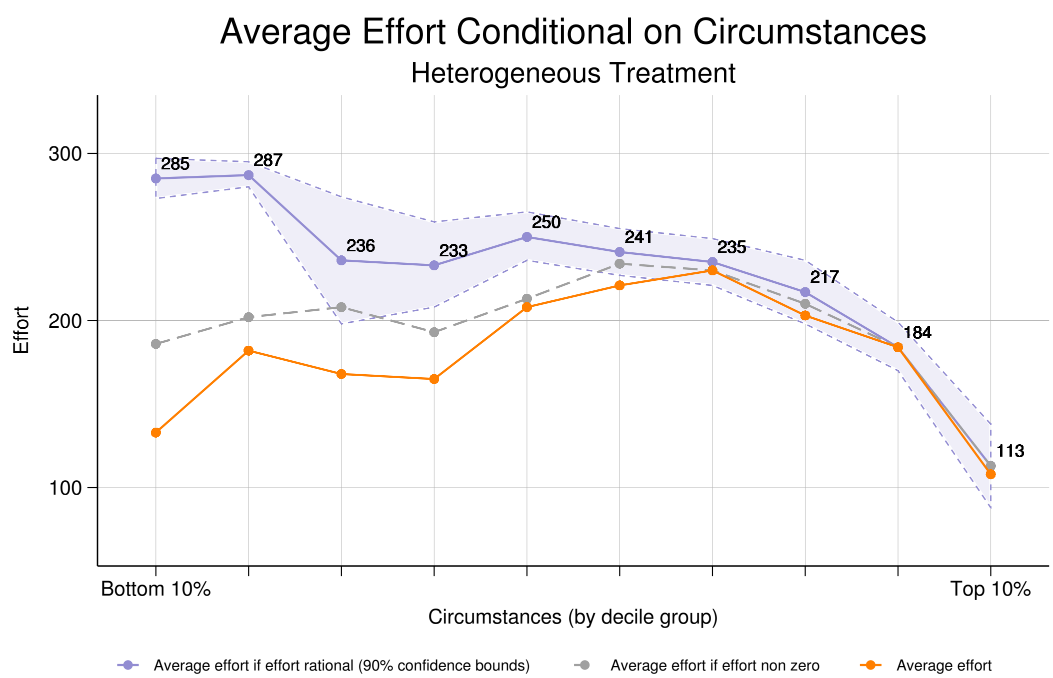


Figure 1.9: Average rational effort conditional on circumstances in the heterogeneous treatment

[42]This includes the observations in which workers rationally exert no effort because they have high effort costs, and the observations in which workers run out of time.

Effect of Luck on Effort

In the *IL* and the *IH* treatment, every worker starts with the same circumstances. The only difference is the expected luck, which is 5 in *IL* and 10 in *IH*. Workers on average correctly position 220 sliders in *IL* and 212 in *IH*.[43] The difference is not significant, which indicates that, contrary to hypothesis 2, workers do not exert less effort when the expected luck increases. Participants may be inattentive and fail to notice that the distribution of luck in the *IH* treatment differ from the other treatments.[44] Another possible reason is that, even if workers observe that the expected luck is higher and understand that this increases their chance of reaching the target, they may not reduce their effort level due to anticipation of regret in case they receive a high luck draw and miss the target because of their lower effort.[45]

1.4.3 Beliefs Conditional on the Outcome

In this section we turn to participant's beliefs conditional on success or failure in reaching the target. After the first instance of each treatment, we ask participants their belief about the average effort, luck, and circumstances (when applicable) of workers who succeed and fail in their session.

Identical Circumstances Treatments

As shown in table 2.3, the expected effort conditional on success is higher in the baseline *IL* treatment compared to the *IH* treatment: 215 vs 197 ($p < 0.05$). The average effort of workers who succeed is 259 in *IL* and 228 in *IH* ($p < 0.01$). So, participants underestimate the effort of workers who succeed in both treatments, but still properly infer that workers who succeed in the baseline exert more effort than in *IH*.

[43] The average success rate is 75.9% in *IH* and 43.7% in *IL*.

[44] To prevent this, participants are told in the instructions to pay close attention to the distribution of luck which is revealed to them at the beginning of each round.

[45] It is also possible that workers do not anticipate that a higher expected luck increases their likelihood of success, so that *ceteris paribus* less effort is required to succeed. However, we show in section 1.4.3 that participants believe that workers who succeed in the *IH* treatment receive higher luck draws than in *IL*.

Table 1.4: Beliefs Conditional on Success (Identical Circumstances Treatments)

		Treatment		
		Identical-Low	Identical-High	Difference
Average effort if success	Truth	259	228	31
	Belief	215	197	18
	Δ Belief − Truth	−44	−31	
Average luck if success	Truth	7.4	12.5	5.1
	Belief	6.8	11.9	5.1
	Δ Belief − Truth	−0.6	−0.6	

Two-sample t-test: $p < 0.10$, $p < 0.05$, $p < 0.01$

Turning to luck, participants estimate that the average luck draw of workers who succeed in the *IH* treatment is 11.9 vs 6.8 in *IL* ($p < 0.01$). Beliefs about the average luck of workers who succeed are accurate in both treatments, even if a small underestimation is again observed. Workers who succeed receive on average a higher luck draw in the *IH* treatment compared to those who succeed in *IL*: 12.5 vs 7.4 ($p < 0.01$).

RESULT 3: *Participants believe that the average effort of workers who succeed is higher in the baseline treatment compared to IH. Compared to the baseline, participants believe that the average luck draw of workers who succeed is higher in IH.*

Appendix **??** provide the table of the beliefs conditional on failure. Within each treatment, participants always believe that workers who succeed on average exert more effort and have a higher luck draw than those who fail. [46]

[46] For example, in the *IH* treatment participants believe that on average those who fail exert 151 effort while those who succeed exert 197 effort. The difference between the beliefs conditional on the outcome within treatment is always highly significant ($p < 0.01$).

Heterogeneous Circumstances Treatment

According to the hypotheses, in the *heterogeneous* treatment the expected circumstances conditional on success should be higher than conditional on failure (H4), which implies that the expected effort conditional on success should be lower than conditional on failure (H5).[47]

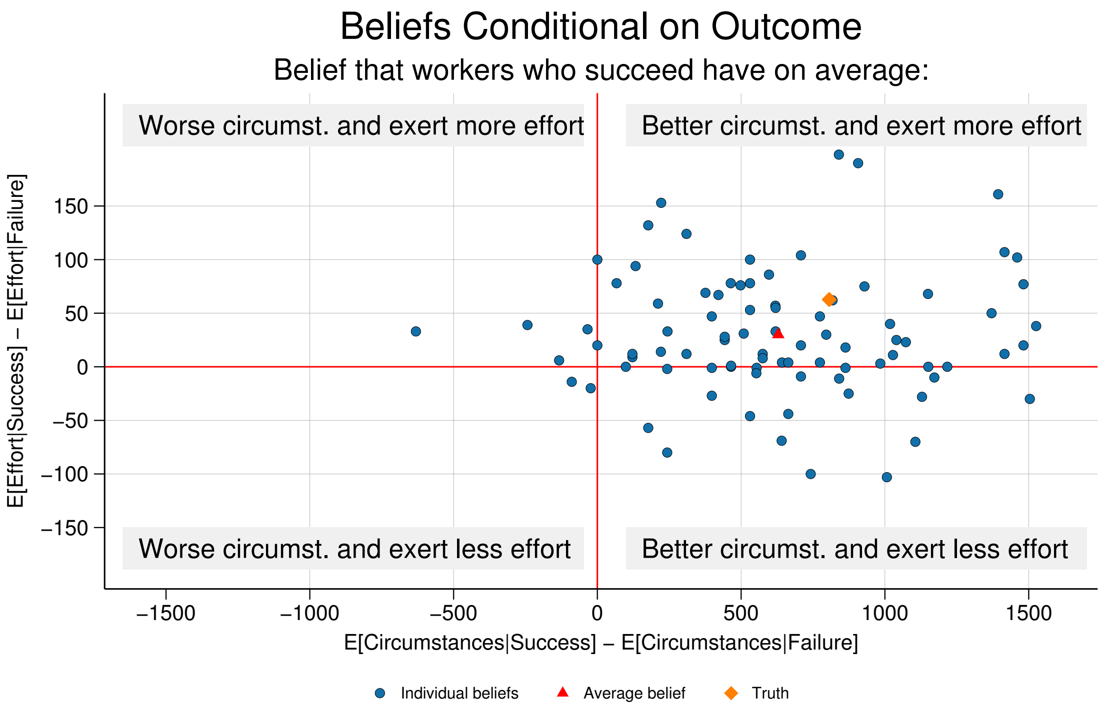


Figure 1.10: Beliefs in heterogeneous circumstances treatment

In figure 1.10, I plot for each participant the difference between the expected effort of workers who succeed and the expected effort of workers who fail (y-axis). One blue dot represents the belief of one participant in the experiment. Above (Below) the horizontal red line, participants believe that workers who succeed in the *Heterogeneous* treatment exert more (less) effort than those who fail. The higher the point, the stronger is the belief that those who succeed exert more effort than those who fail. On the x-axis, I do the same for the beliefs about circumstances. Hence, on the right (left) of vertical red line, participants believe that workers who reach the target on average have better

[47]Recall that this holds if all workers exert positive effort, and that the implication holds because effort is monotonically decreasing in circumstances.

(worse) circumstances than those who miss. The more to the right, the more a participant believe that on average those who succeed have better circumstances than those who fail. Ninety-one percent of participants believe that workers who succeed in the *Heterogeneous* treatment have better circumstances than those who fail.[48] This provides support for H4. On average, they think that those who reach the target have circumstances of 1497 whereas those who miss have circumstances of 868, a difference of 629. In truth, those who reach the target in the *Heterogeneous* treatment have on average a circumstances of 1649 while those who miss have circumstances of 843. Again, subject's beliefs are very accurate.

Among participants who believe that workers who succeed have better circumstances, 70% believe that they also did more effort. This, on the other hand, contradicts H5. In our data, workers who succeed in *Heterogeneous* correctly position 216 sliders, and those who fail correctly position 154 of them. Participants believe that that those who succeed on average correctly position 205 sliders while those who fail correctly position 175.[49]. So, most participant (63%) believe at the same time that those who succeed on average have better circumstances and exert more effort than those who fail: they appear in the top right quadrant on figure 1.10. Roughly a quarter of our participants (in the bottom right quadrant) believe that those who reach the target on average have a better circumstances and exert less effort than those who fail, which is what the model would predict.[50]

Our theoretical model predicts that when circumstances are heterogeneous and all workers exert positive effort, conditional on observing only the outcome, third parties expect those who succeed to have better circumstances and exert less effort than those who fail. When we elicit beliefs in our experiment, third parties instead guess that those who succeed on average have both better circumstances and exert more effort than those who fail, which is what our data shows. In other words, third parties predict where the model fail. The failure of the model is due to the irrational

[48]Visually, 91% of the blue points are on the right side of the vertical red line.

[49]The difference between the circumstances and effort conditional on the outcome are both significant at 1% level (pairwise t-test). The difference between the belief about effort and belief about circumstances conditional on the outcome are also both significant at 1% level (paired t-test)

[50]The exact count of how many participants hold each belief is presented in appendix A.2.2. The average belief is rather close to the truth (pink dot) in our experiment, even if participant slightly underestimate the role of effort and circumstances in success.

effort choice documented in section 1.4.2, which decreases the average effort of workers who fail. Restricting to rounds in which participants exert rational effort, the average effort of workers who fail is higher than those who succeed (234.1 vs 216.3, $p < 0.05$) even though their circumstances are worse on average (1081 vs 1649, $p < 0.01$). [51]

RESULT 4: *In line with our data and the theory, 91% of third parties believe that workers who succeed in the heterogeneous treatment come from better circumstances. In line with our data but not the theory, 70% of third parties believe that workers who succeed exert more effort than those who fail. Overall two-thirds of third parties believe that workers who succeed on average have better circumstances and exert more effort than those who fail.*

1.4.4 Redistribution

Redistribution Conditional on the Outcome

In the second stage, participants act as third parties and decide how to redistribute a bonus of $1 within a pair of workers. In the benchmark decisions, there is no uncertainty about the source of income inequality before third party redistribute. The bonus is allocated by default to one worker in the pair, either randomly (*luck only*) or given to the worker who exerted more effort (*effort only*). Table 2.4 presents the main descriptive statistics for the redistribution behavior in stage 2. In the benchmark cases, when the winner is determined by luck, 64.3% of third parties decide to split the bonus equally between participant, and only 6.9% do not redistribute. When the winner is determined by performance, a lower number of third parties redistribute earnings, as 25.3% of third parties choose not to redistribute. In this case, two-thirds of participants choose a transfer that ultimately leaves the best performer with more income than the worst performer.[52] Conditional

[51] Looking only at the tails of the circumstances distribution, the average effort made by workers who succeed in the top decile (113, n=23) is slightly lower than the average effort made by those *who fail* from the lowest decile (114.8, n=35). Excluding the demotivated, the average effort of participants who miss the target in the lowest decile is 167.5 (n=24), which is much higher than the one of those who succeed from the top.

[52] This includes the people who do not redistribute. The exact distribution of behaviors in those benchmark cases is shown in appendix A.2.5.

on redistributing, there is no significant differences in the average amount transferred between the benchmark cases: average redistributed share is 47.1% in *luck only* and 45.7% in *effort only*. Using third parties answer, I classify third parties as libertarians if they do not redistribute income in *luck only*, and as egalitarian is they split equally the bonus in *effort only* and *luck only*. Finally, I classify as meritocrats third parties who give more to the best performer in *effort only* and equalize earnings in *luck only*. Using this classification, I find 7% of libertarians, 13% of egalitarians and 44% of meritocrats in my sample. [53]

Third parties also have to make one redistribution decision for a randomly selected pair of workers in each treatment, in which the bonus is by default allocated to the worker who succeeds. Third parties only know the worker's outcome, not their effort, luck or circumstances (in *Heterogeneous*).[54]. The redistribution pattern when there is uncertainty about what causes income inequality is closer to the *effort only* benchmark when it comes to the number of third-party who choose an equal split. The level of 50-50 split is 41 percentage points lower than in the *luck only* case. Across our three treatments, 16.1% choose not to redistribute.

A regression with the redistributed share as dependent variable and treatment dummy confirm that there is no difference in how third parties redistribute between treatment (table 2.5). Both third parties preferences and beliefs affect redistribution. The more third party redistribute when they know the causes of income inequality (i.e., in the two benchmark questions), the more they redistribute when there is uncertainty about it. Next, beliefs about the characteristics of the worker who succeed also seem to play a role. The higher the expected *effort of the worker who reach*, the less is transferred ($p < 0.05$).[55] Third parties seem to care about the causes of the winner's success, and they redistribute less when the worker who succeeds exert more effort. This has the flavor of a meritocratic behavior. Unlike effort and luck, beliefs about worker's circumstances have no effect

[53]My sample has slightly more meritocrats, less libertarians and about the same share of egalitarians than in [9].

[54]Pairs that are formed are always comprised of who worker who succeed and one who fail. Participants are reminded of their beliefs conditional on the outcome

[55]The higher the expected luck of the worker who reach in the pair, the more is transferred, although this is only significant at 10% level.

Table 1.5: Redistribution Descriptive Statistics

	Source of Income Inequality					
	Luck only	Effort only	Treatments with uncertainty			
			All	IH	IL	Heter.
Mean redistributed share	0.439	0.341	0.351	0.339	0.349	0.361
	(0.177)	(0.303)	(0.240)	(0.233)	(0.253)	(0.236)
Median redistributed share	0.5	0.3	0.35	0.35	0.35	0.4
Share split 50-50	**0.643**	0.172	0.230	0.230	0.195	0.264
Share do not redistribute	0.069	**0.253**	0.161	0.149	0.172	0.161
Observations	87	87	261	87	87	87

on redistribution in this stage.[56] Finally, I also consider whether the magnitude of difference in third-party's conditional expectations affect redistribution, for instance how much additional effort a third party expect workers who succeed to exert compared to those who fail (see column 2 in table 2.5). The larger the gap between the expected effort of the worker who succeeds and the worker who fail, the less third party redistribute.

RESULT 5: *When third parties only observe worker's outcome, fairness preferences influences redistribution more than beliefs about worker's effort, luck, and circumstances. Third parties redistribute less when the difference between the expected effort of the worker who succeeds and the worker who fails increases.*

Redistribution With Known Circumstances

In stage 3, third parties redistribute income for two ”one-winner” pairs. This time they are told each worker's circumstances in addition to their outcome. Before redistribution, we show third

[56]Looking at each treatment separately, I confirm that beliefs about worker's circumstances do not affect redistribution in the *Heterogeneous* treatment. See Appendix A.2.6, where I look into more details at the determinants of redistribution behavior for each treatment separately.

Table 1.6: Redistributed Share

	(1)	(2)
Belief average effort if succeed	−0.056	
	(0.028)	
Belief average effort if fail	0.033	
	(0.021)	
Belief average luck if succeed	0.935	
	(0.551)	
Belief average luck if fail	−0.557	
	(0.565)	
Belief average circumstances if succeed	−0.084	
	(0.528)	
Belief average circumstances if fail	0.380	
	(0.629)	
Belief difference in effort		−0.040
		(0.019)
Belief difference in luck		0.677
		(0.416)
Belief difference in circumstances		−0.214
		(0.458)
Effort only redistribution	0.166	0.165
	(0.079)	(0.078)
Luck only redistribution	0.353	0.350
	(0.136)	(0.134)
Treatment dummy	Yes	Yes
Subject FE	Yes	Yes
Groups	87	87
Observations	261	261

$p < 0.10$, $p < 0.05$, $p < 0.01$. OLS with subject Fixed Effects

Notes: To improve readability of the coefficients, the belief about circumstances are in hundreds. The ”Belief difference in effort” independent variable is the difference between the expected effort of workers who succeeds and the expected effort of workers who fail. The same goes for the ”Belief difference in luck” and ”Belief difference in circumstances” beliefs: they are the difference in beliefs conditional on the outcome. ”*Luck only* redistribution” refers to the amount transferred from the lucky to the unlucky worker in the *luck only* benchmark case. ”*Effort only* redistribution” refers to the amount transferred from the best to the worst performer in the *effort only* benchmark case.

parties a table indicating the probability of success in the *Heterogeneous* treatment given different combinations of circumstances and effort.[57].

Third parties redistribute much less when the worker who succeeds has lower circumstances, as shown in table 2.6. Around thirty-eight percent of third parties decide *not* to redistribute in that case, which is the modal behavior, and only 10.3% split the bonus equally. By contrast, the most prevalent behavior (33.8%) among third parties is to split earnings equally between workers when the worker who succeeds has better circumstances. On average, third parties redistribute 40.3% of the earnings to the worker who fail when the worker who succeeds has better circumstances, but redistribute only 26.7% of earnings when the winner has worse circumstances. The difference in average redistribution level is significant at 1% level (two-sample t-test).[58] For pairs in which the winner has lower circumstances, the lower levels of redistribution seems to be driven by the belief that the disadvantaged worker who succeeds did about 30% more effort than the one who fail (254 vs 192, $p < 0.01$). This is suggestive of meritocratic behavior, where people want to reward hard work. Third parties may consider that workers who succeed despite having worse circumstances had to work harder and are entitled to their earnings.[59]

A regression with the redistributed share has dependant variable confirm that third parties take circumstances into account. When workers' circumstances are known, the difference between the observed circumstances seems to be good predictor of redistribution behavior, as shown in the table **??**. The larger is the difference in circumstances between the workers in the pair, the more third parties redistribute income.

RESULT 6: *When third parties know worker's circumstances, the larger is the difference in circum-*

[57]For example, third parties learn that correctly placing 200 sliders with a circumstances of 1000 gives 25% chance of reaching the target, and 50% chance if the circumstances is 1500. Third parties are told explicitly that "if two workers exert the same effort, the one who has better circumstances has a greater chance of reaching the target."

[58]When pairs are randomly formed, the case in which the worker who succeeds in the pair has lower circumstances while the one who fail has better circumstances occurs naturally much less often. Hence, forming a pair in which the winner started with worse circumstances is less likely.

[59]In the case where the successful member of the pair has better circumstances, third parties believe that they also work more than those who fail on average (206 vs 186, $p < 0.05$). When the disadvantaged worker succeeds, third parties infer (properly in our data) that the winner exerted more effort.

Table 1.7: Redistribution With Known Circumstances

		All	Member who succeeds in pair has better circumstances	Member who succeeds in pair has worse circumstances
Mean redistributed share		0.380	0.403	**0.267**
		(0.257)	(0.236)	(0.328)
Median redistributed share		0.425	0.45	0.15
Share split 50-50		0.299	**0.338**	0.103
Share do not redistribute		0.149	0.103	**0.379**
Average circumstances	if succeed	1525	1698	663
	if fail	761	655	1290
Mean belief about effort (Truth)	if succeed	214 (245)	**206** (242)	**254** (261)
	if fail	187 (198)	186 (189)	192 (241)
Observations		174	145	29

stances between the workers in the pair, the more they redistribute income. However, third parties redistribute less if the worker who succeed in the pair starts with worse circumstances: 38% of third parties do not redistribute in that case. This is suggestive of meritocratic behavior.

Table 1.8: Redistributed Share

	(1)
Difference in circumstances	0.481
	(0.167)
Belief difference in effort	−0.016
	(0.016)
Belief difference in luck	0.044
	(0.698)
Third party's outcome	−6.574
	(3.724)
Effort only redistribution	0.328
	(0.065)
Luck only redistribution	0.421
	(0.107)
Subject FE	Yes
Groups	87
Observations	174

$p < 0.10$, $p < 0.05$, $p < 0.01$. OLS with Subject Fixed Effects.

Notes: To improve readability of the coefficients, the difference in circumstances is in hundreds. It is the difference between the circumstances of the worker who succeeds and the circumstances of the worker who fails. The third party's outcome is a dummy variable equal to 1 if the third party succeed in the effort task as a worker in stage 3. The belief difference in effort (luck) is the difference between the expected effort (luck) of the worker who succeeds and the worker who fails.

1.5 Conclusion

In this paper, I show that individual circumstances have important effects on effort, people's beliefs about what causes success or failure, and therefore redistribution. I introduce a theoretical model in which the worker's production depends on effort, circumstances and luck. In real life, we tend to believe that success is positively correlated with effort. When I introduce circumstances, success is instead negatively correlated with effort. In my environment, I show that the optimal effort is decreasing in circumstances. Workers with better circumstances are more likely to succeed, which reduces the optimal effort. Hence, the rational belief is that average effort of those who succeed is lower than those who fail. Provided that third parties are able to make the right inference, this should affect the redistribution of people with meritocratic preferences.

I run a real-effort experiment using the slider task introduced by [8] to see how workers adjust their level of effort to various circumstances, and if third parties are able to draw the right inferences about why people succeed or fail. Finally, I look at the way those inferences affect third parties' propensity to redistribute income. Workers in the experiment adjust their effort level to the circumstances they receive. There is a negative correlation between effort and circumstances in the top half of the circumstances distribution. Workers who have circumstances values in the top decile of the distribution exert on average less effort, yet succeed more. I also observe lower levels of effort when workers have unfavorable circumstances. This is partly driven by a significant fraction of subject who choose to provide no effort when they have unfavorable circumstances, which I call the *demotivation effect*.

Circumstances also complicate inferences about what causes inequality. When there is heterogeneity in circumstances, two-thirds of participants believe that workers who succeed on average have better circumstances and exert more effort than those who fail. When third parties observe only the outcome and not effort or luck directly, beliefs have little effect on redistribution. However, when third parties observe differences in circumstances, they take them into account when they redistribute income. The larger the circumstances difference between workers, the higher the level of redistribution. But if the successful worker started with lower circumstances, 38% of third parties do not redistribute income. This behavior, which compensates for luck and rewards effort, is consistent with meritocratic preferences.

By modifying the parameters $\overline{\lambda}$ and δ, the model offers the flexibility to adjust which factor affect the agent's outcome, and generates income inequality. For instance, with identical circumstances and little variance in luck, the workers need to exert high effort to succeed. With high heterogeneity in circumstances and low variance in luck, the workers more likely to succeed are those who have better circumstances. With identical circumstances and high luck variance, the workers more likely to succeed are those with higher luck draws, though some effort is still a necessary condition for success. Each of these configurations may have affect worker's effort, third parties inference, and therefore redistribution.[60]

This paper stresses a crucial point: whether people believe that they live in a society in which there are equal or unequal circumstances will affect their beliefs about what causes inequality. Two people with the same fairness preferences may make very different inferences about what causes income inequality, if one believe that everyone starts with equal chance to succeed whereas the other thinks circumstances are heterogeneous. Related to circumstances, perceptions of social mobility will interact with the beliefs about the source of inequality to affect redistribution [21, 22]. Those who think that social mobility is high believe that everyone has a chance to succeed, even if they start with different circumstances. Mobility is for [23] *"an equalizer of opportunities*".[61] Children of poor families can become rich through hard work (upward mobility), children of rich families can become poor if they do not work hard enough (downward mobility), therefore less redistribution is needed [24]. Our model can also be used to incorporate social mobility by making success feasible only to workers who receive favorable circumstances.[62] This can provide a fruitful avenue for future theoretical and experimental research.

---

[60]Case (1) looks like a meritocratic society in which success is more likely for agent's who exert higher effort, while case (2) resemble a oligarchic society with low mobility, where success depends on individual's exogenous circumstances.

[61]They note that mobility matters especially since "*it helps attenuates the effects of disparities in initial endowments, or social origins, on future income prospects*"

[62]This can be done by weighting the circumstances by a factor greater than one, or by increasing both the target production and the heterogeneity in circumstances

Chapter 2: Correlation Neglect and Overconfidence

Correlation neglect occurs when a decision maker underweights the correlation between signals. We extend a model by [25] and focus on the interaction between correlation neglect and one type of overconfidence, namely overprecision. The model predicts that high neglecters value extra signals more, seek more information than other agents, and are more overconfident. We design and run a laboratory experiment closely replicating the theoretical set-up, which involves endogenous acquisition of highly correlated signals. We identify individual level of correlation neglect using two incentivized methods. Full neglecters in our experiment are willing to pay on average 51% more for three extra signals than other participants. We find that high correlation neglecters are slightly more overprecise than non neglecters. We rule out that the observed behavior is driven by base-rate neglect.

Keywords: Correlation neglect; overconfidence; overprecision; beliefs; experiments

JEL Codes: C91, D83, D91

2.1 Introduction

[26] define overconfidence as three distinct phenomena: overestimation, overplacement, and overprecision. The first is overestimating one's ability, performance, or more generally chance of success. Overestimation of one's skills has been used to explained a variety of phenomena, from excess entry on markets [27] to the high degree of entrepreneurial failures [28]. The second form of overconfidence is overplacing one's performance compared to others, and believing to be better than average. For instance, comparing themselves to other participants in an experiment, around 80% of subjects state that their driving skills is above the median [29]. Our paper focuses on the last type of overconfidence, which is being overprecise about the accuracy of one's belief. People overestimate the precision of what they know. For example, traders and CFO's make predictions about stock market valuation that are often too precise [30, 31]. Overprecision has been shown affect areas such as firms' investment decisions or innovation choices [32, 33].

Yet, if much is known about the effect of overconfidence on decision-making in various domains, little is known about its causes. In the economic literature, [34] consider self-confidence both as an asset leading to better performances, and as a signaling device. [26] explain that, when people have imperfect information about themselves and even noisier information about others, bayesian belief updating can lead to overestimation and overplacement. However, their model can not account for overprecision. Given the pervasive effect of overprecision on economic decision making, identifying its causes might uncover ways to reduce it.

Correlation neglect is a probabilistic mistake made by agents who underestimate or even completely ignore the correlation of information. We borrow the term from [25]. An extreme case of correlation neglect is that of an agent who receives identical pieces of information and treats them as independents. Some concepts closely related to correlation neglect appears in the literature under different names. [35] explain polarization of opinions in a network setting by the fact that people do double-counting of information. In social learning, [36] show that herding may arise as a consequence of ”redundancy neglect”, where people fail to take into account that their predecessors action are correlated. Correlation neglect is thought to play a role in a wide range of settings from belief polarization [35], voting behavior [25, 37, 38], social learning [36], individual financial investments [39], school choice [40], and motivated reasoning [41]. In addition, it may provide an explanation for overprecision. Simply put, bayesian agents may hold overprecise posterior beliefs if they underestimate the correlation between signals.

The degree of an agent's overprecision is often considered exogenous in the literature, with some people being more overprecise than others. By contrast, overprecision arises as the result of correlation neglect in [25]. Agents underestimate the correlation between signals and believe that the information they receive is more informative than it actually is, which leads to overprecision. The model, which focus on political behavior and voting, predicts that overconfidence increases with the number of signals received and leads to political polarization if the signal correlation is high enough. A corollary is that more overconfident people will tend to acquire more information since they believe that additional signals are more informative than they actually are. One goal of this paper is to test experimentally the implications of [25]'s findings which pertain to overprecision, and in particular investigate if correlation neglect is positively correlated with overprecision.

There is a growing experimental literature on correlation neglect. [42] try to identify whether people ignore correlation because they don't detect it, or because they are unable to perform the mathematical operations to account for it. While some participants in their experiment form and update beliefs in a bayesian way, others fully neglect correlations. Differences in ability and the task complexity seem to affect individual levels of correlation neglect [43] shows the presence of correlation neglect in a social learning setting where subjects have to infer from their predecessors' actions, who move in group. In a laboratory experiment simulating financial investments, people appear to ignore correlation between assets in their portfolio choice [39]. Our paper provide additional evidence for correlation neglect. In our highly correlated environment (the correlation between signal is .8), subjects behave as if they perceive a correlation lower than .2 in 40% of our observations. We also contribute to the literature by providing two incentivized methods to estimate individual level of correlation neglect. The first method uses the variance of the posterior belief. The second method exploit participant's valuation for additional signals.

More broadly, this paper contributes to the growing body of literature which documents mistakes in the way people think about probabilities (see [44] for a detailed review) and form beliefs, such as nuisance neglect [45], sample selection neglect [46].

2.2 Theoretical Model

2.2.1 Basic Setup

We follow [25]. There is a state of the world $\theta \in \mathbb{R}$, which is drawn once from a normal distribution with mean 0 and precision τ. Each individual *i* must take an action $a_i \in \mathbb{R}$ and has a quadratic loss utility function:[1]

$$U\left(a_i \mid \theta\right) = -\left(a_i - \theta\right)^2$$

All agents have the same correct prior about the state θ and learn about it over time, with signals $s_{it} = \theta + \varepsilon_{it}$, $\varepsilon_{it} \sim \mathcal{N}[0,1]$. Each agent receives one signal per period. The main feature is that signals are positively correlated such that $\operatorname{corr}\left[\varepsilon_{it}, \varepsilon_{it'}\right] = \rho \in (0,1]$, but each agent *i* underestimates the correlation and instead perceive a correlation of ρ_i. This is what we call *correlation neglect*. Formally, $\operatorname{corr}\left[\varepsilon_{it}, \varepsilon_{it'}\right] = \rho_i < \rho$ with $\rho_i \in [0, \rho)$. Thus, the degree of correlation neglect of an agent is $\rho - \rho_i$, and may differ among individuals.[2]

---

[1] In the original paper, the utility function is $U\left(a, b_i \mid \theta\right) = -\left(a - b_i - \theta\right)^2$. Agents are voters who must take an optimal action a (vote for a policy) given the state of the world θ and an individual preference shifter b_i.

[2] The authors show that the posterior belief follows a normal distribution with mean $\frac{\sum_{t=1}^{n_i} s_{it}}{n_i+\tau(1+(n_i-1)\rho_i)}$, and variance $\frac{1+(n_i-1)\rho_i}{n_i+\tau(1+(n_i-1)\rho_i)}$.

2.2.2 Generalization

We generalize the model to accommodate cases where the state mean is non-zero, and the signals error variance is non-unitary. So, agents have a prior about the state θ which is distributed normally with mean θ_0 and variance σ_0^2. Each signal s_i is distributed with mean θ and variance σ^2. Then, after receiving n signals, the posterior variance of an agent who perceives a signal correlation of ρ_i is (see proof in Appendix B.) :

$$Var(\theta, p_i) = \frac{\sigma_0^2\, \sigma^2(1+(n-1)\rho_i)}{\sigma_0^2\, n + \sigma^2(1+(n-1)\rho_i)} \tag{2.1}$$

Proposition 1 (variance comparative statics):

1. *Fixing the perceived correlation ρ_i, the posterior variance is decreasing convex in the number of signals received n.*
2. *Fixing n, the posterior variance is increasing concave in ρ_i. Hence, it is decreasing convex in correlation neglect $\rho - \rho_i$.*

Proof. See Appendix B.

Intuitively, the posterior variance decreases as the agent gets more information. The first signals bring a lot of information which contributes to a large reduction in posterior variance. As an agent receives more signals, the informative content of each gets smaller and smaller. The agent's posterior variance keeps decreasing, but at a slower rate. The second part of proposition 1 says that the posterior variance is increasing in ρ_i. Therefore, it is decreasing in correlation neglect. The more an agent perceives that the environment is correlated, the larger is their posterior variance. Since the agent perceives that signals are correlated and not very informative, their belief is less precise and they have a larger posterior variance. On the contrary, the more the agent underestimate correlation - the higher their degree of correlation neglect - the smaller will be the posterior variance.

Proposition 2 (variance asymptotic behavior):

As n goes to infinity, the variance limit is:

$$\lim_{n \to \infty} Var(\theta, \rho) = \frac{\sigma_0^2 \sigma^2 \rho}{\sigma_0^2 + \sigma^2 \rho}$$

Proof. See Appendix B.

Because of the correlation between signals, there will always remain some uncertainty about the state even if an agent receives infinitely many signals. This holds for a neglecter, but also for a Bayesian who correctly perceives that $\rho = \rho_i$. By contrast, if there were no correlation between signals (i.e., $\rho = 0$), a Bayesian would eventually end up learning the state, and have a posterior variance of 0. Note that the posterior variance limit is increasing in ρ, so it is smaller for high neglecters.

2.2.3 Marginal Benefit of an Extra Signal

Proposition 3: *After receiving n signals, the marginal benefit of an extra signal is:*

$$Var(\theta, \rho_i) - Var(\theta + 1, \rho_i)$$

Proof. See Appendix B.

With a quadratic loss utility function, the benefit of an additional signal is that it reduces the variance (equivalently, increases the precision) of the posterior belief. After being exposed to n signals, the agent will buy the extra piece of information if the marginal benefit exceeds the marginal cost, that is, if the expected reduction in posterior variance is larger than the cost c of acquiring the extra signal.

We established in proposition 1 that the posterior variance is decreasing convex in n, conditional on ρ_i. Hence, the marginal benefit of an extra signal is decreasing in n. In other words, the value of an extra signal is decreasing as more pieces of information are received. Each additional signal brings less and less information to the agent. Next, the posterior variance is increasing concave in ρ_i, conditional on n. It follows that the marginal benefit is decreasing in ρ_i, and so increasing in correlation neglect. Intuitively, the lower the perceived correlation between signals, the more appealing is each extra signal since it appears more informative for the agent. On the other hand, the less the agent neglects correlation, the lower will be the perceived marginal benefit of the extra information. An agent who properly perceives that signals are highly correlated knows that one additional signal contains little information and therefore value them less. In the extreme, if signals are fully correlated (identical), then all the information is contained in one signal and any additional signal is worthless.

2.2.4 Overconfidence and Number of Signals

The *perceived variance* about θ of an agent who perceives a correlation of p_i, after they have received n signals is denoted $Var(\theta, \rho_i)$. If an agent is bayesian and does not neglect correlation ($\rho_i = \rho$), their posterior variance after receiving n signals is $Var(\theta, \rho)$. We call this the *true variance*.

Definition 1 (overconfidence): *The level of* ***overconfidence*** *of agent i after receiving n signals is the difference between the variance of a bayesian (non neglecter) decision maker, and the agent's perceived variance.*

$$OC_n = Var(\theta, \rho) - Var(\theta, \rho_i)$$

An equivalent definition of overconfidence can be formulated using the posterior belief precision. Then, overconfidence is the difference between the agent perceived precision and the precision of a non neglecter. This makes it clear that the focus of our paper is on overprecision.[3] How does overconfidence change with the number of signals received by an agent, and how does

[3] In the rest of this paper, we use the words overconfidence and overprecision interchangeably.

overconfidence change with correlation neglect?[4] To fix ideas, table 2.1 describes an agent's posterior variance after receiving n or $n + 1$ signals.

We showed in proposition 3 that with a quadratic loss utility function, the marginal benefit of an extra signal is the decrease in variance brought by the additional information.[5] Therefore, the *perceived value* of an extra signal is the difference between the perceived variance after receiving n and $n + 1$ signals. The *true value* of an extra signal is the decrease in variance brought by the additional piece of information for a non neglecter.

Table 2.1: Posterior Variance Conditional on the Number of Signals Received

	Number of signals received				**Marginal benefit of extra signal**	
	n		**n+1**			
Perceived variance	$Var(\theta, \rho_i)$	**(1)**	$Var(\theta + 1, \rho_i)$	**(2)**	Perceived value	**(1) − (2)**
True variance	$Var(\theta, \rho)$	**(3)**	$Var(\theta + 1, \rho)$	**(4)**	True value	**(3) − (4)**
	$OC_n = (3) - (1)$					

Proposition 4 (overconfidence comparative statics):

1. *Fixing ρ_i, overconfidence is increasing in n.*
2. *Fixing n, overconfidence is decreasing in ρ_i. Hence, it is increasing in correlation neglect $\rho - \rho_i$.*

Proof. See Appendix B.

The first part of proposition 4 states that correlation neglecters become more overconfident as they receive more pieces of information. In other words, as agents gets more pieces of information, the difference between the agent's variance and the bayesian variance increases. Correlation neglecters deviate more and more from the true bayesian variance. This is due to the fact the variance is decreasing convex in correlation neglect. Not only a correlation neglecter has a lower variance

[4]For now we consider the signals as exogenous. That is, the agent does not choose the number of signals they receive. In case of endogenous signal acquisition, they may choose how many pieces of information to purchase.

[5]Put another way, the more a piece of information increases the precision of the posterior belief, the more it is valuable.

for any given n, but they always reduce their variance more than a bayesian upon receiving an extra signal. The second part of the proposition states that conditional on n, the more an agent neglect correlation, the higher will be their overconfidence.

Proposition 5 (overconfidence asymptotic behavior): *Overconfidence is monotonically increasing in correlation neglect.*

Proof. We show in proposition 2 that the variance is bounded:

$$\lim_{n\to\infty} Var(\theta,\rho) = \frac{\sigma_0^2\,\sigma^2\,\rho}{\sigma_0^2 + \sigma^2\,\rho}$$

So, as n goes to infinity, the overconfidence converges to:

$$\lim_{n\to\infty} OC_n = \frac{\sigma_0^2\,\sigma^2\,\rho}{\sigma_0^2 + \sigma^2\,\rho} - \frac{\sigma_0^2\,\sigma^2\,\rho_i}{\sigma_0^2 + \sigma^2\,\rho_i} = \frac{\sigma_0^4\,\sigma^2\,(\rho - \rho_i)}{(\sigma_0^2 + \sigma^2\,\rho)(\sigma_0^2 + \sigma^2\,\rho_i)}$$

The limit is monotonically decreasing in ρ_i. Hence, overconfidence is monotonically increasing in correlation neglect. Note that when there is no correlation neglect ($\rho_i = \rho$), the overconfidence is 0. That is, the decision maker perceived posterior variance is the true variance.

2.2.5 Benchmark Cases

To illustrate how the model works, let's look at the case of two agents who are identical except in their correlation neglect level. Assume that the environment is one in which signals are highly correlated with $\rho = 0.9$. One agent perceives a correlation of $\rho_i = .8$, and so is a low correlation neglecter ($\rho - \rho_i = .1$). The second agent is a high neglecter who perceives a correlation of $\rho_i = .1$ ($\rho - \rho_i = .8$). Assume for simplicity that $\sigma_0^2 = \sigma = 1$, and that agents can decide how many signals to buy for a fixed unitary cost of $c = .01$. In the tables below, all variances are posterior variances conditional on receiving n signals.

n	Perceived Var. (1)	Var. reduction (2)	Cost (3)	Gain (2)−(3)	True Var. (4)	Overconfidence (4)−(1)
2	.474				.487	.013
3	.464	.009	.01	−.01	.483	.018
4	.459	.005	.01	−.05	.481	.021

n	Perceived Var. (1)	Var. Reduction (2)	Cost (3)	Gain (2)−(3)	True Var. (4)	Overconfidence (4)−(1)
2	.355				.487	.132
3	.286	.069	.01	.059	.483	.197
4	.245	.040	.01	.03	.481	.235
5	.219	.027	.01	.017	.479	.260
6	.200	.019	.01	.009	.478	.278
7	.186	.014	.01	.004	.478	.292
8	.175	.011	.01	.001	.477	.302
9	.167	.009	.01	−.001	.477	.310

Comparing the benchmark cases, we see that:

(i) Fixing n, the perceived posterior variance is decreasing in correlation neglect. For example, after receiving 2 signals, the low neglecter perceived variance is .474, whereas is it .355 for the high neglecter.

(ii) Fixing n, the signal valuation (i.e., how much variance reduction is expected from an extra signal) is increasing in correlation neglect. For instance, after receiving $n = 3$ signals, a fourth signal decreases the variance of a low neglecter by .009, and that of a high neglecter by .069.

(iii) Given a fixed unitary cost c, the optimal number of signals to buy is increasing in correlation neglect. In our example, a low neglecter finds it optimal to acquire two signals, as the cost of the third signal is higher than its marginal benefit. The optimal number of signal is eight for a high neglecter.

(iv) Fixing n, overconfidence is increasing in correlation neglect. Conditional of having received three signals, the overconfidence level of a low neglecter is .018, whereas it is of .197 for a high neglecter.

(v) Conditional on ρ_i, overconfidence increases as agents gets more signals. Take a high neglecter for example, their overconfidence is of .132 after receiving two signals, increases to .197 after their third signal, then to .235 upon receiving a fourth signal etc.

To sum up, with a quadratic loss utility function, the marginal benefit of extra signals is the expected reduction in posterior variance. This implies that compared to low or non neglecters, high neglecters value extra signals more (H1), and will seek more information than other agents (H2). Intuitively, a correlation neglecter believes that signals are more informative than they actually are, is willing to pay more for each individual signal, and acquire more of them. In addition, if agents can observe each other's information seeking behavior, high neglecters would appear more informed since they consume more information. Hence, we conjecture that they become more influential. When signals are correlated, some uncertainty about the state still remains even if an agent receives infinitely many signals. Hence, due to correlation, the posterior variance do not converge to zero even for a bayesian. We show that overprecision is monotonically increasing in correlation neglect (H3).

2.3 Experiment Design

From the model, we derive the following testable hypotheses:

1. Valuation: fixing n, the signal valuation increases with the degree of correlation neglect (H1)
2. Fixing a unitary cost, high neglecters seek more information (H2). In addition, we conjecture that they become more influential as they seem more informed.
3. Overconfidence (H3)
   a) fixing n, overconfidence is higher for high neglecters (H3a)
   b) fixing ρ_i, overconfidence increases as agents gets more signals (H3b)

In order to test the model, we need an experimental design which allow us to set the true signal correlation (ρ) and identify the correlation perceived by agents (ρ_i).

2.3.1 Implementation

The experiment was conducted at the Missouri Social Science Experimental Laboratory at Washington University in St. Louis, after IRB approval. We run eight sessions of the experiment in September 2021. A total of 58 subjects recruited using ORSEE participated in the experiment. The experiment was coded using oTree [20]. The average earning was \$18.57. Each session including the instructions lasted around one hour.

2.3.2 Experiment Outline

Each participant took part in seven round. Each round was comprised of four stages. In each round, we informed participant that an imaginary box had been randomly selected among a thousand boxes, and that we were interested in the weight of that box. The weight of the boxes is normally distributed with mean 5 000 and variance 250 000. No statistical knowledge was required and participants were given both a verbal and visual description of the distribution in simple terms. For example, ”*around 14% of boxes weight between 4500 and 5000 lbs., less than 2% of boxes weight under 4000 lbs....*” (see the instructions in appendix B.4). In each round, the task of participants was to estimate the weight of the randomly chosen box, with reports provided to them. Each report j in round i in is the sum of three numbers:

$$r_{i,j} = w_i + c_i + s_{i,j}$$

,where $w_i \sim \mathcal{N}[5\,000, 250\,000]$ is the state (i.e., the weight of the box) which is randomly drawn in round i. $c_i \sim \mathcal{N}[0, 40\,000]$ is a common bias affecting all the reports in round i, and $s_{i,j} \sim \mathcal{N}[0, 10\,000]$ is the report j specific noise in round i. Here, $c_i + s_{i,j}$ plays the role of the noise term from our theoretical set up. Introducing a common bias is a way to create correlation across noise terms. The theoretical correlation across reports in a given round is then .8.[6] All numbers, including the box weight, were randomly generated before the start of the first session using MATLAB, and remained identical for all sessions and participants (see table 2.2). At the end of the instructions,

[6] $\rho = \frac{40\,000}{40\,0000+10\,000} = 0.8.$

which lasted around thirty minutes, participants had to answer seven comprehension questions correctly before proceeding. In addition, after reading the instructions, we gave all participants a summary sheet including the visuals of the distributions, which they can refer to at any point.[7]

Table 2.2: Experiment Parameters

	Round i						
Parameter	1	2	3	4	5	6	7
State (w_i)	5305	4513	5421	5196	4482	4898	4942
Common bias (c_i)	247	51	−398	−234	24	−50	360
Noise 1 ($p_{i,1}$)	6	−115	−41	130	−86	−220	−32
Noise 2 ($p_{i,2}$)	−147	55	191	−59	−17	−77	82
Noise 3 ($p_{i,3}$)	−163	157	−39	44	−19	−139	49
Noise 4 ($p_{i,4}$)	−196	−169	41	−50	−87	−39	77
Noise 5 ($p_{i,5}$)	261	−45	−114	10	18	53	78
Noise 6 ($p_{i,6}$)	97	−8	−62	120	127	152	−148
Report 1 ($r_{i,1}$)	5558	4449	4982	5092	4420	4628	5270
Report 2 ($r_{i,2}$)	5405	4619	5214	4903	4489	4771	5384
Report 3 ($r_{i,3}$)	5389	4721	4984	5006	4487	4709	5351
Report 4 ($r_{i,4}$)	5356	4395	5064	4912	4419	4809	5379
Report 5 ($r_{i,5}$)	5813	4519	4909	4972	4524	4901	5380
Report 6 ($r_{i,6}$)	5528	4556	4961	5082	4633	5000	5154

Note: recall that $r_{i,j} = w_i + c_i + p_{i,j}$. For example, report 2 in round 1 is $r_{1,2} = w_1 + c_1 + p_{1,2} = 5305 + 247 - 147 = 5405$. The state, common bias and noises were randomly drawn once and for all before the first experimental session. Reports were identical across participants and across experimental sessions to allow for between-subject comparisons.

2.3.3 Stages

Each round was comprised of four stages. In *Stage 1*, every participant received two reports for free, and then provided a first estimate of the weight of the box. After giving their estimate, they were asked for a 90% confidence interval around it. The 90% confidence interval elicitation was done in an incentive compatible manner, using a random binary choice menu (See figure 3.2).

Stage 2 (bid for one report): Each participant submitted a bid for one extra report. We used a Becker-DeGroot-Marschak procedure to determine if the participant won the bid. A different

[7] We did not give it earlier to make sure to have the subject's full attention while giving them the instructions verbally.

price was randomly generated for each participant. Hence, they were told that they were not in competition with others to submit the highest bid, and that it was in their interest to submit their true valuation for the extra report.[8] We choose to ask first for the valuation for a single extra report in order to control for potential individual heterogeneity. Some participants with the same level of correlation neglect may bid less than others for reasons that are not related to their correlation neglect level (e.g., if they are "cheap"). In addition, by asking right away to bid for several reports, we would have run the risk that some reports may be worthless to some participants. For example, after seeing two reports, a non neglecter should value one extra report at a mere \$0.05. If the participant won the bid, they saw the extra report and were able to revise their estimate and 90% confidence interval.

Stage 3 (bid for three reports): The procedure is identical to stage 2, but this time each participant submitted a bid for three reports. In case they win, a participant saw the three reports and was able to revise their estimate and confidence interval. We choose three extra reports as this would appear to provide a significant amount of extra information to a high neglecter, though the low neglecter would find them of little values.

Stage 4 (nomination): The stage was designed to test hypothesis 2. Participants can pay a small cost of \$0.1 to observe the bids placed by four other subjects in that round, and are shown the latest estimate and confidence interval of one subject of their choosing (their nominee).[9] Upon seeing the estimate and confidence interval of their nominee, participants provide a final estimation of the box weight and report their confidence. The cost of \$0.1 was chosen so that, given our payoff function, it was worth to enter stage 4 for everyone who lost both bids, to the exception of the non neglecters. This is assuming that their nominee won the bid for three extra reports and made proper use of that information for their own estimation.

[8]The possible bidding range was restricted between \$0 and \$2. In order not to inflate participant's bid, we specifically told them: "It is best under this system to bid your true maximum willingness to pay. You will never pay more and can end up paying less. If you value the report at \$2, you should bid \$2. If you value the report at \$0, you should bid \$0."

[9]The four subjects were randomly selected among all subjects who took part in previous sessions of the experiment. For the first session, we used the data of pilot participants. Before choosing a nominee, participants only see the bids and not whether the bid was actually won or not.

In total, in each round, all participants saw at least two reports, and at most six reports (in case they win both bids). Note that the reports given are identical across participants in a given round. The first two reports are the same for all participants, all the winners in stage 2 saw the same report, and all the winners in stage 3 saw the same three reports. This allow us to make between subject comparison of correlation neglect and overconfidence level. In theory, the actual reports values should not affect participants posterior variance (they should of course affect their mean belief), but it is possible that someone seeing two nearly identical reports would be more confident in their estimate. We wanted to avoid this by giving everyone the same reports in each stage. Participants were informed that the reports were identical for all, because this could influence their willingness to participate in the nomination stage.

In terms of design choice, we ask for participant's willingness to pay for extra reports as a continuous measure of valuation. The alternative - asking for a number of signals the participants wish to buy - would have been a coarser measure of valuation. Next, we choose to separate stage 2 and 3 for two main reasons. We ask first for participant's willingness to pay for one extra signal in order to control for participant's preference, as some participant may be cheap and always provide low bids, while some others may be systematically give high bids. In addition, a third report is already worth very little for low neglecters in a highly correlated environment like the one of our experiment where $\rho = .8$. Hence, we would not be able to capture this by asking directly the willingness to pay for four extra reports. We paid participants both for their estimate and their confidence interval. A quadratic scoring rule was used to determine the payoff in round i for the estimated weight $\hat{w_i}$:

$$2 - 0.002\% \ (\hat{w_i} - w_i)^2$$

Hence, the closer the estimate is from the true weight of the box w_i, the higher the payoff. This should encourage people to try to improve their estimate in each stage.[10] In order to obtain participants' perceived variance, we elicited 90% confidence interval around their estimate. This

[10]With this payoff function, participant earned \$2 for a correct estimate, \$1.8 for an estimate 100lbs. away from the true weight, \$1.2 for an estimate 200 lbs. away from the true weight. Subjects were informed that they could not loose money if their estimate was too far (more than 300 lbs.) from the truth.

was done using a random binary choice table which give the choice between 90% chance of $1 or 1$ if the estimate lies in a given interval around the true weight (from ±500 to ±600 lbs, by 10 lbs. increments). Hence, a subject made eleven choices, one for each increment. One of their eleven choice was randomly selected to determine their payoff. In each round, we selected one of the four stage randomly and paid subjects for their estimate and confidence interval in that stage. Thus, each answer potentially mattered for subjects.[11]

Choice	**Option A**			**Option B**
1	90% chance of $1	◉	○	If your estimate is within ± 500 lbs of the weight of the box you get $1
2	90% chance of $1	◉	○	If your estimate is within ± 510 lbs of the weight of the box you get $1
3	90% chance of $1	◉	○	If your estimate is within ± 520 lbs of the weight of the box you get $1
4	90% chance of $1	◉	○	If your estimate is within ± 530 lbs of the weight of the box you get $1
5	90% chance of $1	○	◉	If your estimate is within ± 540 lbs of the weight of the box you get $1
6	90% chance of $1	○	◉	If your estimate is within ± 550 lbs of the weight of the box you get $1
7	90% chance of $1	○	◉	If your estimate is within ± 560 lbs of the weight of the box you get $1
8	90% chance of $1	○	◉	If your estimate is within ± 570 lbs of the weight of the box you get $1
9	90% chance of $1	○	◉	If your estimate is within ± 580 lbs of the weight of the box you get $1
10	90% chance of $1	○	◉	If your estimate is within ± 590 lbs of the weight of the box you get $1
11	90% chance of $1	○	◉	If your estimate is within ± 600 lbs of the weight of the box you get $1

Figure 2.1: Elicitation of 90% Confidence Interval with a Binary Choice Menu

2.3.4 Identification Strategies for Correlation Neglect

We use two different and complementary ways to identify a subject correlation neglect level ρ_i. One method uses participant's reported 90% confidence interval, and the other uses their bids.

[11] If a subject did not participate in the selected stage, we used their latest estimate and confidence interval to calculate their payoff. For example, if stage 4 was randomly selected for payment but the subject did not participate, then their stage 3 answer was used, provided they won the bid. If they did not win the bid in stage 3, then their stage 2 estimate was used etc.

Both measures were incentivized, and are done after the first stage, i.e., when every participant has received two signals.

Using Confidence intervals

The first method uses the size of the reported 90% confidence interval to estimate the perceived posterior variance, denoted $\hat{S_i}$. With a normal prior and normally distributed signals, the subject posterior is normally distributed. Ninety percent of the mass of a normal distribution lies ± 1.64 standard deviation from the mean. Hence, the reported 90% confidence interval should corresponds to around 3.3 standard deviations. For example, if the subject reported 90% confidence interval is of size 1000, then their standard deviation is around 303 (1000/3.3), and their variance is $\hat{S_i} \approx (303)^2$.[12] Once we have recovered the perceived posterior variance, we use the theoretical posterior variance derived in equation (2.1) to estimate each participant perceived correlation (ρ_i) after stage 1.

$$\rho_i = \frac{\sigma_0^2(\hat{S_i} n_i - \sigma^2) + \hat{S_i}\sigma^2}{\sigma^2(n_i - 1)(\sigma_0^2 - \hat{S_i})}$$

Finally, the last step is to compute the subject level of correlation neglect: $\rho - \rho_i$. Table 2.3 below presents the estimated correlation neglect level using our methodology. With our parameter choices, in particular $\rho = .8$, the non neglecter standard deviation after $n = 2$ reports should be 360. Hence, we expect the non neglecter to switch from option A to option B at row 9 in our binary choice table.

Using Bids

Next, we use people's bid to estimate their correlation neglect level. With the quadratic loss payoff function specified in our experiment, the marginal benefit (and hence, optimal bid) for k extra signals, conditional on having observed n signals solves:

$$b = 0.002\%[Var(\theta) - Var(\theta + k)]$$

The correlation neglect level is estimated based on the bid for the first extra signal. For example, conditional on having seen two signals, the non neglecter theoretical marginal benefit for one extra

[12]We assume that the elicited 90% confidence interval is symmetric around the mean belief.

Table 2.3: Identification of Perceived Correlation Using the Reported 90% Confidence Interval

Switching row	Size of 90% CI	Size of one SD	ρ_i	CN: $\rho - \rho_i$	OC level
1	1000	303	-0.03	0.83	57
2	1020	309	0.03	0.77	51
3	1040	315	0.10	0.70	45
4	1060	321	0.17	0.63	39
5	1080	327	0.25	0.55	33
6	1100	333	0.33	0.47	27
7	1120	339	0.42	0.38	21
8	1140	346	0.52	0.28	14
9	1160	352	0.63	0.17	8
10	1180	358	0.74	0.06	2
11	1200	364	0.87	-0.07	< 0

Note: the column ”OC level” is the estimated overconfidence after stage 1. Since a non neglecter should perceive a standard deviation of 360 after receiving $n = 2$ reports, switching from option A to B in row 2 corresponds to a perceived correlation of 0.03 and to an overconfidence level of $360 - 309 = 51$.

signal is \$.047. Hence, any participant bidding less than \$.05 for a third signal is classified as non neglecter. On the other hand, full neglecter bid more than \$.41. We associate each bid value between \$.05 and \$.41 to a correlation neglect level.

2.4 Results

2.4.1 Summary Statistics

The correlation neglect levels measured using the confidence interval and the bids are not correlated (Pearson's r=$-$.06).[13] So, this could mean that they potentially measure a different aspect of correlation neglect. In the rest of the analysis, I define the level of correlation neglect of a participant as the average of the correlation neglect levels measured by their bid and the confidence interval. However, I also report the results for the measures separately for transparency. Using the composite measure of correlation neglect, the estimated density takes the form of a bimodal distribution (see figure 2.2), with around 40% of participants having a correlation neglect level between .25 and .5,

[13]Given the fact that the experiment was complex and contained a lot of instructions, we considered that the first round was a learning round for our participants and therefore dropped it from all our analyses.

and 40% of participants with a level between .6 and .8. Finally, about 10% of participants have a correlation neglect level below .25. [14]

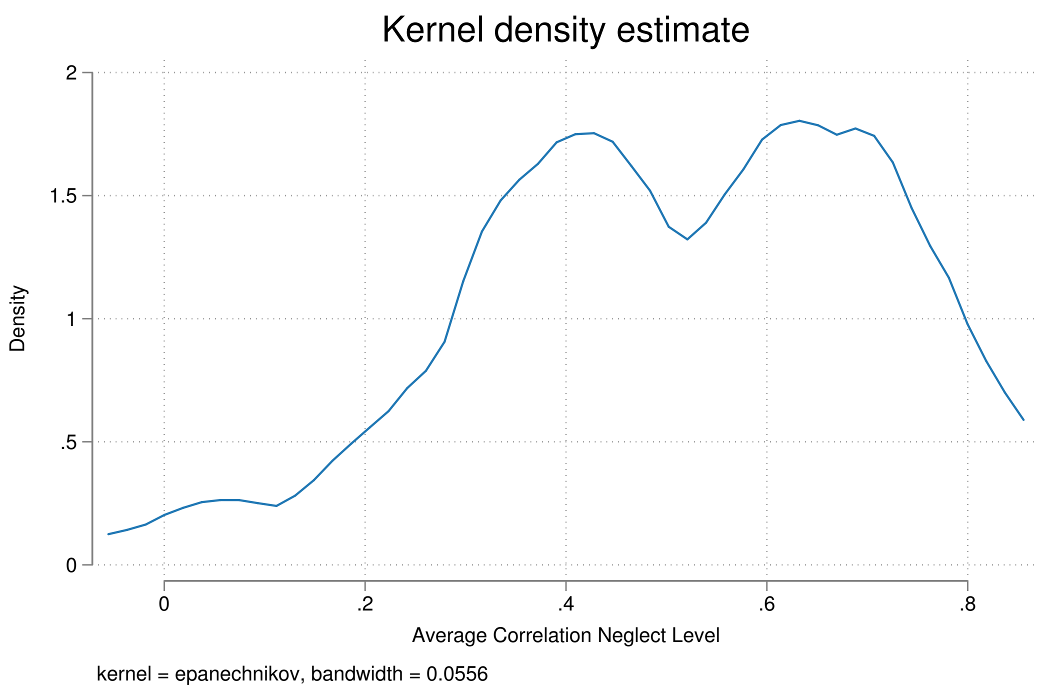


Figure 2.2: Kernel Density Estimate of the Correlation Neglect Level

The average correlation neglect is 0.51, and is fairly stable across rounds (see table 2.4). Within-subject, the correlation neglect level is also very stable, with an average subject standard deviation is of 0.09 across all rounds. This stability is in our opinion a good sign, because we expect correlation neglect to be fairly stable. We would find it surprising if the same subject behave like a non neglecter in one round, and a full neglecter in the next.[15]

The number of reports seen by a participant depends on how many bids they won. Table 2.5 indicates how many reports were seen by participants per round. Across all rounds, people lose both bids in 57% of the case, and thus only received the two initial (free) reports. On the opposite, participants won both bid and saw a total of six reports in 7% of the rounds. A total of five reports

[14]The density estimate for each separate measure of correlation neglect are presented in appendix B.2.

[15]This is not to say that we believe the context in which we measure correlation neglect is irrelevant. But as for other probabilistic mistake such as for instance base-rate neglect, we think it should be fairly stable in a given context or frame.

Table 2.4: Correlation Neglect Level per Round

Round	Mean	SD
2	0.51	0.20
3	0.54	0.18
4	0.52	0.20
5	0.51	0.20
6	0.49	0.21
7	0.49	0.21
Average	0.51	0.20

was seen in 25% of the rounds, which happened when participants only won the second bid. Finally, in 11% of the rounds, people only won the first bid and saw a total of three reports. Overall, participants win the bid for three extra reports more frequently than the bid for one report. The bids for three reports are won in 32% of the case, vs 18% for the bid for one report.

Table 2.5: Number of Reports Seen Per Round

Number of reports	Two	Three	Five	Six
Bids outcome	Lose both bids	Win first bid only	Win second bid only	Win both bids
Frequency	200 (57%)	38 (11%)	87 (25%)	23 (7%)

2.4.2 Signal Valuation and Correlation Neglect

Overall, participants bid significantly more for three reports than for one report ($p<0.01$, paired t-test). Placing higher bids for more reports is expected and provide evidence that participants did not behave randomly and understood the task. We observe a wide range of bids. The maximum bid for one extra report was \$1.88, and \$2 for three extra reports.[16]

Bid for three extra reports

We use a regression with subject fixed-effect to control for individual heterogeneity which could lead some participants to bid systematically less (e.g., if one is stingy) or more (say, if one is curious)

[16]participants were not allowed to bid more than \$2 which was the highest possible randomly generated price in our BDM procedure.

Table 2.6: Average Willingness to Pay for One and Three Reports

Willingness to pay	Mean	SD	Min	Max
One report	\$.33	\$.34	\$0	\$1.88
Three reports	\$.63	\$.47	\$0	\$2.00

than other subjects. The level of correlation neglect has a positive effect on the willingness to pay for three extra reports (p<0.01), controlling for the number of reports seen. Winning the first bid decreases the willingness to pay for the three extra reports (p<0.01).[17] A 0.1 increase in correlation neglect increases the willingness to pay by around 8% (\$0.04). When looking at correlation neglect categories, it is low neglecters - with a correlation neglect lower or equal to .2 - who bid significantly less than other participants (p<0.01).[18] Overall, this tends to support hypothesis 1. Full neglecters are willing to pay \$0.91 for three extra signals, or 51% more than other participants whose average bid is \$0.60 (two sample t-test, p<0.01). On the other hand, low neglecters are willing to pay \$0.36 vs \$0.65 for other participants (two sample t-test, p<0.01).

2.4.3 Correlation Neglect and Information Seeking

In this section, we focus on the decision to participate in stage 4. At most, if a subject has not received any additional reports (i.e., lost both bids), then participating in stage 4 should be valued the same as getting four extra reports. This assumes that they select a nominee that has won both bids and has correctly used the reports to make their estimation.[19] After seeing two reports, getting four additional reports is worth \$0.096 to a non neglecter, and \$1.04 to a full neglecter. Since the cost of entering stage 4 is 0.1\$, then a rational non neglecter should not enter. In fact, it is worth entering stage 4 (i.e., valued more than \$0.1) for all subjects with a correlation neglect level greater than .01.

[17] In Appendix B.2 table B.1, we present the details of the average bids conditional on the first bid outcome.

[18] The effect is not driven by the cutoff of .2, since we find significant negative effect with cutoffs placed at correlation neglect levels of 0.05, 0.1 and 0.25 (p<0.01, p<0.05 for a cutoff at 0.25).

[19] This does not mean that their nominee is necessarily a non neglecter. It simply means that their nominee have correctly updated their prior after receiving the new reports.

Table 2.7: Regression with Dependent Variable: Willingness to Pay for Three Reports

	(1)	(2)	(3)
CN level	0.397		
	(0.135)		
Win one extra report	−0.259	−0.253	−0.273
	(0.046)	(0.046)	(0.047)
Low Neglecter (CN≤0.2)		−0.229	
		(0.076)	
Full Neglecter (CN=0.8)		0.111	
		(0.081)	
CN using bids			0.284
			(0.085)
CN using confidence intervals			0.067
			(0.105)
Constant	0.473	0.678	0.505
	(0.069)	(0.019)	(0.072)
Subject FE	Yes	Yes	Yes
Groups	58	58	58
Observations	348	348	348

Note: Standard errors in parentheses. $p < 0.10$, $p < 0.05$, $p < 0.01$.

The participation rate in stage 4 is 48%, and 62% among subjects who do not win extra reports. We run a logit regression with the decision to participate as dependent variable (see Table 2.9). Overall, the probability of participating in stage 4 does not significantly increases with correlation neglect. However, winning the first bid decreases the probability of participating in stage 4 by 26%, and winning the second bid decreases the probability by 32% (Column (1), p<0.01).

Restricting to participants who did not win any extra reports (57% of our sample), subjects who choose to participate have on average a correlation neglect level of 0.50 vs 0.43 for those who choose not to participate (two-sample t-test, p<0.05). Among people who do not win any extra reports, who choose to participate? Table 2.8 displays the participation rate by categories of neglecters. Only 39% of low neglecters participate, whereas 64% of other subjects do (two sample t-test, p<0.05). Being a low neglecter decreases the probability of entering by 24%, though the effect is only significant at the 10% level (see column (4) in table 2.9). More striking, none of the non neglecter who lost both bids choose to participate, thus behaving rationally.[20]

Table 2.8: Stage 4 Participation Conditional on Receiving No Extra Reports

CN level	Participation Rate
CN≤ 0.2	39% (7/18)
0.2 <CN≤ 0.4	59% (25/42)
0.4 <CN≤ 0.6	68% (48/71)
0.6 <CN< 0.8	64% (38/59)
CN=0.8	60% (6/10)
Average	62% (124/200)

Table 2.9: Logit With Dependent Variable: Stage 4 Participation

	Full sample		*Lost both bids only*	
	(1)	(2)	(3)	(4)
CN Level	1.487		1.697	
	(0.988)		(1.118)	
Win one extra report	−1.182	−1.179		
	(0.393)	(0.397)		
Win three extra reports	−1.487	−1.486		
	(0.400)	(0.400)		
CN using bids		0.735		
		(0.704)		
CN using confidence interval		0.750		
		(0.689)		
Low Neglecter (CN≤0.2)				−1.051
				(0.558)
Full Neglecter (CN=0.8)				−0.193
				(0.786)
Constant	−0.241	−0.241	−0.295	0.599
	(0.500)	(0.503)	(0.542)	(0.289)
Observations	336	336	200	200

Note: Standard errors in parentheses, clustered at participant level.
$p < 0.10$, $p < 0.05$, $p < 0.01$

Overall, the non neglecters in our experiment behave as if they understand that the environment is highly correlated and judge that paying $.1 to observe another subject's estimate is not worth it.

[20]This observation has to be interpreted with caution since across all rounds, only 7 participants have both a correlation neglect level of 0 and lost both bids. Yet, this seems to suggest that non neglecter are reluctant to enter stage 4.

Subjects with a correlation neglect level higher than .2 are more likely to enter stage 4 compared to low or non neglecters.

2.4.4 Correlation Neglecters Become More Influential

If participants choose to enter stage 4, they are presented with a random sample of four potential nominees to choose from.[21] Stage 4 participants choose to observe the highest bidders. The average bid of displayed nominees is \$.56, whereas the average bid of chosen nominees is significantly higher at \$.91 (paired t-test, p<0.01, see table A.1).[22] However, the average correlation neglect level is quite similar for displayed and selected nominees, around .55. Since high neglecters bid more and the highest bidders are nominated, we would expect nominees to have higher correlation neglect level on average. We conjecture that this is what would occur with a larger sample. In our experiment, only 48% of subjects choose to participate in that stage, reducing our sample size.

Table 2.10: Average Bid of Displayed and Selected Nominees

	Average bid (SD)	Average CN Level (SD)
Displayed nominees	\$.56 (\$.21)	.55 (.07)
Selected nominee	\$.91 (\$.48)	.54 (.18)

Do people who choose to enter stage 4 update their belief following the observation of their nominee's estimate and confidence interval? Before entering stage 4, a participant's latest estimate was on average 95 lbs. away from their selected nominee. After seeing the nominee estimate, participants report an estimate which is on average 56 lbs. away from nominee estimate.[23] So, participants update their estimate in the direction of their selected nominee. Among people who did not receive extra reports, the average belief moves 44 lbs. closer to the nominee estimate, and the change in belief is significant (paired t-test, p<0.01).

[21] Recall that participants only see the bids of each potential nominees, not if they actually won the bid or not.

[22] The same hold for participants who lost both bids.

[23] Note: in this section I exclude from the analysis one participant who entered a belief in stage 4 of ”1600” (2930 lbs. away from the truth), presumably a key-in mistake instead of ”4600” which was their latest estimate, and which is only 87 lbs. away from the truth.

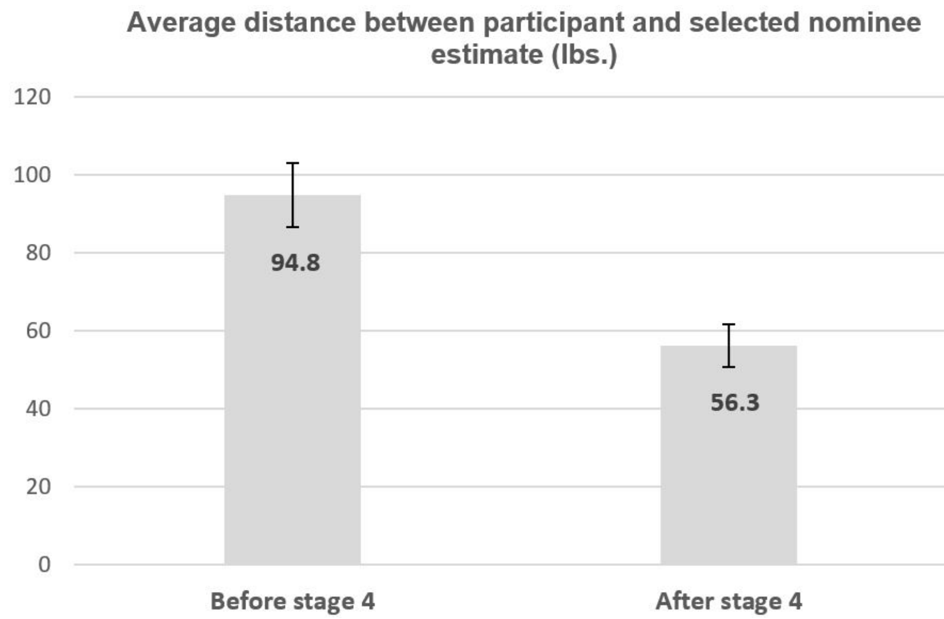


Figure 2.3: Average Absolute Distance Between Participant and Nominee Estimate

To sum up, when given the chance to observe someone's estimate, the highest bidders are chosen. If it were not for the small sample of nominees, we should expect those nominees to have higher degree of correlation neglect, and therefore, be more observed. In addition of being more chosen, they also become more influential since participants revise their estimate and move closer to their selected nominee belief.

2.4.5 Correlation Neglect and Overconfidence

We hypothesized that correlation neglecter would be more overconfident. After seeing two reports, a bayesian (non neglecter) should report a 90% confidence interval of around 1180 lbs., corresponding to a standard deviation of size 360 lbs. If a participant standard deviation is of 310 lbs., their overprecision is evaluated at 50 lbs. Figure 2.4 shows the distribution of overprecision computed for each subject using their latest confidence interval report.[24] Our measure of overconfidence takes into account the number of reports seen, because it is rational that subject's confidence increases

[24]So, if a subject lost stages 2 and 3 bids, and choose not to enter stage 4, then their overconfidence is computed using their stage 1 confidence interval. If a subject only participated in stage 1 and 3, then we use their stage 3 confidence interval report etc.

as they get more signals. For example, a 1160 lbs. confidence interval corresponds to around 10 lbs. of overconfidence after seeing two reports. But after seeing six reports, it is not considered overconfident: 1160 lbs. is in fact the bayesian confidence interval at that point. As shown in figure 2.4, the distribution of overconfidence has three spikes. Overall, around 15% of participants are well calibrated and show no overconfidence. On the other hand, roughly 30% of participants exhibits an overconfidence larger than 50 lbs. Around 20% of people have an intermediate level of overconfidence (20-40 lbs.). In our binary choice menu, the smallest possible 90% confidence interval that a participant can choose has a size of 1000 lbs.[25] This provides an upper bound for our measure overconfidence. We might underestimate if some subjects wished to report smaller confidence intervals.

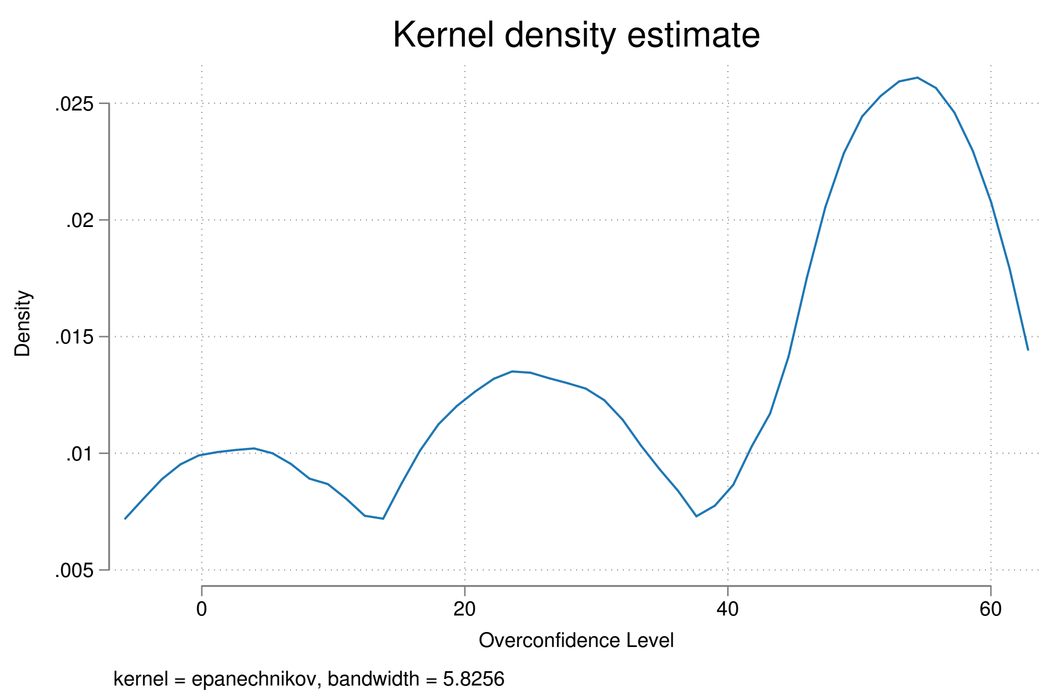


Figure 2.4: Kernel Density Estimate of the Overconfidence Level

We run a fixed-effect regression with overconfidence level as dependent variable (see table 2.11). We restrict our analysis to people who revise their beliefs after stage 1, which accounts for 78% of

[25]This is the case if the subject chooses option B for all rows in figure 3.2.

our sample.[26] As shown in column 1, it appears that correlation neglecters are more overconfident ($p<0.01$). Importantly, both measures of correlation neglect positively affect overconfidence, even if they differ in magnitude (see column 2). Our measure of correlation neglect using confidence interval seems to be a better predictor of overconfidence. However, the size of the effect is relatively small. A .1 increase in correlation neglect reduces the size of one standard deviation by around 4.5 lbs. In other words, the standard deviation of a full neglecter is only about 10% (36 lbs.) smaller than the one of a non neglecter.

Overconfidence Change with Additional Reports

We hypothesized that correlation neglecters would become more overconfident as they receive more reports. Our experiment does not support this hypothesis. Getting one or three extra signals does not significantly increase overconfidence (see column 2 in table 2.11). One possible explanation is the tendency to switch to the center in the binary choice menu used to elicit the confidence intervals. [47] demonstrate that this ”pull-to-center” behavior is common in experiments, and happens regardless of the type of scoring rule used. The midpoint of the table seems indeed to attract a non-negligible share of participants. Among participants who switch between the two options in our table, the midpoint is the switching point for 31% of subjects in the first stage.[27] This could generate inertia and explain why overconfidence does not increase with the number of reports. More simply, participants may have trouble grasping the difference between the 10 lbs. change in range at each row. Asking participants whether they believe there is more than 90% chance that their estimate is somewhere between ± 550 lbs. or ± 560 lbs. from the truth might generate very similar answers.

---

[26]Our measure of correlation neglect is constructed partly using people’s 90% confidence interval at the end of stage 1. So, there would be some endogeneity in using it to predict an overconfidence level which would be measured at the end of stage 1. Therefore, we focus on subjects whose overconfidence levels are estimated after the first stage.

[27]Switching was not required and participants could choose only option A or option B. Including those participants, the midpoint was the choice for 15% of subjects in stage 1.

Table 2.11: Regression with Dependent variable: Overconfidence Level

	Revise belief after stage 1	
	(1)	(2)
CN Level	45.290	
	(6.912)	
Win one extra report	0.881	3.258
	(2.107)	(2.015)
Win three extra reports	−2.515	−2.119
	(1.691)	(1.582)
CN using bids		9.581
		(3.977)
CN using confidence interval		43.315
		(4.890)
Constant	12.799	7.509
	(3.738)	(3.619)
Subject FE	Yes	Yes
Groups	56	56
Observations	272	272

Note: Standard errors in parentheses.
$p < 0.10$, $p < 0.05$, $p < 0.01$

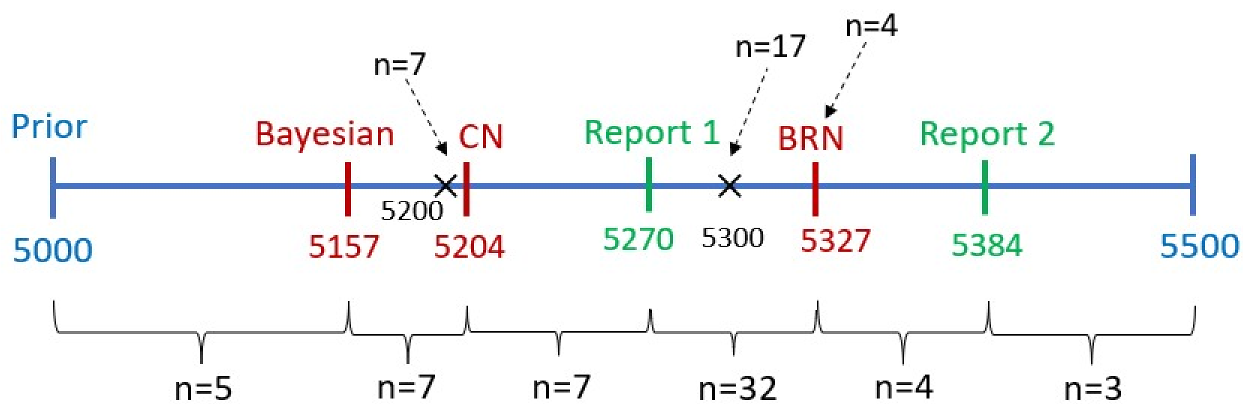
2.4.6 Correlation Neglect and Base-Rate Neglect

One possible concern is that correlation neglecters are in fact base-rate neglecters, i.e., they disregard the prior distribution when forming their beliefs. Our data indicates that this is not the case. First, a full base-rate neglecter estimate would be the average of the reports seen.[28] Indeed, we observe in our data that around a third of subjects seem to consistently give an estimate close to the average of the reports seen. For example, around a third of participants report a stage 1 estimate that is on average ± 10 lbs. away from the average of the first two reports they have received. However, this behavior is not more prevalent among full correlation neglecters than non or low neglecters: 28.6% of low neglecters and 28.1% of full neglecters do so. To illustrate, we display in figure 2.5 the estimates of the participants in stage 1 of the last round (round 7). We choose the last round since

[28] I use the term ”full” base-rate neglecter to refer to an agent who completely disregard their prior. There can be weaker case of base-rate neglect, whereby an agent would underestimate their prior.

by then participants have plenty of time to learn. All participants are shown the following reports: $r_1 = 5270$ and $r_2 = 5384$. Given the prior mean of 5000, the bayesian estimate is 5157 and the full correlation neglecter estimate is 5204. Base-rate neglecters ignore their prior and their estimate is 5327, which is the average of the two reports. Out of 58 subjects, 4 subjects exactly average reports. However, most participants do not ignore their prior, as 81% of subjects submit a belief closer to the prior than the base-rate neglect estimate. A third of participants give an estimate lower than that of the smallest report r_1, again giving an estimate in direction of the prior.

Figure 2.5: Summary of subjects estimates after receiving two reports (round 7)



Notes: Given reports 1 and 2 in the last round, we indicate in red the theoretical estimate for agents who are bayesian, correlation neglecter (CN), or base-rate neglecter (BRN). The number under each bracket indicates the number of subjects who provided an estimate in the given range. For instance, 5 subjects gave an estimate between the prior and the bayesian estimate.

Second, we show in appendix B.3 that an agent who is a base-rate neglecter has a higher posterior variance than a correlation neglecter. In fact, full base-rate neglecters have a posterior variance that is even wider than a Bayesian. This is an important theoretical difference which helps us distinguish between base-rate and correlation neglect. Using each subject mean belief and posterior variance, we can then construct a measure each participant base-rate neglect level based on our theoretical prediction. Our base rate neglect and correlation neglect measures are only weakly correlated (r=.13). As a robustness check, we rerun all our analyses including the base-rate measure instead of the correlation neglect measure. It is never significant, and never qualitatively change our results.

Third, the prior distribution was described at length during the experimental instructions. Each participant was also given a document that displayed the prior (and biases) distribution, and could refer to it at any time. In addition, at the end of each round we told participants the weight of the randomly chosen box and the common bias. All these factors should limit base-rate neglect. There remains the possibility that base-rate neglect and correlation neglect do not affect the same estimates. For example, it could be that base-rate neglect affect the mean, and correlation neglect the variance. This may be worthy of further experimental investigation.

2.5 Conclusion

In our paper, we propose two separate methods to identify individual level of correlation neglect. The intuition behind each method is simple: if we observe an agent signal valuation, or an agent posterior variance, then we should be able to identify individual level of correlation neglect. In the experiment, there is almost no correlation between the individual level of correlation neglect measured by those two methods. This suggests either that they are measuring either a different aspect of the phenomenon, or that one of them fail to properly achieve its goal. In that case, our conjecture is that the measure based on the posterior variance might be closer to the theoretical implication of correlation neglect, whereas the bid measure might be affected by other confounding variables, such as curiosity or stinginess. Overall, the experiment results suggest some heterogeneity in correlation neglect level. About 40% of our sample display substantial level of neglect, greater than .6 out of the maximum possible of .8. The people that we classify as full neglecters are willing to pay on average 51% more for three extra signals than other participants. When additional information is offered at a fixed cost (in stage 4), those with lower level of correlation neglect are less likely to acquire it, as if they understand that the environment is one where signals are highly correlated and therefore additional information has little value. Finally, high correlation neglecters seem also slightly more overprecise than bayesians, even when we only use our correlation neglect measure based on signal valuation. However, the effect is quantitatively small, and we do not observe that overprecision increases as participants are exposed to more information.

Because they overestimate the informativeness of each signal, correlation neglecters are willing to pay more and acquire more information. In the last stage of the experiment, we considered the question of who people would choose to listen to, if they see each other's information valuation. A natural conjecture is that people would choose to listen to people who consume more information, who happen to be correlation neglecters. But are they actually better informed? An interesting question would be to investigate if there exists conditions under which correlation neglecter become better informed about a state, even though they consume too much information compared to the bayesian.

Chapter 3: Guilt-Aversion or Social Image Concerns? Evidence From a Modified Dictator Game

I use a modified dictator game with imperfect information to distinguish between two motives for altruistic behavior: guilt-aversion and social image concerns. A selfish dictator action or nature's move generate the same monetary payoff for the recipient. I exogenously vary the recipient's payoff expectation by introducing uncertainty about the probability with which the dictator is called to play. Increasing the recipient expectation induces guilt-averse dictators to behave more pro-socially. In a second treatment, I reveal to the recipient if their payoff is the result of the dictator or nature's move. Increasing exposure pushes dictator who care about their social image to be more prosocial. I document the extent to which recipient give the benefit of the doubt by allowing them to punish the dictator. I show that it is rational for dictators who care about their social image to exploit the recipient's doubt to their benefit. Dictators start behaving generously only when it is almost certain that the recipient outcome is caused by their action. I argue that this is because there is no social cost of behaving selfishly when the benefit of the doubt is given. This shows that social image concerns are only relevant at the margin. However, some agents still behave pro-socially even when they are given the benefit of the doubt because they are guilt-averse. Finally, I investigate the extent to which people trade self image for social image or guilt aversion, that is their willingness to lie to preserve their social image, or to avoid disappointing others.

Keywords: Guilt aversion; social image; social preferences; experiments

JEL Codes: C91, D83, D91

3.1 Introduction

An extensive literature shows that people have other-regarding preferences, and are motivated by many factors besides their own material payoff. People are willing to forego some of their own payoff for efficiency or to reduce inequality [48–51], reciprocate other's kind or unkind behavior [52], are altruistic and give to charity [53–55], strive to act morally [56, 57], among others. Other motives for prosocial behavior have gained recent interest in the literature, in particular guilt-aversion, social image and self-image. All three can drive prosocial behavior, but they are sometimes mixed.

3.1.1 Guilt-aversion, social image and self-image

First, people feel guilty if they disappoint others and therefore try to meet other's expectations. The existence of *guilt-aversion* has now been documented in many experimental papers [58, 59]. [60] show that in a trust game, trustees who honored the trust were more likely to believe that their partner expected them to do so. Second, *social image* concerns also influences prosocial behavior. People care about their reputation and appearing under a favorable light. What motivates behavior is then appearing generous, honest, trustworthy, fair etc. Many authors argue that social image concerns are triggered by the public nature of the behavior: ”audience effect” [61], ”exposure” [62], ”observability” [63], ”visibility” [64]. Third, people care about doing the right thing and being virtuous, which is called *self-image* concerns. Self-image concerns are often invoked as an explanation for lying aversion, which is caused by an intrinsic cost of lying.[57, 65].

Simple observation of human nature shows that people attach different weights to each of those motivations. Some individuals seem to constantly fear for their reputation, while others do not care for other's opinion as long as they do what they think is right. In addition, those three motivations are not mutually exclusive and may lead to similar behaviors. Two studies will illustrate this point. [66] report that around a third of the subjects in their experiment preferred to take $9 rather than play a $10 Dictator Game. In case participant choose to ”exit” with $9 , the other participant receives nothing and is not even aware that a dictator game was supposed to be played. The resulting allocation for players is (9,0). However, if the dictator is selfish, they should play the dictator game because (10,0) dominates (9,0). If the dictator is inequity averse, they should also play the dictator game rather than exit because (9,1) dominates (9,0). We can justify exiting in the dictator game in two ways. The authors suggest that people ”don’t want to be seen as selfish” and therefore prefer to exit. Another possible interpretation is that people don’t want to enter the game if they think that their allocation choice will disappoint the recipient. Exiting the game is then understood as a way to avoid feeling guilty in case they disappoint the recipient.

In another field study, [67] showed that in a supermarket with multiple entrances, shoppers avoid using the entrance where a charity is posted, especially if it is verbally encouraging donations. If shoppers are inclined to give, avoiding that entrance can occur either because they would feel obliged to give too much compared to what they would give optimally, or because they don’t want to feel guilty for not giving enough. Another interpretation is that shoppers are not inclined to give, but care about their social image and fear being judged as selfish (by the charity employee, or the other shoppers going through the door at the same time) in case they do not give.

The concepts of shame and guilt are often used interchangeably in the literature, which creates confusion. In particular, shame is used to describe the negative utility coming either from a social image or self-image cost. It has initially been argued that it is the public aspect of behavior that leads to shame, whereas guilt is a private emotion (Gehm and Scherer, 1988), but there is little evidence of this ([68] in p.14). There is a stronger support for [69]’s theory which argues that shame is centered on the self (e.g. feeling worthless for ”what I am”), whereas guilt is the result of a bad action (feeling remorse for ”what I did”). Hence, one can feel shame without being observed. In this paper, I use ”guilt-aversion” because the term is widespread in the economic literature. But I will avoid using the word shame, and instead talk separately about social image or self-image costs.

3.1.2 Motivation

The goal of this paper is to distinguish between guilt-aversion and social image concerns as motivating pro-social behavior, controlling for other motives (for instance, inequity aversion). I contribute to the literature on guilt-aversion by exogenously manipulating payoff expectations, and see if agents respond to it. Next, I investigate the degree of visibility which triggers social image costs. Social image concerns are said to matter only when action is fully observable. Yet, for an agent who cares about social image, image costs may depend on the level uncertainty between his action and the outcome in the eyes of the observer. People often seems to take advantage of this ”moral wiggle room” to misbehave [70–72], but little is known about the extent with which they do so. Suppose that Bob is given a sum of money to split between Ann and himself. They both know that Bob will be chosen to make the split with some probability, otherwise the outcome is randomly determined. Whether Ann knows that there is 99% or 1% chance that the amount she receives is the result of Bob’s action will potentially affect the amount Bob chooses to give. If Bob is concerned about his social image, he might be greedy when there is only 1% chance that Ann’s outcome is caused by his choice, but not if there is 99% chance.

If people only care about social image under full (or almost full) visibility, I suggest an explanation for this behavior: the norm of ”giving the benefit of the doubt”. When judging someone’s disposition, we rely not only on the action performed but also on the actor’s intention, as demonstrated by [73]. People may exploit this to their benefit. An agent may undertake an unkind behavior as long as there is a possibility that the action is not fully revealing of the type of person they are. They know that they will be given ”the benefit of the doubt” by the observer. In doing this, they avoid other people judgment and the social image costs associated with it. To my knowledge, only [74] made reference to the idea of the benefit of the doubt.

This paper relates to different parts of the literature. First, it fits in the psychological game theory literature, which which studies how to incorporate belief-dependent preferences into games [75, 76]. Next, it also contributes to the literature exploring the role of emotions such as guilt, shame, anger etc. in games using experiments [59, 77]. Finally, it provides a framework to distinguish between motivations for pro-social behavior. By exogenously manipulating payoff expectations, we can affect guilt-aversion but not social image. On the other hand, we can affect social image but not guilt-aversion by changing whether the dictator's action is observable or not.

3.2 Set-up

3.2.1 Game

There are two possible states of the world $\omega = \{\text{Good,Bad}\}$, where Good is drawn with probability $p \in (0, 1)$. The value of p is common knowledge, but only the dictator (player 1) observes the realization of the state. If the state is Good, the dictator chooses how to split a pie of size 8. He has the choice between two possible actions $a_1 \in \{\text{Share, Take}\}$. He can share equally the pie giving 4 to himself and 4 to the recipient (player 2). Or he can be greedy and "Take", which yields a payoff of 6 for himself and leaves 2 to the recipient. In the Bad state, the pie is only half the size of the Good state pie. In the Bad state, nature moves instead of the dictator. Nature always share the pie equally between the dictator and the recipient, so that each receives 2. The game is designed so that the outcome of the selfish action "Take" in the good state gives player 2 the same monetary payoff as nature's split in the bad state. Hence, upon receiving 2, player 2 can not know with certainty whether it is the result of dictator selfish action in the good sate, or nature in the bad state. And player 1 knows that. I implicitly assume that player 2 only sees her payoff and not the dictator's payoff, and therefore can not learn ω. Let $\pi_2^m(z)$ denote P2's monetary outcome at terminal history $z \in Z$, where Z is the set of terminal histories of the game. If player 2 receives 2, she can choose to punish the dictator: $a_2 \in \{\text{Punish, Don't}\}$. Punishing reduces the dictator's payoff by a half. Figure 3.1 depicts this modified dictator game.

I am now introducing two behavioral features into the player's preference. First, the dictator is guilt averse and does not want to disappoint the recipient. Second, the recipient gives the benefit

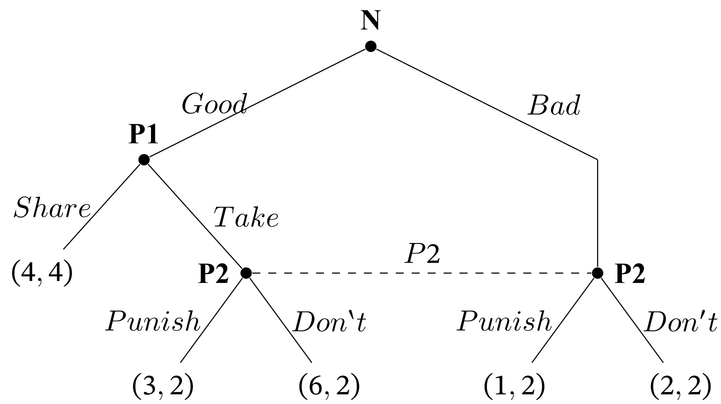


Figure 3.1: A Modified Dictator Game

of the doubt to the dictator and only wish to punish a selfish dictator. That is, if player 2 punishes a dictator in the bad state (where nature made the split) she incurs an emotional cost of $2\pi_2^e(z)$, whereas punishing in the good state (where the recipient low monetary payoff is caused by the dictator selfish move) is rewarded by $\pi_2^e(z)$. This captures the norm of giving the benefit of the doubt, and the fact that the cost for a wrong punishment is proportionally higher than the reward for correctly punishing. This is similar to the idea that it is worse to send an innocent man in jail than letting a guilty person go.

Denote λ the probability that player 2 punish. Let us also denote the dictator strategy in state ω as $\sigma_\omega = P[Share \mid \omega]$. Player 2 has a (first order) belief $\hat{\sigma_\omega}$ about about the probability that player 1 will Share in state G: $\hat{\sigma_\omega} = \mathbb{E}_2[\sigma_\omega]$. Denote $\pi_2^e(z)$ the cost (or reward) of punishing the dictator. Player 2's preferences are given by:

$$u_2(z) = \pi_2^m(z) + \pi_2^e(z)$$

The dictator feels guilty if he disappoints the recipient. I follow [58], and define the recipient's disappointment $D_2(z;\ \hat{\sigma_\omega}, p)$ as the difference between what the recipient expects to receive *before* the game starts, and their final material payoff:

$$D_2(z;\ \hat{\sigma_\omega}, p) = Max\{0, \mathbb{E}_2[\pi_2^m(z)] - \pi_2^m(z)\}$$

We assume that player 2 disappointment can be at most zero. This rules out cases where player 2 would not be disappointed but, on the contrary, feels positively surprised if they receive more than they expected initially. Player 1 preference is given by:

$$u_1(z;\ \theta, \hat{\sigma_\omega}, p, \lambda) = \pi_1(z) - \theta D_2(z;\ \hat{\sigma_\omega}, p)$$

with the dictator's guilt sensitivity $\theta \in \mathbb{R}^+$. When $\theta = 0$, the dictator is not guilt-averse and only cares about their own payoff. The revised game, including both the dictator guilt-aversion and the norm of giving the benefit of the doubt is shown in figure 3.2.

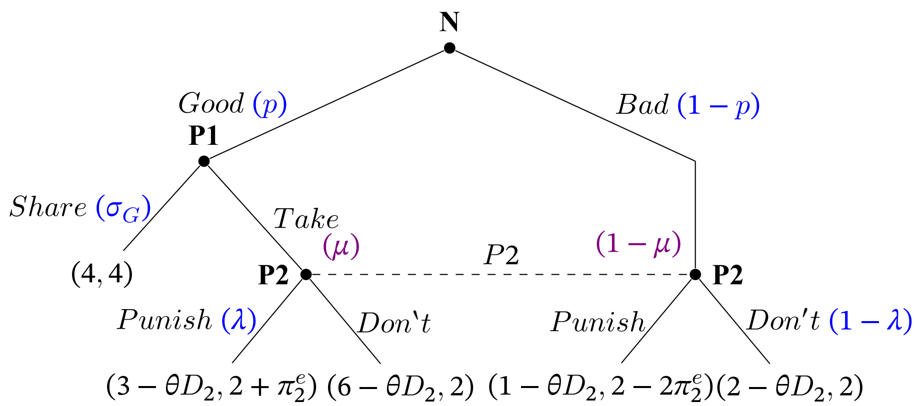


Figure 3.2: Revised Dictator Game with Guilt Aversion and the Benefit of the Doubt

3.2.2 Recipient disappointment

Before the game starts, the recipient expects to receive from the dictator or Nature's split:

$$\mathbb{E}_2[\pi_2^m(z; \hat{\sigma_\omega}, p)] = 2p\sigma_{\hat{G}} + 2$$

Ex ante, the more likely it is that the state is Good, and the stronger the recipient belief that the dictator will share, the higher is the recipient expected monetary payoff. Therefore, the recipient disappointment $D_2(z; \hat{\sigma_\omega}, p)$ is equal to:

$$0 \qquad \text{if} \quad z = \{G, Share\} \tag{3.1}$$

$$Max\{0, \mathbb{E}_2[\pi_2^m] - 2\} \text{ after } h = (\{G, Take\}, \{B\}) \tag{3.2}$$

Notice that the recipient's disappointment is zero if the dictator decides to share in the good state. Recall that we assume that the disappointment can not be positive.[1]

3.2.3 Optimal strategies

The recipient disappointment at each terminal history enters the utility of player 1. Player 1 expected utility also depends on their guilt sensitivity θ and the probability the player 2 punish λ. Because the recipient is never disappointed in case the dictator share in the good state, the dictator payoff in that case is simply their monetary payoff 4.

$$\mathbb{E}[u_1 \mid G, Share] = 4$$

$$\mathbb{E}[u_1 \mid G, Take] = 6 - 3\lambda - \theta\, D_2(G, T)$$

$$\mathbb{E}[u_1 \mid B] = 2 - \lambda - \theta\, D_2(B)$$

Player 1 prefers to Take if $\mathbb{E}[u_1 \mid G, Take] > \mathbb{E}[u_1 \mid G, Share]$. Hence, player 1 optimal strategy is:

$$\begin{aligned}
\sigma_G &= 0 && \text{if} \quad \lambda < \frac{2 - \theta\ D_2(G,T)}{3} \\
\sigma_G &\in (0,1) && \text{if} \quad \lambda = \frac{2 - \theta\ D_2(G,T)}{3} \\
\sigma_G &= 1 && \text{if} \quad \lambda > \frac{2 - \theta\ D_2(G,T)}{3}
\end{aligned} \tag{3.3}$$

The probability that player 1 plays Share is increasing in λ, θ, p and $\sigma_{\hat{G}}$ (since $\mathbb{E}_2[\pi_2^m]$ is increasing in p and $\sigma_{\hat{G}}$). Intuitively, the higher the probability of punishment and the more guilt averse,

[1]This could be the case if the recipient is positively surprised by receiving the highest possible payoff of 4, which exceeds their expectation.

the more player 1 share. Next, player 1 is also more likely to share when player 2's expectations are high, which occurs when the Good state is more likely, and the more player 2 expects player 1 to share. High expectations from player 2 lead to higher disappointment if player 1 decide to take instead.

P2 believes to be at history f$G, Take$g with probability $\mu = \frac{p(1-\sigma_{\hat{G}})}{1-p\sigma_{\hat{G}}}$. The expected utility of each action is:

$$\mathbb{E}[u_2 \mid Punish] = 3\mu\pi_2^e - 2\pi_2^e + 2 \text{ and } \mathbb{E}[u_2 \mid Don`t] = 2$$

Hence, player 2 chooses to punish if $\mu > \frac{2}{3}$. The recipient optimal strategy is then:

$$\begin{aligned} \lambda &= 0 && \text{if} \quad p < \frac{2}{3-\sigma_{\hat{G}}} \\ \lambda &\in (0,1) && \text{if} \quad p = \frac{2}{3-\sigma_{\hat{G}}} \\ \lambda &= 1 && \text{if} \quad p > \frac{2}{3-\sigma_{\hat{G}}} \end{aligned} \tag{3.4}$$

The more player 2 believes that they are in the good state and the dictator played Take, the more likely they are to punish. So, the probability that player 2 punish is decreasing in $\sigma_{\hat{G}}$ and increasing in p.

3.2.4 Equilibria

We use Perfect Bayesian Nash Equilibrium (PBE) as our solution concept. Because the game has no proper subgame, we cannot use Subgame Perfect Nash Equilibrium as solution concept.

Equilibrium with standard preferences

Proposition 1 With standard preferences, that is, without guilt-aversion ($\theta = 0$), there are three equilibria:

1. $\lambda = 0$, $\sigma_G = 0$, when $p < \frac{2}{3-\sigma_{\hat{G}}}$
2. $\lambda = 1$, $\sigma_G = 1$, when $p > \frac{2}{3-\sigma_{\hat{G}}}$.
3. $\lambda = \frac{2}{3}$, $\sigma_G \in (0, 1)$ and $p = \frac{2}{3-\sigma_{\hat{G}}}$

Finally, the consistency condition requires that $\mu = \frac{p(1-\sigma_{\hat{G}})}{1-p\sigma_{\hat{G}}}$.

The intuition for the equilibrium is that the lower the p and the higher $\sigma_{\hat{G}}$, the less likely the recipient believe that they are at f$G, Take$g. In case 1, even though the recipient expects the dictator to be selfish ($\sigma_{\hat{G}} = 0$), it is unlikely that the dictator is called to move. Therefore, the recipient gives the benefit of the doubt and does not punish the dictator because receiving a payoff of 2 is likely caused by nature's move. The dictator exploits this to their benefit to Take (1). On the other hand, the higher the p and the lower $\sigma_{\hat{G}}$, the more likely the recipient believe they are at f$G, Take$g. In the equilibrium 2, as the dictator is likely to be called to move, the recipient chooses to Punish, which prompts the dictator to Share. In case 3, the recipient mixes and punishes two-third of the times, which makes the dictator indifferent between playing Take and Share. This occurs when $p = \frac{2}{3-\sigma_{\hat{G}}}$.

Figure 3.3 displays the equilibrium for each combination of of p and $\sigma_{\hat{G}}$. The blue area is the combination of p and $\sigma_{\hat{G}}$ such that the recipient punish and the dictator plays Share. In the white area, the recipient does not punish and the dictator plays Take. The blue line represents value of p such that the recipient mixes between punishing and not punishing, which makes the dictator indifferent between playing Share and Take.

Equilibrium with non-standard preferences

With non-standard preferences, there exists an equilibrium in which dictators who are guilt-averse enough play Share, even though the recipient does not punish. This occurs when the recipient's expectations are high.

Proposition 2 With non-standard preferences ($\theta > 0$), the equilibrium is:

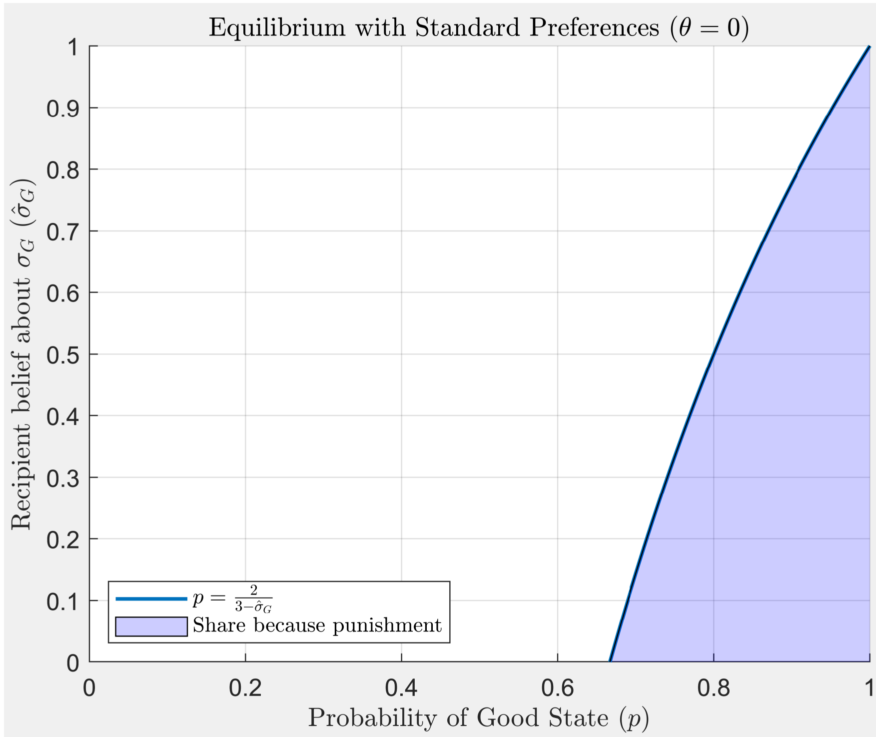


Figure 3.3: Equilibrium With Standard Preferences

$$\lambda = \frac{2}{3}\,\frac{3 - \sigma_{\hat{G}}(1 + 2\theta)}{3 - \sigma_{\hat{G}}}$$

$$p = \frac{2}{3 - \sigma_{\hat{G}}} \in (\frac{2}{3}, 1) \text{ since } \sigma_{\hat{G}} \in (0, 1)$$

The equilibrium exists when:

$$\theta \in (0, \frac{3 - \sigma_{\hat{G}}}{2\sigma_{\hat{G}}})$$

The restriction on the value of θ ensures that $\lambda \in (0, 1)$. In the presence of a guilt-averse dictator, there exists a combination of $(p, \sigma_{\hat{G}})$ (specifically, $p > \frac{1}{\theta\sigma_{\hat{G}}}$) such that a dictator who is guilt averse enough plays Share *even if the recipient does not punish*. If $p \leq \frac{1}{\theta\sigma_{\hat{G}}}$, the dictator is not guilt-averse enough and plays Take. This equilibrium is shown by the area in red in figure 3.4. The bold red curve is $p = \frac{1}{\theta\sigma_{\hat{G}}}$. There still exists an equilibrium where recipient punish (when $p > \frac{2}{3-\sigma_{\hat{G}}}$), which pushes the dictator to Share regardless of the dictator guilt aversion level. This is displayed by the area in blue.

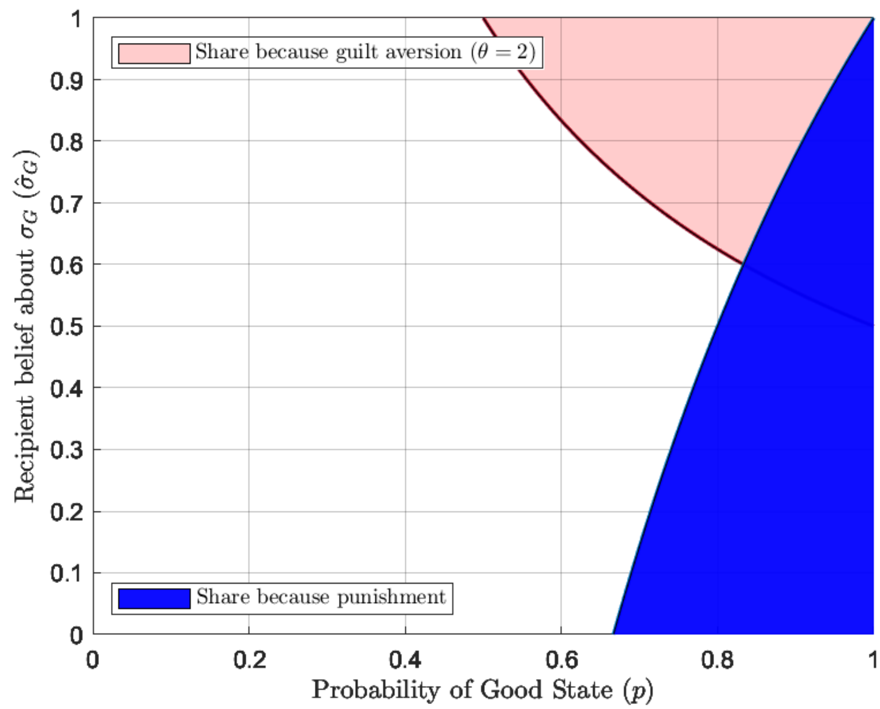


Figure 3.4: Equilibrium With a Guilt-Averse Dictator

3.3 Experimental Design Proposal

There are three treatments, which vary the visibility of the dictator's action, and whether the recipient can punish or not the dictator. I use a within-subject design to compare subjects behavior across treatment. The main idea is that increasing p and $\sigma_{\hat{G}}$ should affect guilt aversion but not social image. On the other hand, increasing visibility should affect social image but not guilt aversion.

In the baseline *no visibility-no punishment* treatment, the recipient never learns the dictator decision, and can not punish him. Thus the dictator's decision to Share or Take will only be affected by p and $\sigma_{\hat{G}}$. In other words, only guilt-aversion should play a role in explaining why participants

		Visibility	
		No	Yes
Punishment	No	Baseline	Treatment 2
	Yes	Treatment 3	

Table 3.1: Experimental treatments

may change behavior as p is modified. Remember that dictator with guilt-averse preferences should Share more as p increases (fixing $\sigma_{\hat{G}}$) because the recipient expects more. After pairing participants, I ask the recipient to guess the percentage of participants in the experiment that they expect to share if the state is Good (Note that I don't condition the recipient belief on p, in which case the question would be "if there is p% chance that the state is good, what is the percentage of participants who plan on sharing"). This is player 2 first order belief about the probability that player 1 shares, and is sent to the dictator. This belief is incentivized and the closer the participant guess is from the experiment average, the higher the payoff. Another possibility is to make use of results from [78] study who elicited the appropriateness of various splits in dictator game.[2] However, this approach doesn't exactly elicit the recipient first order belief about the dictator they are paired with, unless player 2 believes she is paired with someone who will behave like the average person in the population.[3] After sending the recipient's first order belief to the dictator, I elicit the dictator strategy conditional on p using the *strategy method*. The dictator is presented with a series of choices: "If there is p% chance that the state is Good, would you choose Share or Take?" I remind the dictator that the recipient will know the probability with which the state is drawn but not the state itself.

Unlike the baseline, in the *visibility-no punishment* treatment, the recipient is always told the which state is drawn. Thus, this information is made salient when eliciting player 1 strategy. Even though the dictator's choice will be implemented with p% chance, because the recipient learns which state is drawn, they know whether the dictator or nature is responsible for her payoff. By

[2] For example in a standard dictator game with \$10 to split, 83% of their participants considered a (\$5,\$5) split "very socially appropriate", whereas 61% of their participants considered a (\$9,\$1) split "very socially inappropriate".

[3] On the other hand it is also possible that the guilt-averse dictator care about the what the average behavior in the population is i.e. the norm ,and not $\mathbb{E}_2[\pi_2^m]$ (see [79] for a discussion).

comparing within-subject behavior between the baseline and the *visibility-no punishment* treatment, I distinguish between guilt-aversion and social image behaviors. A pure guilt-averse agent should display the same behavior as in treatment one. On the other hand, an agent concerned by social image will share in the *visibility-no punishment* treatment but not in the baseline. The *visibility-no punishment* treatment is designed to measure social image concerns. An alternative implementation to measure the extent with which such concerns affect dictator's behavior would be to offer one of the following option to dictators who choose Share (and hence may be motivated either by social image concerns or guilt aversion):

- Option 1: offer to switch to Take but pay to tell the recipient that the state was "Bad" instead of "Good". In other words, the recipient now believes that low payoff received is due to bad luck (nature) and not the dictator.
- Option 2: offer to switch to Take, and the experimenter informs the dictator that he will compensate the recipient for the loss due to the switch. However, the recipient is still told that the state was Good and that the dictator chose Take, but that her payoff of 10 is the result of the experimenter adding 8 to the 2 that the dictator gave her.

A dictator who initially decide to share because he is motivated by social image concerns would choose option 1, which allow him to avoid social image costs. On the other hand, if he is motivated by pure guilt aversion and not image concerns, he would choose option 2. Option 2 leaves the recipient with 10, and they can't be disappointed with the maximum possible payoff. Option 2 is akin to trading guilt aversion for social image cost. Note that option 1 in this methodology involves some form of deception for the recipient. Option 2 does not involve deception, but can only be used once since in the end the monetary payoff for both player is (18,10). If the dictator is offered the switch every time (and the recipient knows that a switch is possible), then this changes the game.

Finally, the *no visibility-punishment* treatment replicates exactly the game described in the model. The goal is to determine if the recipient gives the dictator the benefit of the doubt, i.e., choose not to punish upon receiving the lowest payoff, and if the dictator exploits this to choose

Take. By comparing behavior with the baseline treatment, I can see which dictators refrain from Taking (which they did in the baseline) out of fear from being punished.

Three main confounding factors can push a dictator to behave pro-socially: inequity aversion, altruism and efficiency. Changing p or visibility should have no effect if our participants have outcome based preferences. Inequity aversion and altruism are outcome based, so a dictator should always play the same action whatever the value of p. So, those types of models can not explain a behavior where the dictator switch from share to take (or vice versa) as p increases. But guilt aversion can explain such a switchversion can. A dictator may also cares about efficiency, i.e., the sum of all players payoffs at the end of the game. To rule this out I fixed the pie size conditional on the state (assuming no punishment). Whether the dictator share or take, the total payoff for both players is the same (8) in the good state if P2 doesn't punish.

3.4 Conclusion

We demonstrate theoretically that in the presence of guilt-aversion, there exists an equilibrium in which a guilt-averse dictator plays Share, even though they expect no punishment from the recipient. This occurs because a guilt-averse dictator do not want to disappoint the recipient, in case they have high expectations.

The game has some similarities with [61] in the sense that I vary the probability with which nature is called to move. However there are some notable differences. In particular I restrict the dictator decision to two possible allocations: either a 50-50 split or giving the recipient the same amount than nature would give when the state is bad. When there is uncertainty (in our case p not equal to 1), around 80% of participants in [61] choose either one of these allocation, despite being able to choose any split between 0 and 20. Therefore, it seems likely that restricting the dictator's decision to those two allocations captures most of the behaviors we would observe if we give dictators the choice to allocate any amount they want to the recipient.

The approach used here is one where presumably participant have a social image that they want to protect, and the recipient has expectations that the dictator will try to meet. Another interesting approach would be for the dictator to start with a low social image, or that the recipient starts with low expectations about the dictator, which the dictator knows. In that case, a guilt-averse player should share less, while a dictator concerned by their social image would try to share more to improve his image.

References

1. Atkinson, A. B. *Inequality: What can be done?* (Harvard University Press, 2015) (cited on p. 2).
2. Cappelen, A. W., Hole, A. D., Sørensen, E. & Tungodden, B. The pluralism of fairness ideals: An experimental approach. *American Economic Review* **97,** 818–827 (2007) (cited on pp. 2, 6).
3. Cappelen, A. W., Konow, J., Sørensen, E. & Tungodden, B. Just luck: An experimental study of risk-taking and fairness. *American Economic Review* **103,** 1398–1413 (2013) (cited on pp. 2, 6).
4. Mollerstrom, J., Reme, B.-A. & Sørensen, E. Luck, choice and responsibility—An experimental study of fairness views. *Journal of Public Economics* **131.** Publisher: Elsevier, 33–40 (2015) (cited on pp. 2, 6).
5. Cappelen, A. W., Mollerstrom, J., Reme, B.-A. & Tungodden, B. A meritocratic origin of egalitarian behaviour. *The Economic Journal* **132.** Publisher: Oxford University Press, 2101–2117 (2022) (cited on pp. 2, 6, 7, 10, 21).
6. Andre, P. *Shallow Meritocracy* tech. rep. (University of Bonn and University of Mannheim, Germany, 2022) (cited on p. 2).
7. Gill, D. & Prowse, V. A structural analysis of disappointment aversion in a real effort competition. *American Economic Review* **102.** Publisher: American Economic Association, 469–503 (2012) (cited on pp. 4, 15).
8. Gill, D. & Prowse, V. Measuring costly effort using the slider task. *Journal of Behavioral and Experimental Finance* **21.** Publisher: Elsevier, 1–9 (2019) (cited on pp. 4, 43).
9. Almås, I., Cappelen, A. W. & Tungodden, B. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy* **128.** Publisher: The University of Chicago Press Chicago, IL, 1753–1788 (2020) (cited on pp. 6, 21, 37).
10. Cappelen, A. W., Moene, K. O., Skjelbred, S.-E. & Tungodden, B. The merit primacy effect. *The Economic Journal* **133.** Publisher: Oxford University Press, 951–970 (2023) (cited on p. 6).
11. Alesina, A. & Giuliano, P. in *Handbook of social economics* 93–131 (Elsevier, 2011) (cited on p. 7).

12. Preuss, M., Reyes, G., Somerville, J. & Wu, J. Inequality of Opportunity and Income Redistribution. *arXiv preprint arXiv:2209.00534* (2022) (cited on p. 8).

13. Andreoni, J., Aydin, D., Barton, B., Bernheim, B. D. & Naecker, J. When fair isn't fair: Understanding choice reversals involving social preferences. *Journal of Political Economy* **128.** Publisher: The University of Chicago Press Chicago, IL, 1673–1711 (2020) (cited on p. 8).

14. Dewatripont, M., Jewitt, I. & Tirole, J. The economics of career concerns, part I: Comparing information structures. *The Review of Economic Studies* **66.** Publisher: Wiley-Blackwell, 183–198 (1999) (cited on p. 8).

15. Corak, M. Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives* **27.** Publisher: American Economic Association, 79–102 (2013) (cited on p. 9).

16. Fehr, D. Is increasing inequality harmful? Experimental evidence. *Games and economic behavior* **107.** Publisher: Elsevier, 123–134 (2018) (cited on p. 9).

17. DellaVigna, S. & Pope, D. What motivates effort? Evidence and expert forecasts. *The Review of Economic Studies* **85.** Publisher: Oxford University Press, 1029–1069 (2018) (cited on pp. 11, 99).

18. Brañas-Garza, P., Jorrat, D., Espín, A. M. & Sánchez, A. Paid and hypothetical time preferences are the same: Lab, field and online evidence. *Experimental Economics* **26.** Publisher: Springer, 412–434 (2023) (cited on p. 22).

19. Hvidberg, K. B., Kreiner, C. & Stantcheva, S. *Social Positions and Fairness Views on Inequality* tech. rep. (National Bureau of Economic Research, 2020) (cited on p. 22).

20. Chen, D. L., Schonger, M. & Wickens, C. oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9.** Publisher: Elsevier, 88–97 (2016) (cited on pp. 22, 56).

21. Alesina, A., Stantcheva, S. & Teso, E. Intergenerational mobility and preferences for redistribution. *American Economic Review* **108,** 521–54 (2018) (cited on p. 44).

22. Alesina, A. & Angeletos, G.-M. Fairness and redistribution. *American economic review* **95,** 960–980 (2005) (cited on p. 44).

23. Benabou, R. & Ok, E. A. *Mobility as progressivity: ranking income processes according to equality of opportunity* 2001 (cited on p. 44).

24. Benabou, R. & Ok, E. A. Social mobility and the demand for redistribution: the POUM hypothesis. *The Quarterly journal of economics* **116.** Publisher: MIT Press, 447–487 (2001) (cited on p. 44).

25. Ortoleva, P. & Snowberg, E. Overconfidence in political behavior. *American Economic Review* **105.** Number: 2, 504–35 (2015) (cited on pp. 45–48, 119).

26. Moore, D. A. & Healy, P. J. The trouble with overconfidence. *Psychological review* **115.** Number: 2, 502 (2008) (cited on p. 46).

27. Camerer, C. & Lovallo, D. Overconfidence and excess entry: An experimental approach. *The American Economic Review* **89.** Number: 1, 306–318 (1999) (cited on p. 46).

28. Koellinger, P., Minniti, M. & Schade, C. "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of economic psychology* **28.** Number: 4, 502–527 (2007) (cited on p. 46).

29. Svenson, O. Are we all less risky and more skillful than our fellow drivers? *Acta psychologica* **47.** Number: 2 Publisher: Elsevier, 143–148 (1981) (cited on p. 46).

30. Barber, B. M. & Odean, T. Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly journal of Economics,* 261–292 (2001) (cited on p. 46).

31. Ben-David, I., Graham, J. R. & Harvey, C. R. Managerial miscalibration. *The Quarterly journal of economics* **128.** Number: 4 Publisher: MIT Press, 1547–1584 (2013) (cited on p. 46).

32. Malmendier, U. & Tate, G. CEO overconfidence and corporate investment. *The journal of finance* **60.** Number: 6 Publisher: Wiley Online Library, 2661–2700 (2005) (cited on p. 46).

33. Hirshleifer, D., Low, A. & Teoh, S. H. Are overconfident CEOs better innovators? *The journal of finance* **67.** Number: 4 Publisher: Wiley Online Library, 1457–1498 (2012) (cited on p. 46).

34. Bénabou, R. & Tirole, J. in *Psychology, Rationality and Economic Behaviour* 19–57 (Springer, 2005) (cited on p. 46).

35. DeMarzo, P. M., Vayanos, D. & Zwiebel, J. Persuasion bias, social influence, and unidimensional opinions. *The Quarterly journal of economics* **118.** Number: 3 Publisher: MIT Press, 909–968 (2003) (cited on pp. 46, 47).

36. Eyster, E. & Rabin, M. Extensive imitation is irrational and harmful. *The Quarterly Journal of Economics* **129.** Number: 4 Publisher: Oxford University Press, 1861–1898 (2014) (cited on pp. 46, 47).

37. Levy, G. & Razin, R. Correlation neglect, voting behavior, and information aggregation. *American Economic Review* **105.** Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 1634–1645 (2015) (cited on p. 47).

38. Moser, J. & Wallmeier, N. Correlation neglect in voting decisions: An experiment. *Economics Letters* **198.** Publisher: Elsevier, 109656 (2021) (cited on p. 47).

39. Eyster, E. & Weizsacker, G. Correlation neglect in portfolio choice: Lab evidence. *Available at SSRN 2914526* (2016) (cited on p. 47).

40. Rees-Jones, A., Shorrer, R. I. & Tergiman, C. Correlation Neglect in Student-to-School Matching. *NBER Working Paper* (2020) (cited on p. 47).
41. Bolte, L. & Fan, T. Q. Motivated mislearning: The case of correlation neglect. *Journal of Economic Behavior & Organization* **217.** Publisher: Elsevier, 647–663 (2024) (cited on p. 47).
42. Enke, B. & Zimmermann, F. Correlation neglect in belief formation. *The Review of Economic Studies* **86.** Number: 1 Publisher: Oxford University Press, 313–332 (2019) (cited on p. 47).
43. Eyster, E., Rabin, M. & Weizsacker, G. An experiment on social mislearning. *Available at SSRN 2704746* (2015) (cited on p. 47).
44. Benjamin, D. J. Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1* **2.** Publisher: Elsevier, 69–186 (2019) (cited on p. 48).
45. Graeber, T. Inattentive inference. *Journal of the European Economic Association* **21.** Publisher: Oxford University Press, 560–592 (2023) (cited on p. 48).
46. Enke, B. What you see is all there is. *The Quarterly Journal of Economics* **135.** Publisher: Oxford University Press, 1363–1398 (2020) (cited on p. 48).
47. Danz, D., Vesterlund, L. & Wilson, A. J. Belief Elicitation and Behavioral Incentive Compatibility. en. *American Economic Review* **112,** 2851–2883. doi:10.1257/aer.20201248 (2022) (cited on p. 71).
48. Charness, G. & Rabin, M. Understanding social preferences with simple tests. *The Quarterly Journal of Economics* **117.** Publisher: Oxford University Press, 817–869 (2002) (cited on p. 77).
49. Charness, G. & Rabin, M. Social preferences: Some simple tests and a new model (2000) (cited on p. 77).
50. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* **114.** Publisher: MIT Press, 817–868 (1999) (cited on p. 77).
51. Bolton, G. E. & Ockenfels, A. ERC: A theory of equity, reciprocity, and competition. *American economic review* **90,** 166–193 (2000) (cited on p. 77).
52. Falk, A. & Fischbacher, U. A theory of reciprocity. *Games and economic behavior* **54.** Publisher: Elsevier, 293–315 (2006) (cited on p. 77).
53. Andreoni, J. Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal* **100.** Publisher: JSTOR, 464–477 (1990) (cited on p. 77).
54. Fisman, R., Kariv, S. & Markovits, D. Individual preferences for giving. *American Economic Review* **97,** 1858–1876 (2007) (cited on p. 77).

55. Levine, D. K. Modeling altruism and spitefulness in experiments. *Review of economic dynamics* **1.** Publisher: Elsevier, 593–622 (1998) (cited on p. 77).

56. Falk, A. & Szech, N. Morals and markets. *science* **340.** Publisher: American Association for the Advancement of Science, 707–711 (2013) (cited on p. 77).

57. Bénabou, R. & Tirole, J. Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics* **126.** Publisher: MIT Press, 805–855 (2011) (cited on p. 77).

58. Battigalli, P. & Dufwenberg, M. Guilt in games. *American Economic Review* **97,** 170–176 (2007) (cited on pp. 77, 81).

59. Attanasi, G., Battigalli, P., Manzoni, E. & Nagel, R. Belief-dependent preferences and reputation: Experimental analysis of a repeated trust game. *Journal of Economic Behavior & Organization* **167.** Publisher: Elsevier, 341–360 (2019) (cited on pp. 77, 80).

60. Charness, G. & Dufwenberg, M. Promises and partnership. *Econometrica* **74.** Publisher: Wiley Online Library, 1579–1601 (2006) (cited on p. 77).

61. Andreoni, J. & Bernheim, B. D. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica* **77.** Publisher: Wiley Online Library, 1607–1636 (2009) (cited on pp. 77, 90).

62. Tadelis, S. The Power of Shame and the Rationality of Trust. *Available at SSRN 1006169* (2007) (cited on p. 77).

63. Dillenberger, D. & Sadowski, P. Ashamed to be selfish. *Theoretical Economics* **7.** Publisher: Wiley Online Library, 99–124 (2012) (cited on p. 77).

64. Butera, L., Metcalfe, R., Morrison, W. & Taubinsky, D. Measuring the welfare effects of shame and pride. *American Economic Review* **112,** 122–68 (2022) (cited on p. 77).

65. Akerlof, G. A. & Kranton, R. E. Economics and identity. *The Quarterly Journal of Economics* **115.** Publisher: MIT Press, 715–753 (2000) (cited on p. 77).

66. Dana, J., Cain, D. M. & Dawes, R. M. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and human decision Processes* **100.** Publisher: Elsevier, 193–201 (2006) (cited on p. 77).

67. Andreoni, J., Rao, J. M. & Trachtman, H. Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy* **125.** Publisher: University of Chicago Press Chicago, IL, 625–653 (2017) (cited on p. 78).

68. Tangney, J. P. & Dearing, R. L. *Shame and guilt* (Guilford Press, 2003) (cited on p. 78).

69. Lewis, B. Shame and guilt in neurosis. *Psychoanalytic review* **58.** Publisher: National Psychological Association for Psychoanalysis, 419–438 (1971) (cited on p. 78).

70. Guth, W., Huck, S. & Ockenfels, P. Two-Level Ultimatum Bargaining with Incomplete Information: An Experimental Study. *Economic Journal* **106.** Publisher: Royal Economic Society, 593–604 (1996) (cited on p. 79).

71. Parra, D., Muñoz-Herrera, M. & Palacio, L. A. The limits of transparency in reducing corruption. *Journal of Behavioral and Experimental Economics.* Publisher: Elsevier, 101762 (2021) (cited on p. 79).

72. Ockenfels, A. & Werner, P. 'Hiding behind a small cake' in a newspaper dictator game. *Journal of Economic Behavior & Organization* **82.** Publisher: Elsevier, 82–85 (2012) (cited on p. 79).

73. Falk, A., Fehr, E. & Fischbacher, U. Testing theories of fairness—Intentions matter. *Games and Economic Behavior* **62.** Publisher: Elsevier, 287–303 (2008) (cited on p. 79).

74. Stüber, R. The benefit of the doubt: Willful ignorance and altruistic punishment. *Experimental Economics.* Publisher: Springer, 1–25 (2019) (cited on p. 79).

75. Battigalli, P. & Dufwenberg, M. Dynamic psychological games. *Journal of Economic Theory* **144.** Publisher: Elsevier, 1–35 (2009) (cited on p. 79).

76. Battigalli, P., Corrao, R. & Dufwenberg, M. Incorporating belief-dependent motivation in games. *Journal of Economic Behavior & Organization* **167.** Publisher: Elsevier, 185–218 (2019) (cited on p. 79).

77. Aina, C., Battigalli, P. & Gamba, A. Frustration and anger in the Ultimatum Game: An experiment. *Games and Economic Behavior* **122.** Publisher: Elsevier, 150–167 (2020) (cited on p. 80).

78. Krupka, E. L. & Weber, R. A. Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association* **11.** Publisher: Oxford University Press, 495–524 (2013) (cited on p. 88).

79. Cartwright, E. A survey of belief-based guilt aversion in trust and dictator games. *Journal of Economic Behavior & Organization* **167.** Publisher: Elsevier, 430–444 (2019) (cited on p. 88).

Appendix A: Circumstances, Effort Choice and Redistribution

A.1 Omitted Proofs

A.1.1 Proof of Proposition 1: Optimal Effort

Proposition 1 *The optimal effort e_i is equal to*

$$e_i^* = \left(\frac{x\,(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}\right)^{\frac{1}{3}} \quad \text{if} \quad \kappa_i > 1-\overline{\lambda}\left(\frac{x\,s_i}{b}\right)^{\frac{1}{2}} \tag{A.1}$$

$$e_i^* = \quad 0 \quad \text{otherwise}$$

Proof:

The agents succeeds when $P_i \geq 1$, that is, if $\lambda_i \geq \frac{1-\kappa_i}{e_i}$. Recall that since $s \sim \mathcal{U}\left[0, \overline{\lambda}\right]$, the probability p that agent i succeeds is given by: [1]

$$p = p(\lambda_i \geq \frac{1-\kappa_i}{e_i}) = \frac{e_i\overline{\lambda} + \kappa_i - 1}{e_i\overline{\lambda}} \tag{A.2}$$

The probability q that the other agent j in the group reaches the target is:[2]

$$q = p\left(\lambda_j \geq \frac{1-\mathbb{E}[k_j]}{e_j}\right) = \frac{e_j\overline{\lambda} + \kappa - 1}{e_j\,\overline{\lambda}}$$

Agent i learns his circumstances κ_i and knows the luck distribution. Hence, he solves:

$$\max_{e_i \geq 0} \; x\,pq + s_i\,x\,p(1-q) + (1-s_i)\,x\,(1-p)q - c(e_i) \tag{A.3}$$

[1] I also derive in Appendix A4 the optimal effort with the more general case where $s \sim \mathcal{U}\left[\underline{\lambda}, \overline{\lambda}\right]$

[2] I assume that i knows κ_i but not κ_j. When agent i knows both κ_i and κ_j, then $q = \frac{e_j\,\overline{\lambda}+k_j-1}{e_j\overline{\lambda}}$. Using this does not change the result of the maximization problem.

Maximizing with respect to e_i gives:

$$c'(e_i) = \frac{x\,(1-\kappa_i)\,s_i}{e_i^2\,\overline{\lambda}} \tag{A.4}$$

In the rest of the paper, I assume that $c(e_i) = b\frac{e_i^2}{2}$, then:[3]

$$e_i = \left(\frac{x\,(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}\right)^{\frac{1}{3}} \tag{A.5}$$

Corner case:
Given their circumstances and the luck distribution, it is optimal for the agent to exert no effort if the optimal effort prescribed by solving problem (A.3) is too small to yield a positive probability of success. Using (A.2), the probability of success is positive if:

$$e_i > \frac{1-\kappa_i}{\overline{\lambda}} \tag{A.6}$$

Any positive optimal effort e_i must satisfy (A.6). If this condition does not hold, it is impossible for the agent to succeed even if they draw the highest possible luck $\overline{\lambda}$, and so they are better off exerting no effort. Note that even if the agent provides no effort, the expected payoff may still be positive. This occurs if they expect the third party to redistribute some income to the person who fail by choosing $1 - s_i > 0$.

Given e_i for our effort cost function, we plug in (A.5) on the left-hand side of (A.6) and can rewrite:

$$\kappa_i > 1 - \overline{\lambda}\left(\frac{x\,s_i}{b}\right)^{\frac{1}{2}} \tag{A.7}$$

This gives a lower bound on the agent's circumstances for which the agent exerts positive effort. If (A.7) does not hold, the circumstances is too small and the agent will never reach the target given the optimal effort induced by $c(e_i)$, hence chooses $e_i = 0$.

To sum up, the optimal effort is:

$$e_i = \text{f}\left(\frac{x\,(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}\right)^{\frac{1}{3}} \quad \text{if} \quad \text{(A.7)} \quad \text{holds}$$
$$0 \qquad \text{otherwise}$$

[3][17] use the power cost function $c(e_i) = \frac{b\,e_i^{1+\gamma}}{1+\gamma}$ to estimate their model parameters, using data from a real-effort experiment. They also consider an exponential cost function $c(e_i) = \frac{b\,exp(\gamma e_i)}{\gamma}$.

A.1.2 Proof of Proposition 2: Optimal Effort Comparative Statics

$$e_i = \left(\frac{x\,(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}\right)^{\frac{1}{3}}$$

Proposition Proposition 2 *If the optimal effort is positive, then it is increasing in x and s_i, and decreasing in b, κ_i and $\overline{\lambda}$.*

Proof.

$$\frac{\partial e_i}{\partial \kappa_i} = \underbrace{-\frac{1}{3}(1-\kappa_i)^{-\frac{2}{3}}}_{\leq 0}\left(\underbrace{\frac{x\,s_i}{b\,\overline{\lambda}}}_{>0}\right)^{\frac{1}{3}}$$

Recall that $\kappa_i \leq 1$ (the circumstances is never greater than the target, which I normalize to 1). Hence, $\frac{\partial e_i}{\partial \kappa_i} \leq 0$.

$$\frac{\partial e_i}{\partial x} = \underbrace{\frac{1}{3}x^{-\frac{2}{3}}}_{>0}\left(\underbrace{\frac{(1-\kappa_i)\,s_i}{b\,\overline{\lambda}}}_{\geq 0}\right)^{\frac{1}{3}}$$

Hence, $\frac{\partial e_i}{\partial x} \geq 0$.

$$\frac{\partial e_i}{\partial s_i} = \underbrace{\frac{1}{3}(s_i)^{-\frac{2}{3}}}_{>0}\left(\underbrace{\frac{x\,(1-\kappa_i)}{b\,\overline{\lambda}}}_{\geq 0}\right)^{\frac{1}{3}}$$

Hence, $\frac{\partial e_i}{\partial s_i} \geq 0$.

$$\frac{\partial e_i}{\partial b_i} = \underbrace{\frac{1}{3}(-\frac{1}{b_i^2})^{-\frac{2}{3}}}_{<0}\left(\underbrace{\frac{x(1-\kappa_i)\,s_i}{\overline{\lambda}}}_{\geq 0}\right)^{\frac{1}{3}}$$

Assuming $b_i > 0$, $\frac{\partial e_i}{\partial b_i^s} \leq 0$.

$$\frac{\partial e_i}{\partial \overline{\lambda}} = \underbrace{\frac{1}{3}(-\frac{1}{\overline{\lambda}^2})^{-\frac{2}{3}}}_{<0}\left(\underbrace{\frac{x(1-\kappa_i)\,s_i}{b_i}}_{\geq 0}\right)^{\frac{1}{3}}$$

Hence, since $\overline{\lambda} > 0$, $\frac{\partial e_i}{\partial \overline{\lambda}} \leq 0$.

A.1.3 Proof of Propositions 3 and 4: Third Party's Conditional Beliefs

First, note that the unconditional beliefs about luck and circumstances are simply

$$\mathbb{E}[\lambda_i] = \frac{\overline{\lambda}}{2} \text{ and } \mathbb{E}[\kappa_i] = k$$

The expected optimal effort is:

$$\mathbb{E}[e_i] = \int_{\kappa-\delta}^{\kappa+\delta} \Big(\frac{x\,(1-\kappa_i)\,s_i}{b_i\,\overline{\lambda}}\Big)^{\frac{1}{3}}\,dk$$

$$= \Big(\frac{x\,s_i}{b_i\,\overline{\lambda}}\Big)^{\frac{1}{3}}\,\frac{3}{4}\Big((1-\kappa+\delta)^{\frac{4}{3}} - (1-\kappa-\delta)^{\frac{4}{3}}\Big)$$

Proposition 3 *The expected circumstances conditional on success and failure are respectively given by:*

$$\mathbb{E}[\kappa_i \mid P_i \geq 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i\,\frac{p(P_i \geq 1 \mid \kappa_i, e_i)\,p(\kappa_i)}{\int_{\kappa-\delta}^{\kappa+\delta} p(P_i \geq 1 \mid \kappa_i, e_i)\,p(\kappa_i)\,dk}\,dk \tag{A.8}$$

$$\mathbb{E}[\kappa_i \mid P_i < 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i\,\frac{p(P_i < 1 \mid \kappa_i, e_i)\,p(\kappa_i)}{\int_{\kappa-\delta}^{\kappa+\delta} p(P_i < 1 \mid \kappa_i, e_i)\,p(\kappa_i)\,dk}\,dk \tag{A.9}$$

Proposition 4 *The expected effort conditional on success and failure are respectively given by:*

$$\mathbb{E}[e_i \mid P_i \geq 1] = \Big(\frac{x\,(1-\mathbb{E}[\kappa_i \mid P_i \geq 1])\,s_i}{b\,\overline{\lambda}}\Big)^{\frac{1}{3}} \tag{A.10}$$

$$\mathbb{E}[e_i \mid P_i < 1] = \Big(\frac{x\,(1-\mathbb{E}[\kappa_i \mid P_i < 1])\,s_i}{b\,\overline{\lambda}}\Big)^{\frac{1}{3}} \tag{A.11}$$

Proofs:

The only information available to the third party is whether the agent reach the target or not. Conditional on observing success or failure, the third party will form a belief about the circumstances of the agent. Upon seeing an agent succeeds, the expected circumstances is:

$$\mathbb{E}[\kappa_i \mid P_i \geq 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i\,p(\kappa_i \mid P_i \geq 1)\,dk$$

$$= \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i\,\frac{p(P_i \geq 1 \mid \kappa_i, e_i)\,p(\kappa_i)}{\int_{\kappa-\delta}^{\kappa+\delta} p(P_i \geq 1 \mid \kappa_i, e_i)\,p(\kappa_i)\,dk}\,dk$$

where $p(P_i \geq 1 \mid \kappa_i, e_i) = \frac{e_i\bar{\lambda}+\kappa_i-1}{e_i\bar{\lambda}}$.

If an agent fails, the third party expects the agent to have drawn a circumstances of

$$\mathbb{E}[\kappa_i \mid P_i < 1] = \int_{\kappa-\delta}^{\kappa+\delta} \kappa_i \, p(\kappa_i \mid P_i < 1) \, dk$$

Finally, the expected (optimal) effort conditional on success or failure is just found by plugging the expected circumstances in the optimal effort function.

$$\mathbb{E}[e_i \mid P_i \geq 1] = \left(\frac{x\,(1 - \mathbb{E}[\kappa_i \mid P_i \geq 1])\, s_i}{b_i\,\bar{\lambda}}\right)^{\frac{1}{3}}$$

Influence of effort cost and luck distribution on the third party's conditional beliefs. The third party's inference are also influenced by changes in the individual cost of effort b_i and the distribution of luck. Increasing the cost of effort b_i decreases e_i, thus reducing the probability of success for all κ_i. As a result, only those with high κ_i to begin with keep succeeding. This increases $\mathbb{E}[\kappa_i \mid P_i \geq 1]$, which causes $\mathbb{E}[e_i \mid P_i \geq 1]$ to drop. Decreasing the upper bound of the luck distribution $\bar{\lambda}$ has two opposite effects. First it increases the optimal effort for all κ_i, which increases the overall chance of success. On the other hand, a decrease in $\bar{\lambda}$ reduces the chance of getting a good λ_i draw which decreases the probability of success $p(P_i \geq 1 \mid \kappa_i, e_i)$. It is not clear which of the two effect dominates.

A.1.4 Optimal Effort When the Luck Lower Bound is Different From Zero

We follow the same reasoning as in A.1.1. When $s \sim \mathcal{U}\left[\underline{\lambda}, \bar{\lambda}\right]$, the probability that i succeeds is:

$$p = p(\lambda_i \geq \frac{1-\kappa_i}{e_i}) = 1 - p(\lambda_i < \frac{1-\kappa_i}{e_i}) = 1 - \frac{\frac{1-\kappa_i}{e_i} - \underline{\lambda}}{\bar{\lambda} - \underline{\lambda}} = \frac{e_i\bar{\lambda} + \kappa_i - 1}{e_i(\bar{\lambda} - \underline{\lambda})}$$

and the probability that j succeeds is

$$q = p\left(s_j \geq \frac{1 - \mathbb{E}[k_j]}{e_j}\right) = \frac{e_j\bar{\lambda} + \kappa - 1}{e_j(\bar{\lambda} - \underline{\lambda})}$$

Substituting p and q in (A.3) and maximizing with respect to e_i gives:

$$c'(e_i) = \frac{x\,(1-\kappa_i)\, s_i}{e_i^2\,(\bar{\lambda} - \underline{\lambda})} \tag{A.12}$$

Assuming that $c(e_i) = b_i \frac{e_i^2}{2}$, (A.12) becomes:

$$e_i = \left(\frac{x\,(1-\kappa_i)}{b_i\,(\bar{\lambda}-\underline{\lambda})} s_i \right)^{\frac{1}{3}}$$

A.1.5 Optimal Effort With Known Circumstances and Known Luck

When there is no uncertainty about λ_i, the agent knows exactly which level of effort is required to reach the target. Hence, if $\mathbb{E}[\pi_i \mid s_i, x, e_i, \kappa_i, \lambda_i] \geq c(e_i)$, the optimal effort is $e_i = 1$ when $\kappa_i = 0$, and $e_i = 1 - \kappa_i$ when $\kappa_i > 0$. If $\mathbb{E}[\pi_i \mid s_i, x, e_i, \kappa_i, \lambda_i] < c(e_i)$, then $e_i = 0$.

A.2 Additional Analyses and Robustness Checks

A.2.1 Effect of Ability, Exhaustion and Time-Limit on Effort

In this section, I show that ability, fatigue and time constraints have little effect on worker's effort. Thus, knowing that participants achieve their desired effort level, we can compare effort between rounds and treatments.

Ability: I compute for each participant the average time spent per slider in each round. For example, a participant who correctly positions 200 sliders in 6 minutes (360 seconds) has an average time per slider of 1.8 seconds. Figure A.1 display the distribution of the best average time per slider for each participant. I use this as a proxy for participant's ability, as it measures the maximal efficiency a participant is able to achieve across all their rounds. At their best, around two-thirds of participants spend between 1.75 and 2.25 seconds per slider.[4]

Exhaustion and time constraint: Next, within subject, effort should not be affected by exhaustion or time constraint, so that any effort variation can be attributed to changes in the experiment parameters i.e., circumstances and luck. There is evidence that time is rarely binding, and of a limited effect of exhaustion. In the stage 1, most participants finish before the end of the 45 minutes that is allowed to do the five rounds. Participants spend on average 38 minutes and 45 seconds in the first round. Only four participants spend no time in the last round of the first stage because they run out of time, having spent the 45 minutes in the first four rounds. Around 93.1% of participants enter round 5 with more than four minutes left to complete the round. Therefore for the vast majority of our subjects, time is not binding. Next, figure A.2 display for each round the average number of sliders solved and the average time spent per slider. Focusing on stage 1, there is no evidence of

[4]More precisely 56 out of 87 (64.3%) participants spend between 1.75 and 2.25 seconds per sliders. Ninety percent of participants spend between 1.7 and 2.7 seconds per sliders.

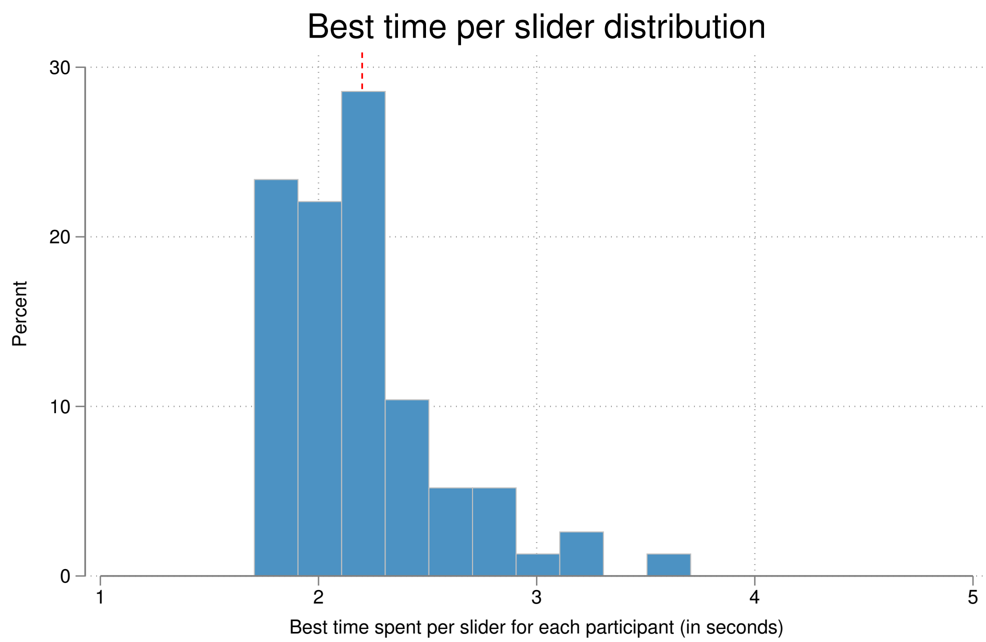


Figure A.1: Distribution of participant's maximal efficiency: smallest time spend per slider across all rounds for each participant

fatigue occurring in our experiment.[5] If anything, it seems that people are getting slightly more efficient with time as they spent 2.8 seconds per slider in the first round and only 2.38 seconds in round five. Those evidences suggest that the difference in within-subject effort levels observed in this experiment are the result of endogenous effort choice driven by our experiment parameters (while between-subject differences are due to unobserved individual effort cost), and not difference in abilities, fatigue or time constraint. In the next section, we show that individual effort in fact responds to changes in our parameters, and that effort is different across treatments.

A.2.2 Beliefs Conditional on The Outcome for all Treatments

[5] I focus on the first stage because participants spend around 10 minutes between the end of the stage 1 and stage 3 in which the last round of slider task occur. This delay may help people replenish their cognitive resources which could partially explain an increase in productivity in the last round.

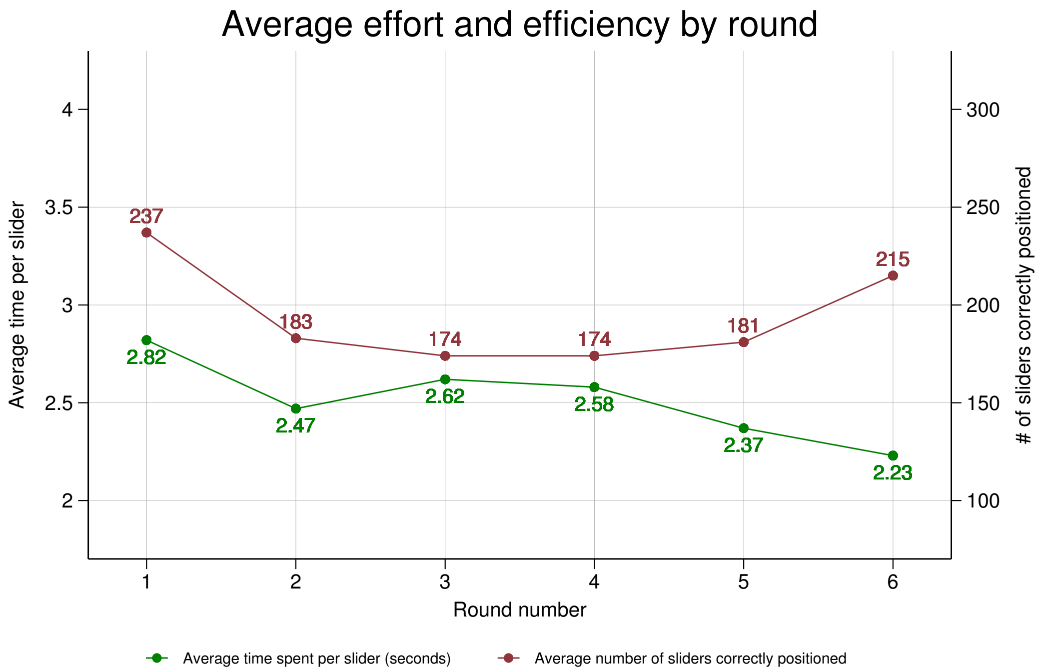


Figure A.2: Average number of sliders solved and average time spent per slider for each round
Notes: rounds 1 to 5 occur in the first stage, while round 6 occurs in stage three at the end of the experiment.

Table A.1: Beliefs conditional on *success*

		Treatment			Absolute difference	
		IL	IH	Heter.	IL vs IH	IL vs Heter.
Average effort if succeed	Truth	259	228	216	31	43
	Belief	215	197	205	18	10
	Δ Belief − Truth	−44	−31	−11		
Average luck if succeed	Truth	7.4	12.5	6.8	5.1	0.6
	Belief	6.8	11.9	6.5	5.1	0.3
	Δ Belief − Truth	−0.6	−0.6	−0.3		
Average circumstances if succeed	Truth	1250	1250	1649	0	399
	Belief	n/a	n/a	1497	n/a	n/a
	Δ Belief − Truth	n/a	n/a	−152		

Two-sample t-test: $p < 0.10$, $p < 0.05$, $p < 0.01$

Table A.2: Beliefs conditional on *failure*

		Treatment			Absolute difference	
		IL	IH	Heter.	IL vs IH	IL vs Heter.
Average effort if fail	Truth	190	161	154	29	37
	Belief	164	151	175	13	11
	Δ Belief − Truth	−26	−10	+21		
Average luck if fail	Truth	3.0	4.8	3.7	1.8	0.7
	Belief	3.6	6.1	3.8	2.5	0.3
	Δ Belief − Truth	+0.6	+1.3	+0.1		
Average circumstances if fail	Truth	1250	1250	843	0	407
	Belief	n/a	n/a	868	n/a	n/a
	Δ Belief − Truth	n/a	n/a	+25		

Two-sample t-test: $p < 0.10$, $p < 0.05$, $p < 0.01$

A.2.3 Heterogeneous Circumstances Treatment: Beliefs Conditional on the Outcome

The table below classify participants according to their beliefs about the circumstances and the level of effort made by the participants, depending on their outcome. For example, 55 participants believe that those who succeed on average have better circumstances *and* exert more effort than those who fail.

Table A.3: Subject classification according to their beliefs conditional on the outcome (Heterogeneous treatment).

		Circumstances belief			
		Better if succeed	Same	Worse if succeed	**Total (%)**
Effort belief	Higher if succeed	55	2	4	**61 (70.1)**
	Same	4	0	0	**4 (4.6)**
	Lower if succeed	20	0	2	**22 (25.3)**
	Total (%)	**79 (90.8)**	**2 (2.3)**	**6 (6.9)**	**87 (100)**

Table A.4: Beliefs conditional on outcome (Heterogeneous)

		Outcome		Difference
		Succeed	Fail	Succeed vs Fail
Average effort	Truth	216	154	63
	Belief	205	175	30
	Δ Belief − Truth	−11	+21	
Average luck	Truth	6.8	3.7	3.1
	Belief	6.5	3.8	2.6
	Δ Belief − Truth	−0.3	+0.1	
Average circumstances	Truth	1649	843	807
	Belief	1497	868	629
	Δ Belief − Truth	−152	+25	

Two sample t-test (truth) and paired t-test (beliefs): $p < 0.10$, $p < 0.05$, $p < 0.01$

A.2.4 Effect of Participants' Outcome on Their Belief

Does participants' performance affect their conditional beliefs? We identify two instances where conditional beliefs seem affected by the participant's outcome. In the baseline treatment, the belief about the average effort exerted by workers who succeed is significantly lower among participants who fail to reach the target compared to participants who succeed (205 vs 228, $p < 0.05$). In the *Heterogeneous* treatment, participants who miss the target provide a lower estimate for the average circumstances of workers who fail compared to their peers who reach the target (733 vs 982, $p < 0.01$).[6]

A.2.5 Fairness Preferences

The table below classify participants according to their redistribution behavior in the *luck only* and *effort only* benchmark cases. In *effort only* the redistributed share corresponds to the share sent from the best performer to the worst performer. In *luck only* the share corresponds to the share sent from the lucky to the unlucky. Third parties can redistribute up to 100 points (worth $1 in the experiment), hence this renders the interpretation of the redistributed amount of points straightforward in terms of percentage.

[6]Both types of participants underestimate the truth in the baseline treatment, which is 259. In the *Heterogeneous* treatment, the average circumstances of workers who miss is 843, so participants who miss the target underestimate the circumstances of workers who fail, while participants who reach the target overestimate them.

Table A.5: Redistribution behavior in the benchmark cases

		redistributed share: *effort only*				
		0	1-49	50	51-100	**Total (%)**
redistributed share: *luck only*	0	4	0	1	1	**6 (6.9)**
	1-49	5	9	2	2	**18 (20.7)**
	50	13	25	11	7	**56 (64.3)**
	51-100	0	2	1	4	**7 (8.0)**
	Total (%)	**22 (25.3)**	**36 (41.4)**	**15 (17.2)**	**14 (16.1)**	**87 (100)**

A.2.6 Redistribution Conditional on the Outcome

The table below reports the results of a regression for each separate treatment. Note that I do not use the circumstances beliefs or difference in circumstances beliefs as regressors for the treatments with identical circumstances, precisely because circumstances are fixed in those cases.

Table A.6: Regression of Belief and Preference on Redistributed Share

	Identical-Low		*Identical-High*		*Heterogeneous*	
	(1)	(2)	(3)	(4)	(5)	(6)
Belief difference in effort	−0.052		−0.074		−0.072	
	(0.038)		(0.042)		(0.050)	
Belief difference in luck	2.996		1.034		−0.962	
	(1.292)		(0.722)		(1.161)	
Belief difference in circumstances					−0.000	
					(0.556)	
Belief average effort if succeed		−0.073		−0.081		−0.108
		(0.054)		(0.064)		(0.065)
Belief average effort if fail		0.045		0.074		0.067
		(0.044)		(0.042)		(0.050)
Belief average luck if succeed		4.078		1.544		−2.175
		(1.860)		0.946		(1.500)
Belief average luck if fail		−1.465		−0.676		−0.585
		(2.143)		(0.827)		(1.610)
Belief average circumstances if succeed						0.002
						(0.006)
Belief average circumstances if fail						0.003
						(0.008)
Luck only redistribution	0.366	0.358	0.469	0.0492	0.270	0.265
	(0.199)	(0.198)	(0.110)	(0.120)	(0.149)	(0.157)
Effort only redistribution	0.174	0.185	0.080	0.079	0.220	0.192
	(0.120)	(0.122)	(0.082)	(0.082)	(0.104)	(0.107)
Constant	5.808	−1.376	7.933	0.095	22.001	40.021
	(8.400)	(19.124)	(8.366)	(15.864)	(8.290)	(17.711)
Observations	87	87	87	87	87	87

$p < 0.10$, $p < 0.05$, $p < 0.01$. Robust standard errors in parentheses.

Notes: the ”Belief difference in effort” independent variable is the difference between the expected effort of workers who succeeds and the expected effort of workers who fail. Same goes for the ”Belief difference in luck” and ”Belief difference in circumstances” beliefs. ”*Luck only* redistribution” refers to the amount transferred from the lucky to the unlucky worker in the *luck only* benchmark case. ”*Effort only* redistribution” refers to the amount transferred from the best to the worst performer in the *effort only* benchmark case. A Breusch-Pagan test for heteroskedasticity was not significant for the *Identical-High* and *Hetergogeneous* treatment, but was highly significant for the *Idential-Low* treatment. Hence, I adopt a conservative approach and use robust standard errors for all the treatments.

A.3 Experiment Instructions

Welcome to the experiment. You will now take part in an experiment on decision making. In addition to the show-up fee of $5, you can earn additional money which will be paid in cash at the end of the experiment. How much you earn depends on your decisions, the decisions of others, and on chance.

Instructions

Overview

There are **three stages** in this experiment. We will give you instructions at the beginning of each stage.

This **first stage consists of 5 rounds**. During each round, you will perform the same task.

The task involves a screen with 300 sliders. Each slider is initially positioned randomly. You can use the mouse in any way you like to move each slider. You can readjust the position of each slider as many times as you like. Your *slider score* will be determined by the number of sliders that are positioned exactly in the middle when you decide to stop working on the task. A green dot will appear on a slider when it is correctly positioned. How many sliders you decide to correctly position is up to you. You don't have to correctly position all the sliders in each task.

You may take as much time as you want working on a task, but you have 45 minutes to finish the 5 rounds of the first stage. A timer will indicate how much time you have left. If you finish early, you may use your phone silently while others finish the stage.

In each round, **if you reach a production target of 2500 you will receive 100 points. If you don't reach the target you will receive 0 points.** Your production is calculated as follows:

1. First, in each round you will receive a starting line, which is a randomly selected integer with a minimum value of 0 but never greater than the target. You will always know what your starting line is before adjusting the sliders.
2. Next you will receive a multiplier value, which is a randomly selected integer with a minimum value of 0 but never more than 20. Unlike the starting line, you will NOT know the value of your multiplier before adjusting the sliders.
3. After you complete the task, your production will be calculated as follows: first we will multiply your slider score (the number of sliders correctly positioned) by your multiplier. Then we will add your starting line. The formula is as follows:

Production = Starting line + (Multiplier * Slider score)

Examples

Example 1:
Suppose your randomly selected starting line (which you will know) is **1500**, and that your

randomly selected multiplier (which you won't know before adjusting the sliders) is **5.**

If you correctly position 100 sliders, your slider score is **100**. Then your production will be **1500** + **5***100** = 2000. Since your production is less than the target production of 2500, you miss the target and receive 0 points.

Instead, if you correctly position 250 sliders, your slider score is **250**. Then your production will be **1500** + **5***250** = 2750. Since your production is equal to the target of 2500, you reach the target and receive 100 points.

As shown in this example, you might not need to correctly position all the 300 sliders to reach the target.

Example 2:
Suppose your randomly selected starting line is **500**, and that you correctly position 300 sliders, so your slider score is **300**.

If your randomly selected multiplier is **5**, then your production will be **500** + **5***300** = 2000. Since your production is less than the target, you receive 0 points.
Instead, if your randomly selected multiplier is **10**, then your production will be **500** + **10***300** = 3500. Since your production is greater than the target, you receive 100 points.

As shown in this example, you might not reach the target even if you correctly position all the 300 sliders.

Starting line and multiplier

Ranges

You will always be told the possible ranges for your starting line and multiplier. But the ranges may change depending on the round, so it is important to check these at the beginning of each round because they will affect your production value.

In some rounds, your multiplier will be selected between 0 and 10. In other rounds it will be selected between 0 and 20.

In some rounds, your starting line is set at 1250. In other rounds it will be selected between 0 and 2500.

Random selection

When your multiplier or starting line is randomly selected in a given range, **each number in the range (boundary points included) has an equal probability of being selected.**

For instance, if your multiplier is randomly selected between 0 and 10, then each number (0, 1,..., 9, 10) has an equal chance (1/11) of being selected.

If your starting line is randomly selected between 0 and 2500, then each number (0, 1,..., 2499, 2500) has an equal chance (1/2501) of being selected.

Earnings

You can earn points in this experiment in two ways :

First, you can earn points based on whether you succeed or fail in reaching the target.

In each round, **if you reach the target you will receive 100 points.** If your production exceeds the target, it does not matter by how much your production exceeds the target. In all such cases you will receive the same 100 points.

On the other hand, **if you don't reach the target you will receive 0 points.** It doesn't matter whether you miss the target by a lot or almost reach it.

In addition, you can also earn points based on the accuracy of your guesses to some questions. Since you will be asked to make multiple guesses, you may in fact earn more points from your guesses than from reaching the target. For each guess, you will be paid according to the following rule:

100 points - Penalty

If the difference between your guess and the true value is less than 40, the penalty is equal to:

$$6\%*(\text{Guess} – \text{Truth})^2$$

Otherwise, the penalty equals 100 points.

The formula means that the closer your guess is to the true value, the smaller your penalty will be and the more points you will receive, regardless of whether your guess is above or below the true value. Since the maximum penalty is 100 points, you will never lose points if your guess is too far from the truth.

Example:
Suppose we ask you to guess the percentage of participants who reach the target in a given round. If 50% of participants reach the target, and your estimate is 70% (or 30%), the difference

is 20 percentage points. The penalty is 6%*(70-50)² = 6%*(20)² = 6%*400 = 24 points. Thus, you earn 100 - 24 = 76 points.

Using this formula, if 50% of people reach the target, your earnings from your guess are:

If your guess is	Difference between your guess and the truth	Penalty (pts)	Earnings (pts)
50	0	0	100
55 or 45	5	1.5	98.5
60 or 40	10	6	94
70 or 30	20	24	76
80 or 20	30	54	46
90 or 10	40	96	4
less than 10 or more than 90	>40	100	0

To calculate your final payment at the end of the experiment, we will convert your total number of points earned into dollars at a rate of 1 point = $0.01.
For instance, if you accumulated 1000 points over the course of the experiment, you will receive a payment of 1000*0.01=$10, in addition to the show-up fee, for a total of $15.

Stage 2

In the next screens, you will be presented with a series of decisions to make. Each decision involves sharing a bonus of 100 points between two participants.

These choices are hypothetical, meaning they won't affect your earnings or the earnings of other participants. However, try to answer the questions as if they had real consequences.

Stage 3

In this stage, there will be **one last slider task**. The time limit is 10 minutes. As before, if you reach the target you will receive 100 points. If you don't reach the target, you will receive 0 points.

After completing the slider task, **you will be paired with another participant**. A pair will always consist of one member who reached the target, and one who didn't. So, if you reached the target you will be paired with someone who did not, and vice versa.

Your pair will be randomly assigned to at least one other participant, who will decide how to allocate a bonus of 100 points between the two of you. The other participant will be told who reached the target in the pair, and their decision will directly affect your earnings.

You won't know the identities of the other person in your pair or the person making the redistribution decision. If your pair is shown to more than one participant, we will randomly choose one of the decisions made for your pair to determine the amount you will earn.

You will also have to decide how to allocate a bonus of 100 points between the members of one or two randomly chosen pairs of other participants. Your decision will directly affect their earnings.

You won't know the identities of these participants, and they won't know who made the redistribution decision for them. You will never have to make a redistribution decision for a pair to which you belong.

Additional information

We would now like to give you additional information regarding how likely it is to reach the target given different combinations of starting line and slider score. This is displayed in the table below, for the case that the multiplier value is between 0 and 10.

		Slider score						
		1	50	100	150	200	250	300
Starting line	0	0%	0%	0%	0%	0%	0%	17%
	500	0%	0%	0%	0%	0%	20%	33%
	1000	0%	0%	0%	0%	25%	40%	50%
	1500	0%	0%	0%	33%	50%	60%	67%
	2000	0%	0%	50%	67%	75%	80%	83%
	2500	100%	100%	100%	100%	100%	100%	100%

This table tells you, for instance, that participants who correctly position 200 sliders have a 75% chance of reaching the target with a starting line of 2000 (whether they succeed or not depends on the value of their multiplier). With the same slider score of 200, they have a 50% chance of reaching the target with a starting line of 1500, a 25% chance with a starting line of 1000, and a 0% chance with a starting line of 500 and below. **So, if two participants make the same effort and obtain the same slider score, the one who draws a higher starting line has a greater chance of reaching the target.**

Looking at the table, we also see that if participants want, for example, a 50% chance of reaching the target, then they must correctly position 300 sliders if their starting line is 1000, 200 if their starting line is 1500, and only 100 if their starting line is 2000. **So, participants drawing lower starting lines must put in more effort to have the same chance of reaching the target as those with higher starting lines.**

Appendix B: Correlation Neglect and Overconfidence

B.1 Model generalization

In this section we generalize the model with a state $\theta \sim \mathcal{N}\left[\theta_0, \sigma_0^2\right]$ and each signal (report) $r_i \sim \mathcal{N}\left[\theta, \sigma^2\right]$, $i \in [1, n]$. However, signals are correlated $corr(r_i, r_j)_{i,j \in [1,n], i6=j} = \rho$.

B.1.1 Posterior Belief Distribution

Assuming that people have a prior about the state which is distributed normally with mean θ_0 and variance σ_0^2. Assuming that each signal (report) r_i is distributed with mean θ and variance σ^2. After n signals, the posterior distribution for an agent who perceives a ρ_i correlation is proportional to

$$\mathcal{L}(\theta_1, ..., r_n) = \mathcal{L}(r_1, ..., r_n \mid \theta)\,\mathcal{L}_0(\theta)$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\begin{pmatrix}\theta - r_1\\ \cdots \\ \theta - r_n\end{pmatrix}^T \begin{pmatrix}1 & \rho_i & \cdots & \rho_i\\ \rho_i & 1 & \cdots & \rho_i\\ \vdots & \vdots & \ddots & \vdots\\ \rho_i & \cdots & \rho_i & 1\end{pmatrix}^{-1} \begin{pmatrix}\theta - r_1\\ \cdots \\ \theta - r_n\end{pmatrix}\right] \times \exp\left[-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right]$$

$$= \exp\left[-\frac{1}{2\sigma^2}\left(\frac{n_i\theta^2 - 2\theta\sum_i r_i}{1 + (n_i - 1)\rho_i} + C\right)\right] \times \exp\left[-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right]$$

, where C is a constant with respect to θ.

$$= \exp -\frac{1}{2}\left(\frac{\sigma_0^2(n_i\theta^2 - 2\theta\sum_i r_i)}{\sigma_0^2\sigma^2(1 + (n_i - 1)\rho_i)} + \frac{\sigma^2(1 + (n_i - 1)\rho_i)(\theta - \theta_0)^2)}{\sigma_0^2\sigma^2(1 + (n_i - 1)\rho_i)} + C\right)$$

$$= \exp -\frac{1}{2}\left(\frac{\sigma_0^2 n_i \theta^2 - 2\theta\sigma_0^2 \sum_i r_i + \theta^2\sigma^2(1+(n_i-1)\rho_i) - 2\theta\theta_0\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\sigma^2(1+(n_i-1)\rho_i)} + C\right)$$

Note that $\theta_0^2\sigma^2(1+(n_i-1)\rho_i)$ has been ignored in the numerator since it is constant w.r.t θ.

$$= \exp -\frac{1}{2}\left(\frac{\theta^2[\sigma_0^2 n_i + \sigma^2(1+(n_i-1)\rho_i)] - 2\theta[\sigma_0^2 \sum_i r_i + \theta_0\sigma^2(1+(n_i-1)\rho_i)]}{\sigma_0^2\sigma^2(1+(n_i-1)\rho_i)} + C\right)$$

$$= \exp -\frac{1}{2}\,\frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)}\left(\theta - \frac{\sigma_0^2 \sum_i r_i + \theta_0\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)}\right)^2$$

Which is distributed according to:

$$\sim \mathcal{N}\left[\frac{\sigma_0^2 \sum_i r_i + \theta_0\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)}, \frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)}\right]$$

,where $\frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i+\sigma^2(1+(n_i-1)\rho_i)}$ is the posterior *variance*. The posterior precision can be rewritten as: $\frac{1}{\sigma_0^2} + \frac{n_i}{\sigma^2(1+(n_i-1)\rho_i)}$

B.1.2 Posterior Variance Comparative Statics

Proposition 1:

1. *Fixing ρ_i, the posterior variance is decreasing convex in n*
2. *Fixing n, the posterior variance is increasing concave in ρ_i. Hence, decreasing convex in correlation neglect $\rho - \rho_i$.*

Proof: For simplicity, consider the variance after $n+1$ signals:

$$V(\theta_i + 1, \rho_i) = \frac{\sigma_0^2\,\sigma^2(1+n_i\rho_i)}{\sigma_0^2\,(n_i+1) + \sigma^2(1+n_i\rho_i)}$$

We assume that $n \geq 2$, $\rho_i \in [0,1)$ and $\sigma_0^2, \sigma^2 > 0$

$$\frac{\partial V}{\partial \rho_i} = \frac{\sigma_0^4\,\sigma^2 n_i(n_i+1)}{(\sigma_0^2\,(n_i+1)+\sigma^2(1+n_i\rho_i))^2} > 0$$

$$\frac{\partial V}{\partial n_i} = \frac{\sigma_0^4\,\sigma^2(\rho_i-1)}{(\sigma_0^2\,(n_i+1)+\sigma^2(1+n_i\rho_i))^2} \leq 0$$

$$\frac{\partial^2 V}{\partial \rho_i^2} = \frac{-2\sigma_0^4\,\sigma^2 n_i^2(n_i+1)}{(\sigma_0^2\,(n_i+1)+\sigma^2(1+n_i\rho_i))^3} < 0$$

$$\frac{\partial^2 V}{\partial n_i^2} = \frac{-2\sigma_0^4\,\sigma^2(\rho_i-1)(\sigma_0^2+\sigma^2\rho_i)}{(\sigma_0^2\,(n_i+1)+\sigma^2(1+n_i\rho_i))^3} \geq 0$$

$$\frac{\partial^2 V}{\partial n_i \partial \rho_i} = \frac{\sigma_0^6\sigma^2(n_i+1)+\sigma_0^4\sigma^4+\sigma_0^2\sigma^4 n_i(2+\sigma^2\rho_i-2\rho_i)}{(\sigma_0^2\,(n_i+1)+\sigma^2(1+n_i\rho_i))^3} > 0$$

B.1.3 Posterior Variance Asymptotic Behavior

Proposition 2:
As n goes to infinity, the variance limit is:

$$\lim_{n\to\infty} Var(\theta,\rho) = \frac{\sigma_0^2\,\sigma^2\,\rho}{\sigma_0^2+\sigma^2\,\rho}$$

Proof: Let $V(\theta_i,\rho) = \frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho)}{\sigma_0^2\,n_i+\sigma^2(1+(n_i-1)\rho)} \equiv \frac{f(n)}{g(n)}$

Using L'Hôpital's rule, the limit of the true variance is:

$$\lim_{n\to\infty} V(\theta_i,\rho) = \frac{f'(n)}{g'(n)} = \frac{\sigma_0^2\,\sigma^2\,\rho}{\sigma_0^2+\sigma^2\,\rho}$$

Because of the correlation between signals, there will always remain some uncertainty about the state even if an agent receives infinitely many signals. This holds even for a Bayesian who correctly perceives that ($\rho = \rho_i$), or for a neglecter. By contrast if there were no correlation between signals i.e. $\rho = 0$, a Bayesian would eventually end up learning the state and with a posterior variance of 0.

B.1.4 Marginal Benefit of Extra Signal

Proposition 3: *After receiving n signals, the marginal benefit of an extra signal is:*

$$Var(\theta,\rho_i) - Var(\theta+1,\rho_i)$$

Proof.

We follow here [25] (we refer to their proposition 8). First, note that the agent's utility after receiving n signals is:

$$-\int (a_i-\theta)^2 f(\theta)d\theta = -a_i^2 - \mathbb{E}(\theta^2) + 2a_i\mathbb{E}(\theta)$$

Since $Var(\theta) = \mathbb{E}(\theta^2) - \mathbb{E}(\theta)^2$, then the utility can be rewritten as:

$$\begin{aligned}
&- a_i^2 - \left(Var(\theta) + \mathbb{E}(\theta)^2\right) + 2a_i\mathbb{E}(\theta) \\
&= -a_i^2 - \mathbb{E}(\theta)^2 + 2a_i\mathbb{E}(\theta) - Var(\theta) \\
&= -\left(a_i - \mathbb{E}(\theta)\right)^2 - Var(\theta) \\
&= -\left(a_i - \mathbb{E}(\theta)\right)^2 - Var(\theta)
\end{aligned}$$

Therefore, the marginal benefit of the $n+1^{th}$ signal for an agent who perceives a correlation of ρ_i is $Var(\theta,\rho_i) - Var(\theta+1,\rho_i)$.

B.1.5 Overconfidence Change With Correlation Neglect and n

Proposition 4:

1. *Fixing ρ_i, overconfidence is increasing in n*
2. *Fixing n, overconfidence is decreasing in ρ_i. Hence, it is increasing in correlation neglect $\rho - \rho_i$*

Proof:

$$OC_{n_i} \equiv V(\theta_i,\rho) - V(\theta_i,\rho_i)$$

$$= \frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho)} - \frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)}$$

$$= \frac{\sigma_0^4\,\sigma^2\,n_i(n_i-1)(\rho-\rho_i)}{[\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho)][\sigma_0^2\,n_i + \sigma^2(1+(n_i-1)\rho_i)]}$$

Fixing n, overconfidence is decreasing in ρ_i, i.e. increasing in CN $(\rho-\rho_i)$ since $V(\theta_i,\rho_i)$ is increasing in ρ_i.

What about the change in overconfidence with n? The overconfidence change after receiving one extra signal is:

$$OC_{n_i+1} - OC_{n_i} = V(\theta_i + 1, \rho) - V(\theta_i + 1, \rho_i) - V(\theta_i, p) + V(\theta_i, \rho_i)$$

$$= V(\theta_i + 1, \rho) - V(\theta_i, \rho) - (V(\theta_i + 1, \rho_i) - V(\theta_i, \rho_i))$$

We can rewrite:

$$V(\theta_i + 1, \rho) - V(\theta_i, \rho) = \frac{\sigma_0^2 \sigma^2(1 + n_i\rho)}{\sigma_0^2 (n_i + 1) + \sigma^2(1 + n_i\rho)} - \frac{\sigma_0^2 \sigma^2(1 + (n_i - 1)\rho)}{\sigma_0^2 n_i + \sigma^2(1 + (n_i - 1)\rho)}$$

$$= \frac{\sigma_0^4 \sigma^2(\rho - 1)}{[\sigma_0^2 n_i + \sigma^2(1 + (n_i - 1)\rho)][\sigma_0^2 (n_i + 1) + \sigma^2(1 + n_i\rho)]}$$

and

$$V(\theta_i + 1, \rho_i) - V(\theta_i, \rho_i) = \frac{\sigma_0^4 \sigma^2(\rho_i - 1)}{[\sigma_0^2 n_i + \sigma^2(1 + (n_i - 1)\rho_i)][\sigma_0^2 (n_i + 1) + \sigma^2(1 + n_i\rho_i)]}$$

So, the change in overconfidence caused by one extra signal is:

$$OC_{n_i+1} - OC_{n_i} =$$

$$\frac{\sigma_0^4 \sigma^2(\sigma_0^4 n^2 + \sigma_0^4 n + 2\sigma_0^2 \sigma^2 n + \sigma_0^2 \sigma^2 + \sigma^4)(\rho_i - \rho)}{[\sigma_0^2 n_i + \sigma^2(1 + (n_i - 1)\rho)][\sigma_0^2 (n_i + 1) + \sigma^2(1 + n_i\rho)][\sigma_0^2 n_i + \sigma^2(1 + (n_i - 1)\rho_i)][\sigma_0^2 (n_i + 1) + \sigma^2(1 + n_i\rho_i)]} > 0$$

Hence, overconfidence is increasing in n.

B.1.6 Overconfidence Asymptotic Behavior

Recall that the variance is bounded:

$$\lim_{n\to\infty} V(\theta_i, \rho) = \frac{\sigma_0^2 \sigma^2 \rho}{\sigma_0^2 + \sigma^2 \rho}$$

So, the overconfidence limit is:

$$\lim_{n\to\infty} OC_{n_i} = \frac{\sigma_0^2 \sigma^2 \rho}{\sigma_0^2 + \sigma^2 \rho} - \frac{\sigma_0^2 \sigma^2 \rho_i}{\sigma_0^2 + \sigma^2 \rho_i} = \frac{\sigma_0^4 \sigma^2 (\rho - \rho_i)}{(\sigma_0^2 + \sigma^2 \rho)(\sigma_0^2 + \sigma^2 \rho_i)}$$

B.2 Additional Analyses

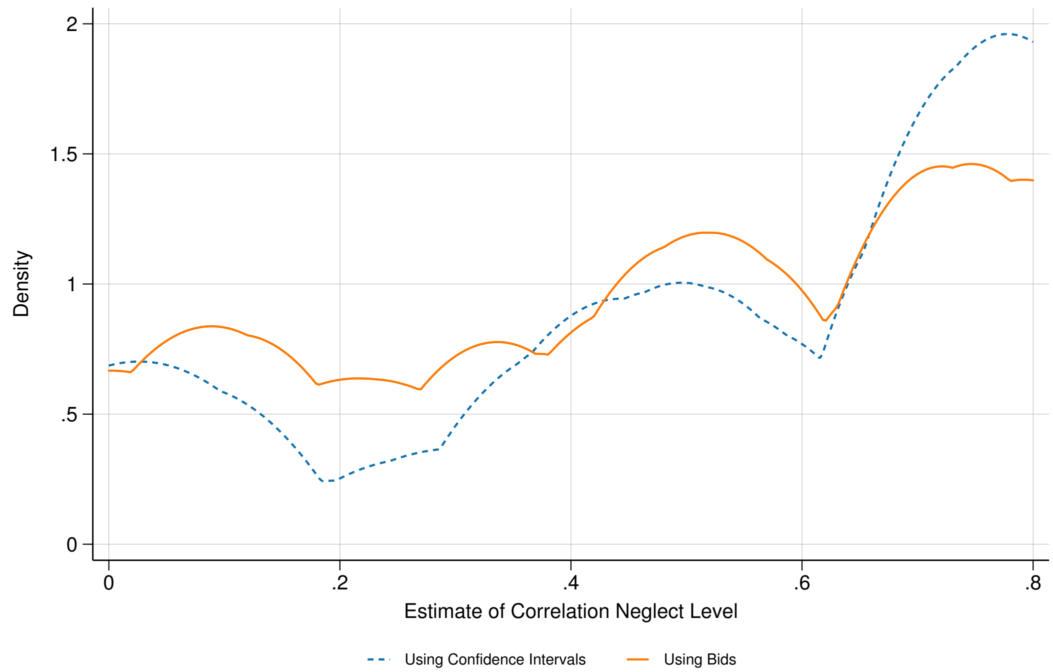


Figure B.1: Kernel Density Estimate of Correlation Neglect

Table B.1: Willingness to Pay for Three Additional Reports Conditional on First Bid Outcome

	Lose first bid		Win first bid	
CN level	Mean WTP (SD)	Obs.	Mean WTP (SD)	Obs.
CN≤ 0.2	$.37 (.39)	20	$.2 (0)	1
0.2 <CN≤ 0.4	$.66 (.38)	59	$.55 (.45)	4
0.4 <CN≤ 0.6	$.53 (.46)	97	$.50 (.46)	23
0.6 <CN< 0.8	$.66 (.38)	88	$.84 (.75)	24
CN=0.8	$.98 (.41)	23	$.73 (.51)	9
Total	$.62 (.43)	287	$.67 (.43)	61

B.3 Base-Rate Neglect

What is the posterior if we allow for base-rate neglect (below, BRN), which occurs when agents underweight their prior by $\delta \in (0, 1)$.

After 1 report:

$$\mathcal{L}(\theta_1) = \mathcal{L}(r_1 \mid \theta)\,\mathcal{L}_0(\theta)^{\delta}$$

$$= \exp -\frac{1}{2\sigma^2}\left(\frac{n_i\theta^2 - 2\theta r_1}{1+(n_i-1)\rho_i} + C\right) \times \exp\left[-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right]^{\delta}$$

which follows:

$$\sim \mathcal{N}\left[\frac{\sigma_0^2 r_1 + \delta\,\theta_0\,\sigma^2(1+\rho_i)}{\sigma_0^2\; + \sigma^2(1+\rho_i)}, \frac{\sigma_0^2\; + \sigma^2(1+\rho_i)}{\sigma_0^2 + \delta\sigma^2(1+\rho_i)}\right]$$

After n reports:

$$\mathcal{L}(\theta_1, \ldots, r_n) = \mathcal{L}(r_1, \ldots, r_n \mid \theta)\,\mathcal{L}_0(\theta)^{\delta^n}$$

$$= \exp -\frac{1}{2\sigma^2}\left(\frac{n_i\theta^2 - 2\theta\sum_i r_i}{1+(n_i-1)\rho_i} + C\right) \times \exp\left[-\frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right]^{\delta^n}$$

which follows:

$$\sim \mathcal{N}\left[\frac{\sigma_0^2\sum_i r_i + \delta^n\,\theta_0\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2 n_i\; + \delta^n\sigma^2(1+(n_i-1)\rho_i)}, \frac{\sigma_0^2\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2 n_i\; + \delta^n\sigma^2(1+(n_i-1)\rho_i)}\right]$$

Note that when $\delta = 1$, we revert to the former distribution without BRN. The variance denominator is decreasing for high base-rate neglecter. This implies that the variance of a base-rate neglecter is actually increasing in BRN. The smallest posterior variance are reported by agents that are full correlation neglecters agents but **not** base-rate neglecters. The larger variance would be reported by agents that are full base-rate neglecters but **not** correlation neglecters. Bayesians, and agents that have both to some extent CN and BRN have a variance between those two extremes.

An implication is that as $\delta \to 0$:

$$\frac{\sigma_0^2\sum_i r_i + \delta^n\,\theta_0\,\sigma^2(1+(n_i-1)\rho_i)}{\sigma_0^2 n_i\; + \delta^n\sigma^2(1+(n_i-1)\rho_i)} \to \frac{1}{n_i}\sum_i r_i$$

So, the full BRN estimate will be the average of the reports seen, regardless of the CN level. The CN point estimate will lie between the full BRN estimate and the bayesian report, which will be the closest to the prior. Finally, the closer the reports average is from the prior, the closer to each other are the BRN, CN and Bayesian estimates.

B.4 Experiment Instructions

[The following instructions were read out loud]

Welcome to the experiment!

You will now take part in an economic experiment. In addition to the show-up fee of $5, you can earn additional money which will be paid at the end of the experiment. How much you earn depends on your decisions.

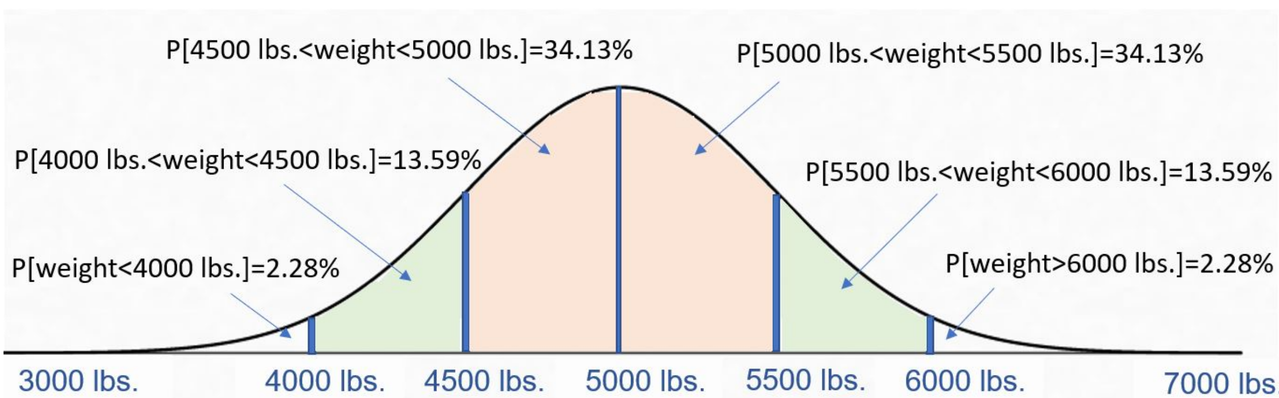
At the end of the instructions you will be asked to answer a few questions, so make sure that you listen carefully.

Remember that your participation is voluntary. You are free to leave this experiment at any time by simply closing the webpage. No payment will be made in that case.

Task

There are 7 rounds in the experiment. In each round, an imaginary box is randomly chosen by a computer program among 1000 (imaginary) boxes. Your task will be to find out the weight of the randomly chosen box.

Here is what we know about those 1000 boxes. First, we know that average weight of the 1000 boxes is 5000 lbs. There are 50% of boxes that weighs more than 5000 lbs, and 50% that weighs less than 5000 lbs. In addition, the graph below summarizes the percentage of boxes whose weight lies in a given range.



The weight of a box can be anywhere between 3000 and 7000 lbs, but it is more likely to be close to 5000 lbs. In fact, as shown in the graph by the orange area, there is around 68% of boxes that weighs between 4500 and 5500 lbs. There are around 14% of boxes that weighs between 4000 and 4500 lbs, and around 14% of boxes that weighs between 5500 and 6000 lbs. There are around 2% of boxes that weighs less than 4000 lbs, and around 2% of boxes that weighs more than 6000 lbs.

Formally, the weight of the boxes is normally distributed with mean 5000 and standard deviation 500.

Your task is to estimate the weight of the randomly chosen box, with the information that will be provided to you. **A new box will be randomly drawn in every round.**

In each of the 7 round, there are 4 stages, and in each stage you have to provide an estimation.

Learning about the weight of the box

In each round there is one red robot and many blue robots, which send reports about the weight of the randomly chosen box. The robots are different in each round. Each blue robot, which we will refer to by their number (robot 1, robot 2 etc.) generates its own report. This is how the reports of the blue robots are generated:

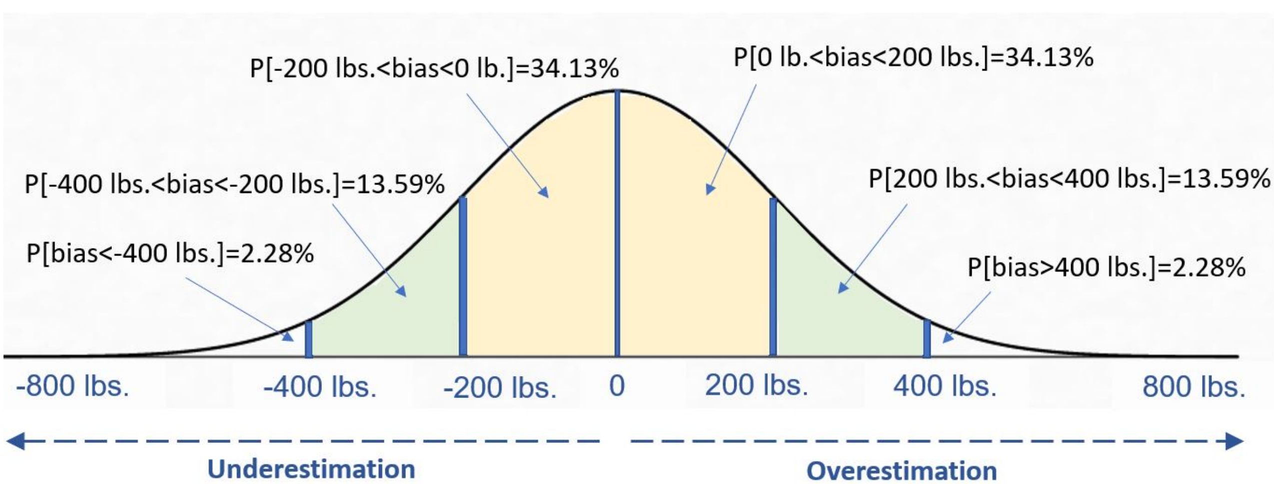
1) First, the randomly chosen box is weighted by the red robot. However, the red robot is biased. It can overestimate or underestimate the weight of the box, up to 800 lbs.

2) Next, blue robots one by one take a look at the weight displayed by the red robot. The red robot has a scale which displays the weight it measures to the blue robots. However the needle on the red robot scale is constantly moving around the (biased) weight, up to 400 lbs below or above it. Because the needle is moving, each blue robot potentially observes a different weight. So, they can also overestimate or underestimate the weight displayed by the red robot up to 400 lbs. The weight that each blue robots briefly sees is what it reports.

So, a blue robot report will be biased in two ways: one because of the red robot bias, and the other caused by the moving needle on the red robot scale.

Example: Suppose the randomly chosen box weighs 4430 lbs, and the red robot *overestimates* the weight by 200 lbs. If at the moment when blue robot 1 observes the red robot, the needle on the scale is 150 lbs *below* the weight estimated by the red robot, then robot 1 *underestimates* the weight by - 150 lbs. Then robot 1 report will be 4480 = 4430 + 200 − 150. If at the moment when blue robot 2 observes the red robot, the needle on the scale is 100 lbs *above* the weight estimated by the red robot, then robot 2 *overestimates* the weight by 100 lbs. Then robot 2 report will be 4730 = 4430 + 200 + 100.

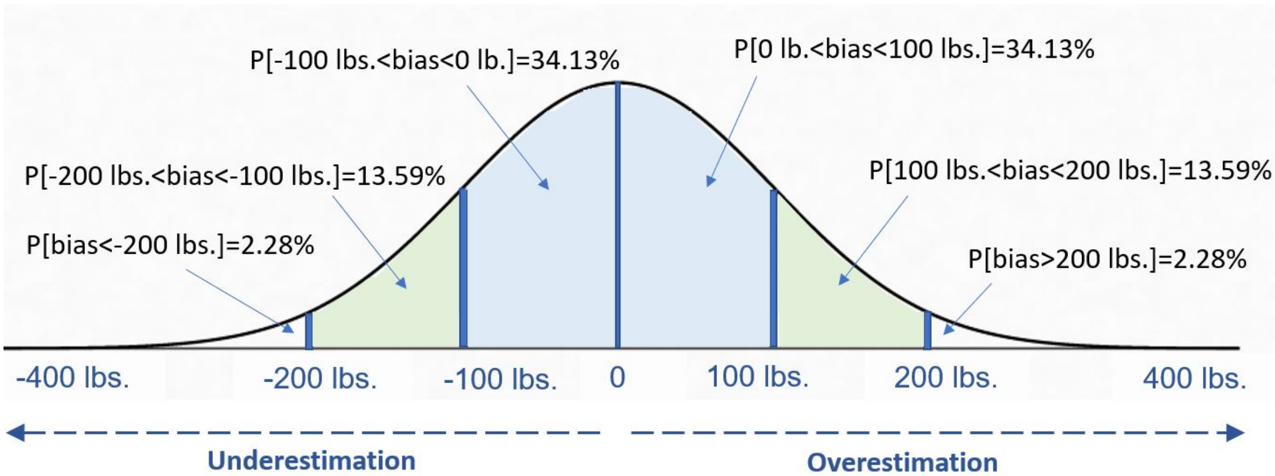
The picture below describes what are the chances that the **red robot bias** goes in a given direction and be a of certain size:



Probability that the red robot bias lies in each range

The red robot has 50% chance to overestimate the weight, and 50% chance to underestimate it. Each bias (either over or underestimation) can be up to 800 lbs, but it is more likely to be closer to 0 than 800 lbs. In fact, as shown in the picture by the area in yellow, there is around 68% chance that the red robot bias is between −200 and 200 lbs. There is 14% chance that the overestimation is between 200 and 400 lbs, and 2% of chance that it is bigger than 400 lbs. There is 14% chance that the underestimation is between −200 and −400 lbs, and 2% of chance that it is bigger than −400 lbs.

The picture below describes what are the chances that the **needle bias** goes in a given direction and be a of certain size:



Probability that the needle bias lies in each range

When each blue robot take a look at the needle on the red robot scale, the needle has 50% chance to be above the weight estimated by the red robot, and 50% chance to be below it. Each bias (either over or underestimation) can be up to 400 lbs, but it is more likely to be closer to 0 than 400 lbs. In fact, as shown in the picture by the area in blue, there is around 68% chance that the needle bias is between −100 and 100 lbs. There is 14% chance that the overestimation is between 100 and 200 lbs, and 2% of chance that it is bigger than 200 lbs. There is 14% chance that the underestimation is between −100 and −200 lbs, and 2% of chance that it is bigger than −200 lbs.

It is important to remember that the red robot bias will affect every reports the same way, but the bias caused by the needle will affect each blue robot report differently.

Bidding for information

At multiple times during this experiment you will have to submit a bid for one or several reports.

Your bid is the most you would be willing to pay for the report. It means that if you indicate a bid of \$B, you would buy the report if we offered it for a price less than or equal to \$B, and you would not buy the report if we offered it for a price greater than \$B. You can bid between \$0 to \$2. A price will be randomly drawn between \$0 and \$2. If your bid is above or equal to the price, you buy the report and pay the price, which will be deducted from your earnings. If your bid is below the price, you do not buy the report.

You are not in competition with other participants to submit the highest bid. The price randomly drawn is different for each participant. Anyone who bids more than the price randomly drawn for him/her buys the report(s).

Example: Suppose we ask you to bid for one of the robot report, for which you are willing to pay at most \$0.5. If you bid more - say \$0.6 - and the price randomly generated is \$0.55, you get the report but pay \$0.55. So you end up paying \$0.05 more than the most you are willing to pay. On the other hand, if you bid less than the most you would be willing to pay - say \$0.4 - and the price randomly generated is \$0.45, you will not get the report. But if you had bid \$0.5, you would have gotten the report and paid "only" \$0.45.

In summary, it is best under this system for you to bid your true maximum willingness to pay. You will never pay more and you could end up paying less. If you value the report at \$2, you should bid \$2. If you value the report at \$0, you should bid \$0.

Earnings

You can earn money in two ways in each stage of a round. At the end of each round, the computer will randomly select a stage and you will be paid according to the choices you made in that stage.

First you will be paid depending on your estimate of the weight of the box, according to the following rule:

$$\$2 - Penalty$$

If the difference between your estimate and the true weight is less than 300 lbs, the penalty is equal to:

$$\$0.00002\ (Estimate - Trueweight)^2$$

Otherwise the penalty equals \$2.

The formula simply means that the closer your estimate is from the true weight of the box, the smaller will be your penalty. It does not matter if your estimation is above or below the true weight. Because the penalty is at most \$2, you can never lose money if your estimation is too far from the truth (that is, above 300 lbs).

Example: If the randomly chosen box weighs 5000 lbs, and your estimate is either 4700 or 5300 lbs, the difference is 300 lbs. The penalty is $\$0.00002*(300)^2 = \$0.00002*90\,000 = \$1.8$. Therefore, you earn \$2 − \$1.8 = \$0.2.

Using this formula, if the box weighs 5000 lbs, your potential earnings are:

If your estimate is	Difference between your estimate and the box weight	Penalty	Earnings
5000 lbs	0 lb	\$0	\$2
4900 or 5100 lbs	100 lbs	\$0.2	\$1.8
4800 or 5200 lbs	200 lbs	\$0.8	\$1.2
4700 or 5300 lbs	300 lbs	\$1.8	\$0.2
less than 4700 or more than 5300 lbs	>300 lbs	\$2	\$0

Second, every time you make an estimate, you will also have to indicate your preference between two options. Option A gives you 90% chance to get $1. Option B gives you $1 if your estimate lies in a given range around the true box weight. You will make your choices in a table similar to the one displayed below.

For example, look at *choice 1* in the table. If you pick option A you get $1 with 90% chance. If you pick option B you get $1 if your estimate is less than 500 lbs away from the box weight. Therefore, you should pick option B if you think there is more than 90% chance that your estimate is less than 500 lbs away from the true weight of the box.

The same is true for all choices. For *choice 2*, you should pick option B if you think there is more than 90% chance that your estimate is less than 510 lbs away from the true weight of the box. Otherwise, you should pick option A.

Choice	Option A	Option B
1	90% chance of $1	If your estimate is within ± 500 lbs of the box weight you get $1
2	90% chance of $1	If your estimate is within ± 510 lbs of the box weight you get $1
3	90% chance of $1	If your estimate is within ± 520 lbs of the box weight you get $1
4	90% chance of $1	If your estimate is within ± 530 lbs of the box weight you get $1
5	90% chance of $1	If your estimate is within ± 540 lbs of the box weight you get $1
6	90% chance of $1	If your estimate is within ± 550 lbs of the box weight you get $1
7	90% chance of $1	If your estimate is within ± 560 lbs of the box weight you get $1
8	90% chance of $1	If your estimate is within ± 570 lbs of the box weight you get $1
9	90% chance of $1	If your estimate is within ± 580 lbs of the box weight you get $1
10	90% chance of $1	If your estimate is within ± 590 lbs of the box weight you get $1
11	90% chance of $1	If your estimate is within ± 600 lbs of the box weight you get $1

You have to make a choice for each row. As soon as you pick option B in one row, all the choices below that row are filled automatically with option B for convenience. You can still modify your choice if you wish. You are allowed to pick A for all choices if you prefer 90% chance of $1 for each choice. You are allowed to pick B for all choices if you think that there is more than 90% chance that your estimate lies in the given range for each choice. One of the 11 choice will be randomly selected by the computer and you will be paid according to your chosen option.

To summarize, you can earn money in two ways in each stage. The first way is by submitting an estimate of the weight of the box. The other way is by choosing the option you prefer (A or B) for the 11 choices given to you.

Each round has 4 stages. At the end of each round, one of the 4 stage is randomly chosen to determine your earnings for that round.

Your final payoff will be the sum of all the money earned during the 7 rounds.

Sequence of stages in a round

There is 7 rounds and each round has 4 stages.

In the beginning of each round, a box is randomly drawn by the computer among the 1000 boxes. Next, the box is weighted by the red robot. Then, each blue robot generates its report.

In stage 1, you will be given two reports. Then you will submit your first estimate.

In stage 2, you will be able to bid for one extra report. If your bid is higher than the randomly generated price, you will see the extra report and revise your estimates. If your bid is lower, your payoff for that stage is determined using your stage 1 estimate.

In stage 3, you will be able to bid for 3 extra reports. If your bid is higher than the randomly generated price, you will see the 3 extra reports and revise your estimates. If your bid is lower, your payoff for that stage is determined using your latest estimate.

In stage 4, for a fee of $0.1 you will be given some information about the bids of 4 participants who took part in this experiment. You will be able to see the stage 3 estimate of the participant of your choice. Then you will be able to revise your estimates. If you do not wish to participate, your payoff for that stage is determined using your latest estimate.

[End of the instructions. Next, participants moved to the comprehension questions before starting round one.]