THE UNIVERSITY OF CHICAGO

ESSAYS IN BEHAVIORAL AND LABOR ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
ANTHONY DITTA

CHICAGO, ILLINOIS
MARCH 2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Many thanks to family, friends, classmates, teachers, advisors, and all combinations thereof.

ABSTRACT

The three chapters of this dissertation are independent of one another.

Gender, Risk, and Signaling: Preferences and Beliefs

In many signaling environments, female senders take fewer risks than male senders. To what extent are these different behaviors a result of different preferences versus different beliefs? In this chapter I present an experiment which parses out the effect of preferences and beliefs on risk taking by male and female senders in a signaling environment. In particular, I compare risk preferences to three beliefs: confidence in one's own abilities, general beliefs about risk, and anticipation of discrimination. Consistent with previous findings, I find that women take fewer risks than men in this signaling environment. However, there is no evidence for differences in risk preferences between men and women. Instead, there is strong evidence for differences in beliefs. In particular, women are less confident in their own abilities than men. There is some evidence for more pessimistic beliefs about risk among women than men and no evidence of anticipated discrimination.

Sense of Control and Willingness to Pay For Non-Instrumental Control

Control can be desirable for two reasons: its instrumental value and its intrinsic value. Separating the instrumental value of control from its intrinsic value can be difficult since most control comes with both. But, it is important to distinguish between them since they play very different roles in economic decision-making. To manage this problem, I focus on control with no instrumental value, which I call non-instrumental control. In particular, in this chapter I present an experiment in which I measure the effect of changes in participant's sense of control on their willingness to pay for non-instrumental control. The results are

inconclusive. The sense of control manipulation did not significantly affect participants' sense of control or their willingness to pay for non-instrumental control. However, participants' willingness to pay was significantly positive with notable and predictable heterogeneity.

The "Blue Wall of Silence" and Police Careers

Many police departments in the United States have an unwritten code "which prohibits disclosing perjury or other misconduct by fellow officers, or even testifying truthfully if the facts would implicate the conduct of a fellow officer" (Chin and Wells, 1997). In this chapter, I use administrative data on Chicago police officers and the investigation of complaints against them to create a proxy for instances of officers testifying for or against other officers. I validate the proxy using a subset of data for which validation is possible, and test for correlations between the proxy and important career outcomes: leaving the police force, being promoted to higher pay grades, being accused in future complaints, and having those complaints sustained in investigations. I find that officers who do not testify against other officers are less likely to leave the force and more likely to be accused in future complaints, and officers who do testify against other officers are more likely to have complaints sustained against them.

CHAPTER 1

GENDER, RISK, AND SIGNALING: PREFERENCES AND BELIEFS

1.1 Introduction

In many signaling environments, female senders take fewer risks than male senders. Female college students are more likely than male college students to change majors after a bad grade in physical sciences (Rask and Tiefenthaler, 2008) and economics (McEwan et al., 2021). Female surgeons take on less risky patients than their male peers (Sarsons, 2017). And, female economists write their papers with clearer language than male economists (Hengel, 2022).

To what extent are these different behaviors a result of different preferences versus different beliefs? There is good reason to believe that both play a role, but distinguishing between the two with observational data is challenging – especially if one allows beliefs to be inaccurate or their updating to be biased.

So, in this chapter I present an experiment which parses out the effect of preferences and beliefs on risk taking by male and female senders in a signaling environment. In particular, I compare risk preferences to three beliefs: confidence in one's own abilities, general beliefs about risk, and anticipation of discrimination.

In the experiment, there are two groups of participants: senders and receivers that are randomly paired together. The senders are randomly assigned to have their gender revealed or hidden from their receiver. They take a multiple-choice quiz, and provide their beliefs about their own performance and their second-order beliefs about how their receiver will believe they performed. They learn that their receiver will get a signal of their performance (results from a sub-sample of the quiz questions) and are given the choice between a flat payment or a risky payment based off of the receiver's beliefs about their performance after

the signal.

The preferences and beliefs are parsed as follows. A sender's stated beliefs about their quiz performance is their confidence in their own abilities. Any differences in second-order beliefs among senders whose gender is hidden are attributed to general beliefs about risk. Additional differences in second-order beliefs among senders whose gender is revealed are attributed to anticipation of discrimination. Differences in sender risk taking conditional on beliefs are attributed to differences in preferences.

Consistent with previous findings, I find that women take fewer risks than men in this signaling environment. However, there is no evidence for differences in risk preferences between men and women. Instead, there is strong evidence for differences in beliefs. In particular, women are less confident in their own abilities than men. There is some evidence for more pessimistic beliefs about risk among women than men and no evidence of anticipated discrimination.

1.1.1 Motivation

Risk in signaling appears in many domains. Signals contain risk if they are noisy or if the senders have imperfect information about themselves or the receivers of their signals. Consider the examples from the introduction. College students have imperfect information about their own ability in various subjects, and grades are a noisy signal of ability. Surgery outcomes are subject to chance and thus a noisy signal of surgeon ability. Writers cannot be certain how their writing will be interpreted by readers, so there is uncertainty about the receiver's response to signals from academic writing.

And, many behaviors normally thought of as risk contain signal value. For example, drug use and crime are often cited as prototypical risky behaviors, but they can also signal moral and ethical values to one's social group (Gallupe and Baron, 2014; Wikström, 2010). Even everyday risks like telling a joke or speaking up in class (Byrnes et al., 1999) signal a

person's ability to be funny or scholarly.

Distinguishing preferences from beliefs matters a great deal when considering welfare. If different preferences are the sole cause of differences in market outcomes, there is little scope for inefficiency or welfare improvements (at least in this simple environment). For example, economists often invoke differences in risk preferences as an explanation for differences in labor market outcomes (Bertrand, 2011; Croson and Gneezy, 2009; Eckel and Grossman, 2008b; Lindquist and Säve-Söderbergh, 2011). If preferences are indeed the root of the differences, markets are allocating efficiently, and policy intervention would be misguided. However, if beliefs play a role, there could be major welfare losses. This is most obvious when people's beliefs are inaccurate, but accurate beliefs can lead to welfare losses, too: for example with coordination failures in statistical discrimination models (Moro and Norman, 2004). And there are equity concerns: if women expect discrimination and behave accordingly, market outcomes may be unfair – even if the discrimination is accurate in equilibrium.

The potential for large welfare losses means that researchers and policy-makers would do well to be skeptical about attributing differences in behavior to differences in preferences. This is especially true as stereotyping claims such as "women are more risk averse than men" can persist even without sufficient evidence due to bias in individuals like confirmation bias and bias in groups such as publication bias (Nelson, 2014).

1.1.2 Related Literature

Risk preferences will influence signalling behavior if there is uncertainty as discussed in the examples above. A large body of experimental (Croson and Gneezy, 2009; Eckel and Grossman, 2008a) and observational (Bertrand, 2011) work purports to show that men and women have "fundamental differences" in risk preferences (Croson and Gneezy, 2009): women being more risk averse than men.[1] And, as mentioned above, risk preferences are often

1. These preferences and beliefs do not always imply less risk taking by signal senders. For example, they could affect the informativeness of the signal – perhaps so much so that populations with more risk-averse

invoked as an explanation for gender differences in market outcomes.

However, reviews in psychology (Byrnes et al., 1999) and economics (Nelson, 2014, 2015) find that this difference is actually quite varied and relatively small on average. The difference may also be sensitive to framing (Schubert et al., 1999). And, even to the extent that there is a difference in risk-taking behavior, these studies may not show differences in preferences as they do not account for differences in the "background risk" in men and women's lives (Guiso and Paiella, 2008). In addition, if risk preferences were fundamentally connected to gender, one would expect to find a biological connection.[2] One attempt to find such a connection is in hormone differences. For example (Sapienza et al., 2009) look at risk preferences and testosterone. However, their analysis is correlational and the results are more nuanced than simply more testosterone implying more risk seeking preferences. Experimental research finds similarly nuanced results (Cueva et al., 2015) or no association at all (Zethraeus et al., 2009). So, risk preferences alone are an imperfect explanation for gender differences in behavior.

Men and women may also differ in their beliefs about risk and signalling. I consider three possible differences in beliefs: confidence in one's own abilities, general beliefs about risks, and anticipation of discrimination.

With regard to their own abilities, men may be more overconfident than women. This has been found in both experimental (Niederle and Vesterlund, 2007) and observational (Barber and Odean, 2001; Bengtsson et al., 2005; Huang and Kisgen, 2013) environments. However, a recent meta-analysis finds no statistically distinguishable difference in overconfidence between men and women (Bandiera et al., 2022), so this effect may not be particularly robust.

General beliefs about risk are straightforward: women often perceive risks as more "risky"

---

preferences and beliefs take more risks in signaling. I rule out these effects in my experimental design, so I do not discuss them further in the chapter, however the interested reader may see more detail in Appendix 1.A.

2. This is true even as gender is socially constructed since there is a high correlation between the social constructs and biological features.

than men (Gustafson, 1998; Slovic, 1999). In other words, women have a more negative outlook about the potential outcomes of risk than men. This is typically shown with environmental and health risks (Davidson and Freudenburg, 1996; Flynn et al., 1994), but many of the proposed mechanisms (e.g. power, status, trust) apply to signaling, too.

Women may anticipate discrimination which dissuades risk in signaling.[3] For example, in many domains, women are praised less for good signals and/or punished more for bad signals than men (punishment of misconduct by financial advisers in Egan et al., 2022; CEO dismissal in Gupta et al., 2020; referrals to surgeons in Sarsons, 2017; and career benefits for co-authored research in Sarsons et al., 2021), in which case women could correctly believe that signaling is less valuable for them than for men. Then even if they have the same risk preferences as men, women would take fewer risks. Expected discrimination in signaling has been shown to deter women from sharing their true career aspirations (Bursztyn et al., 2017) and negotiating wage increases (Exley et al., 2020).

Although it may be important, I do not directly investigate risk as feelings (Loewenstein et al., 2001). Risk may induce more negative feelings for women and positive feelings for men (for example due to a lifetime of experience with being praised less for successes and punished more for failures as a woman or praised more and punished less as a man). It is not totally clear how risk as feelings would show up in this experiment. It could show up in participants' general beliefs about risk, in which case I would attribute it to beliefs. Or it could show up as excess risk avoidance or risk seeking conditional on beliefs, in which case I would attribute it to preferences.

While I am focusing on signal senders, a number of experimental studies investigate the beliefs and behaviors of signal receivers with regard to gender. They find that receivers discriminate against women when they have less information (Bohren et al., 2019) and in

3. Experiencing discrimination can lead to anger, which can lead to more risk taking (Jamieson et al., 2013). I will not consider this psychological response. Instead, I focus on forward-looking behavior of economic agents.

male-stereotyped topics (Bordalo et al., 2019); however with more information or in female-stereotyped topics, men are discriminated against. Coffman et al. (2021) find that gender per se does not appear to affect receiver beliefs, but rather average beliefs about groups. To the extent that senders are aware of these beliefs and behaviors, they will affect their behavior in the present experiment and inform its generalizability outside the lab.

1.2 Experimental Design

The experiment was conducted online using Qualtrics and the Prolific platform. The full text of the experiment can be found in Appendix 1.C. There were two waves. In the first wave, participants were signal senders. In the second wave, participants were signal receivers. participants could not participate in both waves.

1.2.1 Senders

Sender participants were randomly assigned to either the "gender hidden" or "gender revealed" group. Participants' gender was determined from their self-identification on their Prolific profile.

They were then given 7 minutes and 30 seconds to take a 10-question multiple choice quiz. The quiz questions were based on middle school and high school level standardized test mathematics questions – some were adapted from the questions in Bordalo et al. (2019). The questions were written so that participants could not easily search for the answers using a search engine (e.g. by finding questions with the same wording in an online study guide). To incentivize effort, participants were told that they would be paid \$0.10 per correct answer with the possibility of getting another payment based on their performance. In addition, they were told that their results would be shown to another participant.

After the quiz, participants were asked to state their beliefs about how many of the 10 questions they answered correctly. For this and all other belief questions, participants were

incentivized to accurately report their beliefs by offering them a payment of $0.10 if their stated belief is correct.

The senders then learned that they would be paired with a receiver. Those in the gender hidden group were told that the receiver initially knows "nothing about you or your answers," while those in the gender revealed group were told that the receiver knows "your gender and nothing else about you or your answers." The senders were asked to give their second order beliefs about how many questions their receiver would believe they answered correctly.

They then learned about the quiz signal: 3 of the 10 questions would be randomly chosen, and the number of these questions that they answered correctly would be reported to their receiver. The senders did not learn which questions were chosen or what the realization of their signal was. Senders in the gender revealed group were reminded that their receiver would see their gender in addition to the quiz signal. They then reported their contingent second order beliefs for each possible realization of the signal (i.e. how many of the remaining 7 questions the receiver would believe they answered correctly after seeing that they answered 0, 1, 2, or 3 of the 3 sample questions correctly).

Finally, the senders chose their payment. All senders were given $1.95 for completing the survey plus any bonuses for correct quiz answers and correct beliefs. In addition, the senders were offered a choice between an additional flat payment of $0.50 or a risky payment of $0.20 per question that their partner believes they answered correctly after seeing the quiz signal. The senders were told that their receiver would not see their choice of payment and the payment would be received at the same time, regardless of their choice.

The experiment was presented to 350 sender participants, of whom 11 failed attention checks. So, 339 participants completed the tasks. Table 1.1 breaks down the sample size by group and gender. Table 1.2 summarizes the basic demographics in each group.

	Gender Revealed	Gender Hidden
Female	85	88
Male	84	82

Table 1.1: Sample size by group and gender

1.2.2 Receivers

The signal receivers are not the focus of analysis. They were included only so the senders would be acting in a real signaling environment. As such, I will only briefly discuss them here.

The receivers were randomly paired with senders. Receivers whose sender was in the gender hidden group were told nothing about their sender, and those whose sender was in the gender revealed group were told their sender's gender. The receivers were shown the quiz questions and asked to guess how many their sender answered correctly. The receivers were then shown their sender's signal (the number correct out of three randomly chosen sample questions), and asked to guess how many of the remaining questions their sender answered correctly.

1.2.3 Comments

I make a number of choices in the experimental design to simplify the decisions faced by the participants with the hope that they will answer more accurately when the decisions are easy to understand. For example, I offer a simple, flat reward for answering the questions about beliefs correctly. This eschews more complicated but incentive compatible elicitation methods like quadratic scoring (Brier, 1950) or binarized scoring (Hossain and Okui, 2013). Given that the more complicated methods don't always elicit true beliefs (Danz et al., 2022), simplicity was preferred. The questions which ask for participants' beliefs could be inter-

	Gender Revealed	Gender Hidden
Age	26.8	27.2
	(7.0)	(8.4)
Employment status:		
Full-time	0.32	0.32
	(0.47)	(0.47)
Part-time	0.18	0.17
	(0.38)	(0.37)
Not in paid work	0.05	0.03
	(0.23)	(0.17)
Unemployed	0.19	0.20
	(0.39)	(0.40)
Other/unknown	0.25	0.28
	(0.44)	(0.45)
N	169	170

Table 1.2: Sample demographics by group

preted differently by different people. The payoff-maximizing choice would be to report one's modal belief, but they could also be interpreted as the mean or median. The key assumption I make is that this interpretation does not vary systematically by gender.

Similarly, I do not attempt to elicit the entire distribution of participants' beliefs. People's ability to accurately report distributions of beliefs depends strongly on the elicitation method used (Goldstein and Rothschild, 2014). While they can be accurate, I believe that the additional information would be of limited value in this case.

To limit the effect of priming or experimenter demand, I attempt to limit the number of times I mention gender to the participants so that it is brought up only when it is necessary to fully inform them (it is mentioned twice to the gender revealed group and never to the gender hidden group). Since the experiment was conducted online, I cannot use clever methods like the sound of a person's voice to signal gender as in (Bordalo et al., 2019). Benjamin et al.

(2010) find that priming gender did not have a significant effect on risk-taking, so I expect that it is not a major factor here.

To rule out time preferences, I explicitly told participants that their payment would arrive at the same time regardless of whether they chose the partner payment or the flat payment. Without this explicit statement, there might be some concern that participants believed that choosing the partner payment would mean they had to wait until their partner saw their results while the flat payment would arrive immediately.

During the experiment, senders do not learn the results of the quiz, which questions would be chosen for the signal, or the realization of their signal. This means that they face all of the types of uncertainty mentioned above: incomplete information about their own skill, noise in the signal, and uncertainty about their receiver's beliefs. This was done so that I could evaluate participants' belief about each type of uncertainty.

The structure of the signal (randomly picking a subset of questions and revealing to the receiver how many of those questions the sender answered correctly) was chosen to make the objective probabilities of each signal simple to calculate. This is used in the model in Appendix 1.B.

1.3 Results

The risky choice in this context is the sender's choice to base their payment on their partner's posterior belief about them. I confirm that women take significantly fewer risks than men in this signaling environment. Figure 1.1 summarizes this result. In this and all other figures, the black capped bars denote 95% confidence intervals.

To decompose this difference into preferences and beliefs, I perform a logistic regression of choosing the risky payment on participants' stated beliefs and a fixed effect for gender. The results of this regression are presented in Table 1.3. The coefficient on gender is insignificant indicating that there is no significant difference in risk-taking by gender conditional on

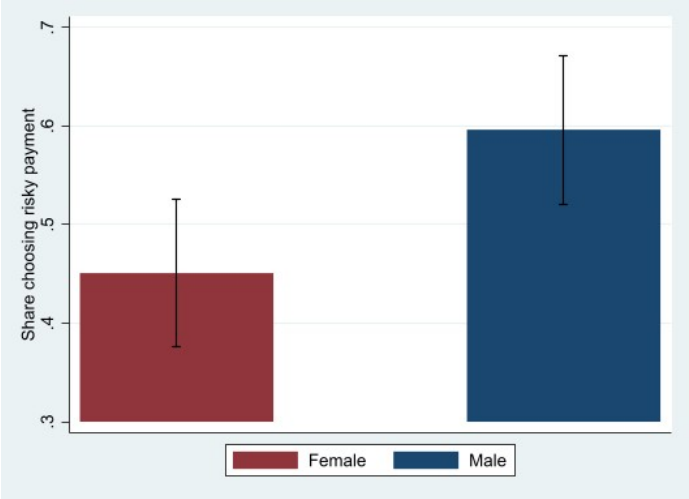


Figure 1.1: Share choosing risky payment by gender

beliefs. In other words, there is no significant difference in risk preferences as measured by this regression. I estimate simple utility functions by gender in Appendix 1.B and find the same result.

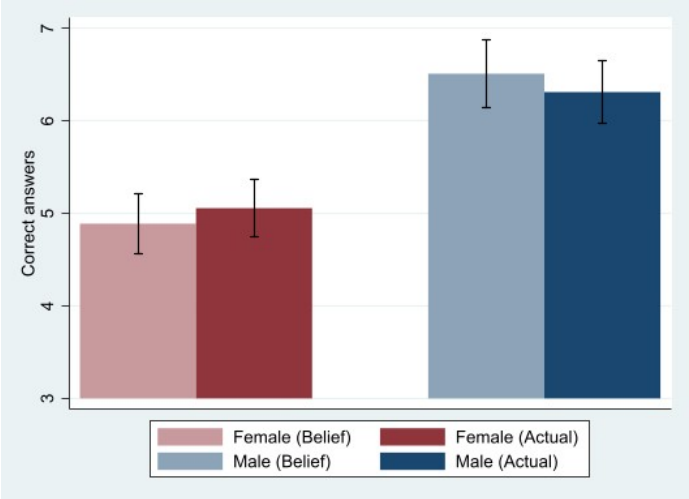


Figure 1.2: Believed and actual quiz results by gender

While preferences do not differ significantly by gender, beliefs do. To compare differences in participants' beliefs about their own ability, I compare their stated beliefs about their own score. Figure 1.2 shows the believed and actual quiz results by gender. Women are significantly less confident in their own abilities than men. However, in this case, the difference

	Risky Payment
Female	-0.19
	(0.24)
Belief about own score	0.21
	(0.06)
Contingent second order beliefs:	
Signal = 0	0.07
	(0.10)
Signal = 1	0.13
	(0.17)
Signal = 2	0.01
	(0.16)
Signal = 3	0.12
	(0.12)

Table 1.3: Risky payment logistic regression results

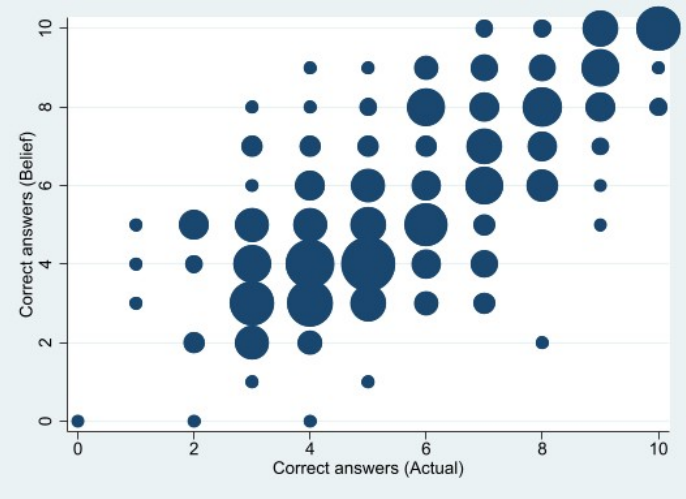


Figure 1.3: Believed and actual quiz results

in beliefs largely matches difference in performance. On average, women are slightly underconfident and men slightly overconfident relative to their true performances, but these beliefs do not differ significantly from the true results. In addition to being accurate on average, participants were reasonably accurate at the individual level. Figure 1.3 shows the scatter

plot of true score versus believed score for the whole sample; they are highly correlated ($r = 0.74$).

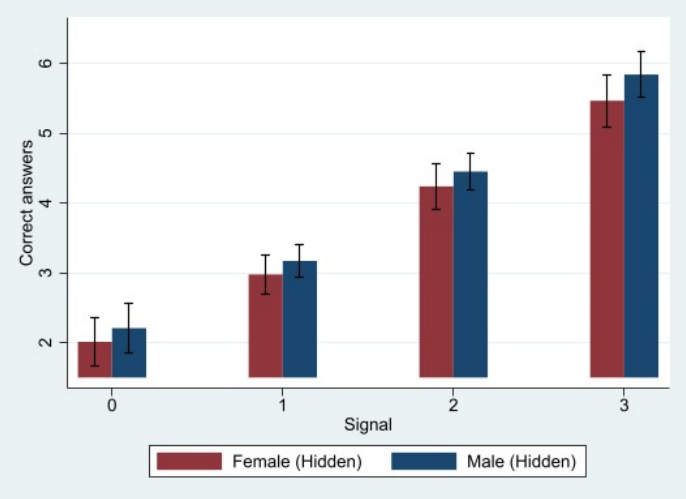


Figure 1.4: Contingent belief about partner's guess by gender and signal (Gender hidden)

To compare general beliefs about risk, I compare participants' beliefs about their receiver's beliefs following each possible signal when their gender is hidden. Figure 1.4 shows these contingent second-order beliefs. Women have lower beliefs for all possible signals, which might indicate more pessimistic beliefs about risk. However, the differences are not individually statistically significant, nor are they significantly different in a joint test ($p = 0.48$).

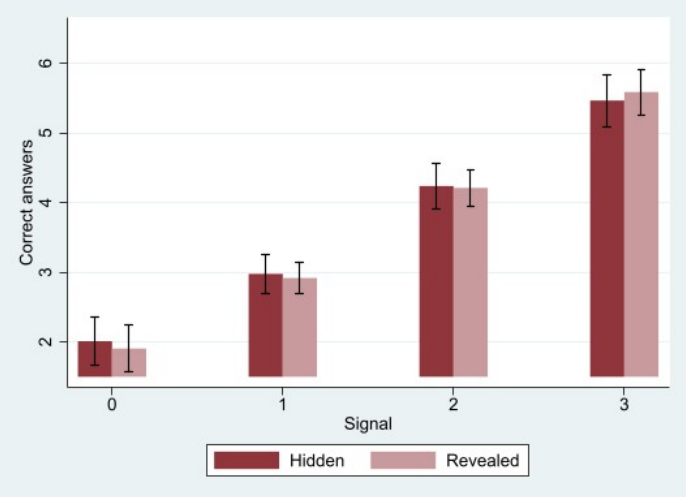


Figure 1.5: Contingent belief about partner's guess by group and signal (Females)

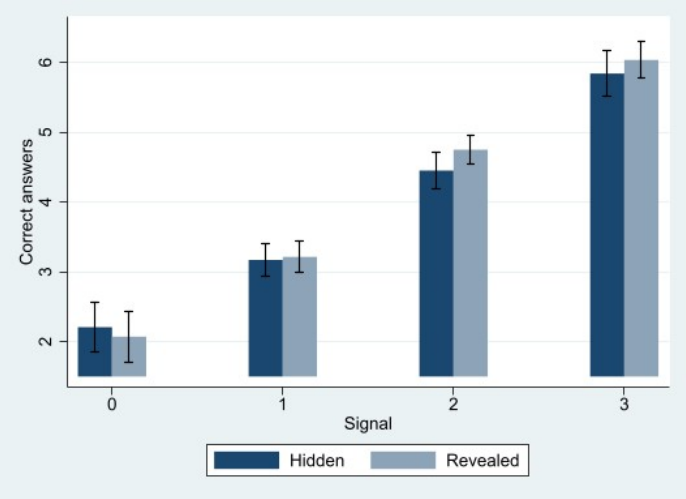


Figure 1.6: Contingent belief about partner's guess by group and signal (Males)

To identify anticipation of discrimination, I measure the difference in contingent second-order beliefs between participants whose gender is hidden and those whose gender is revealed. Figures 1.5 and 1.6 show these belief for females and males respectively. Women whose gender is revealed have lower beliefs than those whose gender is hidden for 3 of the 4 signals, and men whose gender is revealed have higher beliefs than those whose gender is hidden for 3 of the 4 signals, which might indicate some anticipated discrimination against women and towards men. However, these differences are small in magnitude and not individually statistically significant or significantly different in a joint test for men ($p = 0.96$) or women ($p = 0.35$).

Figure 1.7 shows the effect of anticipated discrimination on risk-taking. Women whose gender is revealed take slightly fewer risks and men whose gender is revealed take slightly more risks, however the difference between the hidden and revealed groups is not significant in either case.

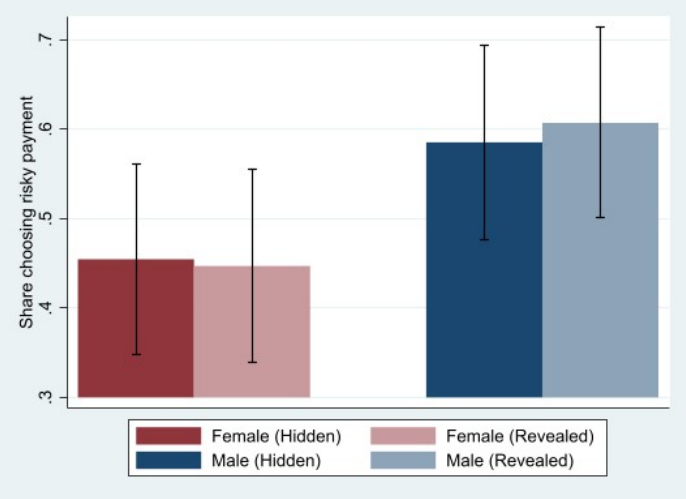


Figure 1.7: Share choosing risky payment by gender and group

1.4 Conclusion

I find that women take fewer risks than men in this signaling environment; women are significantly less likely to base their payment on their partner's belief about them than men. This difference is driven by differences in beliefs. Most importantly, women are less confident in their own abilities than men as evidenced by their significantly lower belief in their own quiz performance. There is some evidence for more pessimistic beliefs about risk among women than men in their contingent second-order beliefs and no evidence of anticipated discrimination. After accounting for beliefs, there is no significant different in risk taking between men and women, so there is no evidence of difference in risk preferences by gender.

All told, the lack of difference in risk preferences by gender is intriguing and contrary to the conventional wisdom in the economics literature. However, the results about beliefs are more mixed and apply to a narrow domain. In this case, the difference in confidence between men and women accurately reflects a true difference in scores, so these results do not shed particular light on cases where men and women are known to be of equal ability (e.g. Sarsons 2017). And, women and men do not anticipate discrimination against themselves or in their favor, so this is not informative about contexts in which discrimination is important. Future

work remains to be done to evaluate risk-taking in a variety of signaling environments.

APPENDICES

1.A Risk aversion and risk taking in signaling environments

It is often the case (as in this chapter) that choosing to send a signal is risky, either because the signal is noisy or because the senders have imperfect knowledge about themselves. One notable feature of signaling is that risk taking by senders could increase with the sender's risk aversion. Risk aversion could make the choice to send a risky signal more informative: all else equal, more risk averse senders must have more confident beliefs to choose to signal, and if those confident beliefs accurately reflect the sender's type, this is informative for the receiver. I rule this out in my experiment as the receivers do not learn whether or not their sender chose to signal (and the senders are told their receivers will not learn this). That said, this is an important fact when thinking about gender and risk taking in signaling environments: more risk aversion alone does not imply less risk taking, so the differences in risk taking between men and women are not necessarily explained even if one grants the assertion that women have more risk averse preferences than men. Here I present a simple model which illustrates this point.

A sender has type $\theta \in \{0, 1\}$ hidden from the receiver, and may choose to send a signal S which is distributed

$$Pr(S = \theta) = 1 - Pr(S = 1 - \theta) = p$$

where $p \in (1/2, 1)$. With slight abuse of notation, I will say that $S = \emptyset$ if the sender chooses not to signal. If the sender chooses to signal, they receive a utility bonus or penalty of v with CDF F. This cost is independent of θ and is known to the sender before choosing whether or not to signal. Regardless of whether or not the sender chooses to signal, they receive utility $u(w)$ where w is the receiver's posterior mean belief about θ after seeing the realization of S. The receiver has the prior belief $Pr(\theta = 1) = \pi$. The receiver does not observe v but knows its CDF F.

Consider a threshold strategy: senders of type $\theta = t$ signal if and only if $v > v_t$. Let $\mathbf{v} = (v_0, v_1)$. The receiver's posteriors would be the following.

$$Pr(\theta = 1|S = \emptyset) = \frac{F(v_1)\pi}{F(v_1)\pi + F(v_0)(1-\pi)} \equiv w(\emptyset, \mathbf{v})$$

$$Pr(\theta = 1|S = 0) = \frac{(1-F(v_1))(1-p)\pi}{(1-F(v_1))(1-p)\pi + (1-F(v_0))p(1-\pi)} \equiv w(0, \mathbf{v})$$

$$Pr(\theta = 1|S = 1) = \frac{(1-F(v_1))p\pi}{(1-F(v_1))p\pi + (1-F(v_0))(1-p)(1-\pi)} \equiv w(1, \mathbf{v})$$

If this is an equilibrium strategy, senders of type $\theta = t$ with $v = v_t$ will be indifferent between signaling and not signaling. So, the equilibrium $\mathbf{v}^* = (v_0^*, v_1^*)$ would solve the following.

$$u(w(\emptyset, \mathbf{v}^*)) = pu(w(0, \mathbf{v}^*)) + (1-p)u(w(1, \mathbf{v}^*)) + v_0^*$$

$$u(w(\emptyset, \mathbf{v}^*)) = (1-p)u(w(0, \mathbf{v}^*)) + pu(w(1, \mathbf{v}^*)) + v_1^*$$

The ex ante probability that a sender chooses to signal is $(1-F(v_0^*))(1-\pi) + (1-F(v_1^*))\pi$. The key takeaway from this example is that this probability is not always decreasing as preferences u become more risk averse. For example, consider the following parameterization:

$$\begin{aligned} \pi &= 1/2 \\ F(v) &= (v+1)/2, \ v \in [-1, 1] \\ u(w) &= \begin{cases} \alpha w + (1-\alpha)\frac{1}{2} & \text{if } w < \frac{1}{2} \\ w & \text{if } w \geq \frac{1}{2} \end{cases} \end{aligned}$$

Where $\alpha \geq 1$ is a risk aversion parameter: $\alpha = 1$ is risk neutrality, and risk aversion increases with α. Figure 1.A.1 shows the ex ante probability that a sender chooses to signal

as a function of α for two levels of p: 0.6 and 0.8.

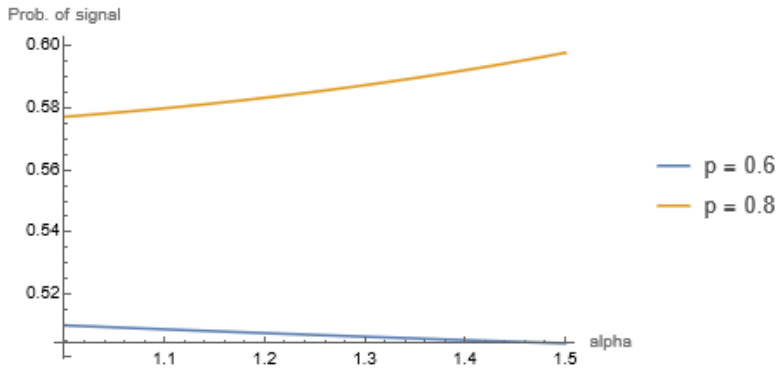


Figure 1.A.1: Ex ante probability sender chooses to signal

For the less informative signal ($p = 0.6$), the standard intuition holds: more risk averse senders (higher α) are less likely to choose to signal. However, for the more informative signal ($p = 0.8$) more risk averse senders are more likely to choose to signal.

1.B Estimating a simple model of preferences

In this appendix, I identify preference parameters by gender and show that there is no significant difference between men and women's risk preferences.

For preferences, I assume that every member of gender $g \in \{female, male\}$ has the same isoelastic utility function

$$u_g(x) = \frac{x^{1-\rho_g} - 1}{1 - \rho_g}$$

where $\rho_g > 0$ and $\rho_g \neq 1$. For beliefs, I assume that participant i is certain about the number of quiz questions they answered correctly ($confidence_i$) and about their receiver's posterior belief after signal s ($posterior_{is}$).

Then person i with gender $g(i)$ will choose the risky payment if

$$\sum_{s=0}^{3} Pr(S_i = s) u_{g(i)}(0.20 \times posterior_{is}) + \epsilon_i > u_{g(i)}(0.50)$$

where $\epsilon_i \sim N(0, \sigma_{g(i)})$ is a random utility shock, 0.20 is the dollar reward to the sender per question that the receiver believes they got correct, and 0.50 is the dollar reward for the flat payment. The probability of a given signal $Pr(S_i = s) = \text{Hypergeometric}(10, confidence_i, 3, s)$.

The likelihood of observation i is

$$\begin{cases} \Phi\left(u_i^{diff}/\sigma_{g(i)}\right) & \text{if } risky_i = 0 \\ 1 - \Phi\left(u_i^{diff}/\sigma_{g(i)}\right) & \text{if } risky_i = 1 \end{cases}$$

where $u_i^{diff} = u_{g(i)}(0.50) - \sum_{s=0}^{3} Pr(S_i = s) u_{g(i)}(0.20 \times posterior_{is})$ and $risky_i$ is an indicator for i choosing the risky payment. The maximum likelihood parameter estimates are show in Table 1.B.1. The test of $\rho_{female} = \rho_{male}$ yields a p-value of 0.64. In other words, there is no significant difference between the risk preference parameters for men and women.

ρ_{female}	0.63
	(0.28)
ρ_{male}	0.14
	(1.00)
σ_{female}	2.77
	(1.67)
σ_{male}	1.13
	(0.34)

Table 1.B.1: MLE estimation results

1.C Experiment text

The full text of the experiment begins on the next page.

Quiz Instructions

On the next page, you will be asked 10 multiple choice questions based on math skills. You will receive a bonus payment of $0.10 for each correct answer and could receive another bonus based on your results. Not answering a question will be treated the same as answering incorrectly.

You will be given 7 minutes and 30 seconds to answer all the questions. A timer will show how much time you have left. The page will automatically advance when time is done, or you may finish early by advancing to the next page. You will not be allowed to return to these questions once you advance.

You may find it useful to have paper and pencil available to write down calculations.

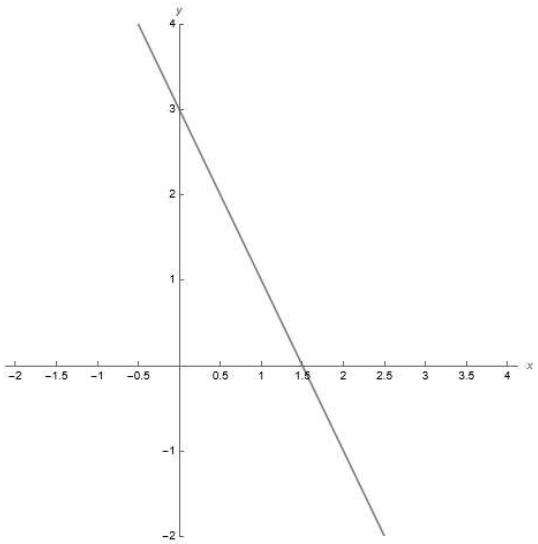
You will not learn your results immediately; instead you will be asked to guess how many questions you answered correctly. The results of some questions (the number you answer correctly) will be shown to another participant.

When you are ready to begin, go to the next page.

Quiz

Instructions: Answer the questions in the given time. The page will automatically advance when time is done, or you may finish early by advancing to the next page. You will not be allowed to return to these questions once you advance.

0 7 3 0



What is the equation for the line on the graph above?

- ○ y = 1.5 + 0.5 x
- ○ y = 2 - 3 x
- ○ y = 1.5 + 3 x
- ○ y = 3 + 2 x
- ○ y = 3 - 2 x

Some bamboo can grow 1.5 inches in an hour. How many inches would it grow in a week if it grows at that rate? (Use 24 hours in a day and 7 days in a week.)

- ○ 10.5
- ○ 36
- ○ 150
- ○ 252
- ○ 1024

Consider this list of numbers: 1, 2, 4, 8, 10. Which of the following is true of the mean and median of this list?

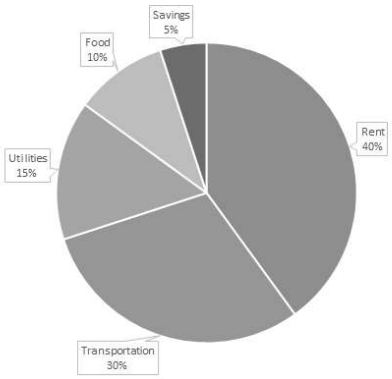
- ○ The median is larger than the mean
- ○ The mean is larger than the median
- ○ The mean and median are the same
- ○ The median cannot be determined with the information given
- ○ The mean cannot be determined with the information given

A builder buys a square piece of land for 70 thousand dollars. A second builder wants to buy a square piece of land whose sides are twice as long as the sides of the first builder's land. How many thousands of dollars will the second builder's piece of land cost if it has the same price per square mile?

- ○ 105
- ○ 140
- ○ 210
- ○ 280
- ○ 490

If I can type 30 words per minute, how many minutes will it take me to write 75 words?

- ○ 0.4
- ○ 1.5
- ○ 2.5
- ○ 15
- ○ 45



The graph above shows the distribution of spending in Tom's monthly budget. If Tom spent $800 in rent, how much did he spend on food?

- $200
- $10
- $100
- $1200
- $350

A person starts facing north. First, they turn 60 degrees clockwise. Then, they turn 150 degrees counterclockwise. In which direction are they facing after these turns?

- North
- South
- East
- West
- Cannot be determined with the information given

$2^{30} + 2^{30} + 2^{30} + 2^{30} =$

- 8^{120}
- 8^{30}
- 2^{32}
- 2^{30}
- 2^{26}

Your friend has 3 cans of orange soda, 4 cans of grape soda, and 8 cans of lemon-lime soda in her refrigerator. What are the chances that she gets an orange OR grape soda if she takes one can from her refrigerator at random?

- 7 / 15
- 1 / 8
- 3 / 4

- ○ 7 / 8
- ○ 3 / 7

A door has a lock which opens with a particular sequence of four numbers. Each number is from 1 to 36 (inclusive), and numbers could appear more than once in the sequence. Which expression below shows the number of possible sequences to open the lock?

- ○ 36^4
- ○ 35^4
- ○ 36×4
- ○ $36 \times 35 \times 34 \times 33$
- ○ 4^{36}

Results

How many of the 10 questions do you think you answered correctly?
If your guess is correct, you will receive a bonus payment of $0.10.

[ ]

To maintain the quality of our data, we need to be sure that participants are thoroughly reading the instructions. As an attention check, please select "Somewhat agree" in the text box below.

Based on these instructions select an option from the choices below.

- ○ Completely agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Completely disagree

First Guess

After completing this survey, you will be partnered with another participant who will try to guess how many questions you answered correctly. Your partner will receive a bonus payment if they guess correctly about you.

At first, your partner will see the questions you saw (but not your answers) and learn [your gender and nothing else / nothing] about you or your answers.

With only that information, how many of the 10 questions do you think your partner will guess that you answered correctly?
If your guess is correct, you will receive a bonus payment of $0.10.

[ ]

Signal Info

After your partner's first guess, we will randomly choose a sample of 3 questions, and your partner will see how many of those 3 questions that you answered correctly. Any 3 out of the 10 questions is equally likely to be the sample.

Your partner will see your results from 3 out of the 10 questions. For how many questions does your partner not know your results?

Partner Score

After seeing your results for the sample questions [and your gender], your partner will be asked to guess how many of the remaining 7 questions you answered correctly. As before, your partner will receive a bonus payment if they guess correctly about you.

Which of the following is most true of your partner's guess about you?

- It could be as low as 0
- It could be as high as 7
- It is chosen by your partner after seeing your sample question results
- All of the above

Conditional Guess

Here we ask you to consider the possible results from the sample questions and guess what your partner's guess about you would be in each case. Your partner will only see one of these results (some of the results may not be possible for you; for example, if you answered all the questions correctly, the only possible result is that your partner will see that you got 3 out of the 3 sample answers correct).
If your guess about your partner is correct, you will receive a bonus payment of $0.10.

If your partner sees that you got **0 out of the 3** sample answers correct, how many of the other 7 questions do you think they will guess that you answered correctly?

If your partner sees that you got **1 out of the 3** sample answers correct, how many of the other 7 questions do you think they will guess that you answered correctly?

If your partner sees that you got **2 out of the 3** sample answers correct, how many of the other 7 questions do you think they will guess that you answered correctly?

If your partner sees that you got **3 out of the 3** sample answers correct, how many of the other 7 questions do you think they will guess that you answered correctly?

Payment

Thank you for participating in this survey. After completing this page, you will be given $1.95 plus all the bonuses already stated.

In addition you will get one final bonus payment. To receive this payment, select your payment type below.

If you select "Flat Payment," you will receive $0.50.

If you select "Partner Guess Payment," you will receive $0.20 for each of the 7 questions that your partners guesses you answered correctly after seeing your results from the sample questions. So, for example, if your partner guessed that you answered 4 of the 7 questions correctly, this would be $0.80.

Your partner will not see which payment type you choose.

Either way, you will receive your payment within 48 hours.

Which payment type would you like?

- ○ Flat payment
- ○ Partner Guess Payment

CHAPTER 2

SENSE OF CONTROL AND WILLINGNESS TO PAY FOR NON-INSTRUMENTAL CONTROL

2.1 Introduction

For the most part, people want to have control over their outcomes. However, there are exceptions to this rule, and the strength of this desire varies (Burger, 1992). Mismatch between a person's desire for control and actual (or perceived) level of control can lead to an array of negative consequences: psychological distress, physiological stress, and poor performance (Evans et al., 1993).

When thinking about control as "the extent to which an agent can intentionally produce desired outcomes and prevent undesired ones" (Skinner, 1996), desire for the right amount of control should not be surprising to economists: the ability to determine one's outcomes or allow outcomes to be determined by someone or something else is instrumental to utility maximization. However, people also desire control because having control is intrinsically valuable.

Separating the instrumental value of control from its intrinsic value can be difficult since most control comes with both. But, it is important to distinguish between them since they play very different roles in economic decision-making.

To manage this problem, I focus on control with no instrumental value, which I call non-instrumental control. Any willingness to pay for non-instrumental control must reflect its intrinsic value. And any determinant of willingness to pay for non-instrumental control must be a determinant of the intrinsic value of control.

In particular, in this chapter I present an experiment in which I measure the effect of changes in participant's sense of control[1] on their willingness to pay for non-instrumental

1. The meaning of constructs like "sense of control" can vary by context (Skinner, 1996). Here I am using

control. Sense of control is manipulated with a writing task, and the non-instrumental control is over a guess of the outcome of a coin toss (either the participant's choice or a computer's choice).

The results are inconclusive. The sense of control manipulation did not significantly affect participants' sense of control or their willingness to pay for non-instrumental control. However, participants' willingness to pay was significantly positive with notable and predictable heterogeneity.

2.1.1 Differentiating between instrumental and non-instrumental control

I will consider a conception of control similar to Weisz (1986): "causing an intended event." From this perspective, control can be desirable for two reasons. First is its instrumental value: by having control over their outcomes, people can bring about outcomes that better satisfy their preferences. Second is its intrinsic value: people may prefer to be in control of their outcomes for its own sake – regardless of the desirability or undesirability of the outcomes themselves. This is closely related to concepts from the literature such as a "control premium" (Owens et al., 2014), an "intrinsic value of decision rights" (Bartling et al., 2014), and an "intrinsic value of choice" (Bodadilla-Suarez et al., 2017).

The instrumental value of control depends on the ends (outcomes), while the intrinsic value depends on the means (producing outcomes). This distinction appears in psychology (Skinner et al., 1988) and is important to economic analysis.

I will refer to control that has any instrumental value as "instrumental control" and control that has no instrumental value as "non-instrumental control." Of course it is possible for one thing to provide both instrumental and non-instrumental control. For example, owning one's one business affords many choices not available to a person that works for someone else; some of these choices provide instrumental value, some are purely non-instrumental, and

---

it as an individual's general belief about their control over their environment.

some contain elements of both. There is likely to be such overlap in most non-experimental choices.

This can be further complicated by the fact that people's subjective beliefs may not match objective reality. Even if control is objectively non-instrumental, people may not perceive it as such. For example they may hold irrational beliefs about their effect on outcomes (e.g. superstitions, magical thinking, or fallacious reasoning). This is discussed in the literature on illusion of control (Langer, 1975). Charness and Gneezy (2010) and Bartling et al. (2014) find that illusion of control does not play a major role their experiments, but even if people do not experience the illusion of control, they may know that control is non-instrumental and still acquiesce to superstition or magical thinking (Risen 2016) and treat it as instrumental, which further blurs the distinction between instrumental and non-instrumental control.

I will consider an environment where the distinction is very clear, but I acknowledge that this may not be particularly realistic.

2.1.2 Motivation

I focus on non-instrumental control as it allows me to cleanly identify the intrinsic value of control. As noted by Bartling et al. (2014), measuring the intrinsic value of control requires separating it from the instrumental value. By definition, non-instrumental control has no instrumental value, so any willingness to pay for non-instrumental control reflects its intrinsic value.

The intrinsic value of control may be very important, but has "been largely overlooked" in economics (Owens et al., 2014). (By contrast, the instrumental value of control is central to standard economic analysis; it is implicit in the preferences and constraints economic agents are assumed to have.[2]) There are many economic environments in which control might have intrinsic value. Bartling et al. (2014) provide a useful summary; they name corporate finance

2. To connect the language of economics to the psychology of control, constraints can be seen as the environmental affordances and behavioral competencies that an individual has (Evans et al., 1993).

and governance, entrepreneurship and self-employment, principal-agent delegation, and high-performance work systems. The intrinsic value of control in these environments could be large. For example, Hurst and Pugsley (2011) show that non-pecuniary benefits, including a greater sense of control, "play a first-order role" in the decision to become self-employed, and Moskowitz and Vissing-Jørgensen (2002) find that the value of these non-pecuniary benefits amount to 143 percent of total annual income. Not all of the non-pecuniary benefits relate to control, and much of the value of control will be instrumental, however it's possible that the intrinsic value of control plays a large role.

In addition, the intrinsic value of control may influence the market for buying versus renting. For example, Drew (2014) reports that control over one's living space is a major determinant of the decision to buy a home instead of renting. In their survey, 68% of respondents call it a major reason to buy a home (making it the fourth most important reason behind only being a good place to raise children, being a safe place, and having more space). Similarly, people may choose to buy a vehicle rather than use public transit or ride sharing to gain a sense of control. This shows up in idioms like "in the driver's seat" to mean "in control."[3] Again, some of the value of the control in these cases is instrumental, but the intrinsic value could matter, too.

The intrinsic value of control could also affect both supply and demand for in-kind versus cash transfers. Cash affords more control to the recipient, so may afford recipients value beyond the instrumental value (Simon, 2019; Skovdal et al., 2013). However giving cash gives less control to the provider than does in-kind transfers, so an intrinsic value of control may restrict supply of cash.

In particular, I am using this experiment to measure the effect of changes in sense of control on people's willingness to pay for non-instrumental control (and thus the effect of changes in sense of control on the intrinsic value of control). The effect is relevant to impor-

3. https://www.dictionary.com/browse/in-the-driver-s-seat

tant economic environments.

For one thing, many factors could change a person's sense of control. The most obvious examples are major events like job losses and unemployment (Baum et al., 1986; Goldsmith et al., 1996), diagnoses of chronic illnesses (Williams and Koocher, 1998), and natural disasters (Sneath et al., 2009), but more common occurrences matter, too: things like everyday hassles and anxiety (Ryon and Gleason, 2014) or family conflict (Newcomb and Harlow, 1986).

There could also be important interactions between changes in sense of control and the intrinsic value of control (and changes to willingness to pay for non-instrumental control). For example, sense of control is associated with a number of dimensions of socio-economic status (Mirowsky and Ross, 1990). Thus a loss in socio-economic status may lead to a loss in sense of control. If this increases the intrinsic value of control and willingness to pay for non-instrumental control, it may contribute to a poverty trap. Specifically, it could be that paying for non-instrumental control (which by definition doesn't improve outcomes) reduces people's ability to get out of poverty, and being in poverty increases people's willingness to pay for non-instrumental control. Thus, non-instrumental control might play a similar role to temptation goods in the consumption habits of the poor (Banerjee and Mullainathan, 2010).

2.1.3 Related Literature

While the psychology literature on control is vast and multifaceted (Skinner, 1996), it does not usually explicitly distinguish between the instrumental and intrinsic value of control. Even pathological attempts to gain control, for example eating disorders (Surgenor et al., 2002) or obsessive-compulsive checking (Gagné and Radomsky, 2017), might be (perceived as) instrumental to the person employing them if they have pathological preferences or beliefs.

A closely related literature on "decision seeking" (Beattie et al., 1994) or the "lure of choice" (Bown et al., 2003) finds that people seek out the ability to make choices even when those choices don't improve outcomes. While this is similar to an intrinsic value control, it is distinct: the intrinsic value of control is not about the number of choices per se. Instead, it is about control: the connection between choices and outcomes.

In economics, a small literature attempts to distinguish between the instrumental and intrinsic value of control. A big picture finding is that people are willing to choose hypothetical life choices which lead to a lower subjective well-being if those choices give more control over their lives (Benjamin et al., 2012). More specifically, people are willing to forfeit expected earnings in order for their payoff to be a function of their own choice rather than someone else's (Owens et al., 2014). In a principal-agent environment, principals are willing to forfeit expected earnings to avoid delegating to a human agent (Bartling et al., 2014), and this result holds with computer agents over both gains and losses (Bodadilla-Suarez et al., 2017).

This work leaves an open question: what factors determine the intrinsic value of control and people's willingness to pay for non-instrumental control? Motivated by results from psychology, I consider one such factor: changes in people's existing sense of control.

The effect of changes in sense of control on subsequent desire for control is somewhat mixed. Greenberger and Strasser (1986) summarize it as a desire for a balanced ratio between desired control and (perceived) possessed control. So if desire for control is fixed, a decrease in sense of control will lead to more desire to gain control (Brehm, 1966; Wicklund, 1974) and an increase in sense of control will lead to less desire to gain control. This is called reactance. On the other hand, if desire for control is allowed to vary, a decrease in sense of control could lead to so much less desire for control that people will exhibit less effort to gain control (Wortman and Brehm, 1975), and vice versa for an increase in sense of

control.[4] By comparing the distribution of willingness to pay for non-instrumental control after experimentally manipulating participants' sense of control, I can measure effects in both directions.

It is also worth noting that there are cases in which the intrinsic value of control may be negative and thus people might be willing to pay to avoid having non-instrumental control. The negative consequences of psychological distress, physiological stress, and poor performance mentioned above can come from having too much control as well as too little (Evans et al., 1993). People might not want to have control in choices between fair and unfair outcomes (Bartling and Fischbacher, 2012) or when the outcomes could be negative and thus control would lead to regret (Loomes and Sugden, 1982).

2.2 Experimental Design

The experiment was conducted online using Qualtrics and the Prolific platform. The main text of the experiment can be found in Appendix 2.A.

Participants were randomly assigned to one of three groups: high control, low control, or baseline. All three groups were asked to write a response of at least 60 words to a prompt. Those in the high control/low control groups were asked to describe an incident in their lives in which they had a great deal/very little control. This treatment and the language of the prompt follow Kraus et al. (2009). Those in the baseline group were asked to describe a building they had seen recently.

After writing the response, the participants were asked questions about their sense of control as a manipulation check. They were asked to rate their agreement on a scale of 1 to 5 with the following statements: "I can do just about anything I really set my mind to,"

4. People might also divert their desire from primary control (bringing the environment into line with their wishes) to secondary control (bringing themselves into line with environmental forces) as discussed by (Rothbaum et al., 1982). The effect on willingness to pay for (primary) non-instrumental control is the same as if they simply desired less control, so I do not distinguish between these mechanisms.

"Other people determine most of what I can and cannot do," "Whether or not I am able to get what I want is in my own hands," and "There is little I can do to change many of the important things in my life." These statements are validated as measures of sense of control in Lachman and Weaver (1998).

Participants were then offered the chance to win a bonus payment of $0.50 (in addition to the baseline payment of $1.25 for completing the survey) if they or a computer could correctly guess the outcome of a fair coin toss. They were asked which they preferred: payment based on their own choice or payment based on the computer's choice (they were informed that the computer would guess heads or tails randomly with equal probability). To ensure incentive compatibility in stating this preference, half the participants were given their preferred option at no cost (this was determined randomly after the participants completed the survey). After revealing their preference, participants stated their willingness to pay for their preferred option using a Becker-DeGroot-Marschak mechanism. The payment would come out of the bonus if they guessed correctly, so no payment was made if they did not get the bonus.

The experiment was presented to 450 people, of whom 18 failed attention checks or left the experiment without completing it. So, 432 participants completed the tasks: 147 in the low control group, 148 in the baseline group, and 137 in the high control group. Table 2.2.1 summarizes the basic demographics of each group in the sample.

2.2.1 Comments

The choice to use a coin toss for the payoff mechanism was made to ensure that the control in this experiment was transparently non-instrumental. Tossing a coin is a common randomization technique, and most people will know that the ability to choose "heads" or "tails" has no instrumental effect on the outcome. Indeed, people often use the phrase "coin toss" to refer

	Low control	Baseline	High control
Age	29.3	29.0	28.3
	(9.9)	(9.7)	(9.0)
Female	0.52	0.50	0.49
	(0.50)	(0.50)	(0.50)
Employment status:			
Full-time	0.35	0.41	0.39
	(0.48)	(0.49)	(0.49)
Part-time	0.14	0.13	0.15
	(0.35)	(0.33)	(0.35)
Not in paid work	0.05	0.07	0.03
	(0.23)	(0.25)	(0.17)
Unemployed	0.14	0.18	0.19
	(0.35)	(0.38)	(0.39)
Other/unknown	0.31	0.22	0.24
	(0.47)	(0.42)	(0.43)
N	147	148	137

Table 2.2.1: Sample demographics by group

to a random outcome with equal odds.[5] To verify this, participants were asked to report their belief that the outcome would be heads: 86% reported the correct answer (i.e. 50%),[6] and the results reported below are robust to excluding those who answered incorrectly.

The benefit of using transparently non-instrumental control is the assurance that participants' willingness to pay is in fact for non-instrumental control rather than a mistaken attempt to gain instrumental control. Related experimental work (Owens et al., 2014; Bartling et al., 2014; Bodadilla-Suarez et al., 2017) uses more complicated mechanisms, which runs the risk of participant misunderstanding and thus mis-measuring the effect.

The weakness of this approach is in external validity. As mentioned above, in most

5. https://www.dictionary.com/e/slang/coin-flip/

6. An additional 5% of participants answered within 10% of the correct answer. Since participants were allowed to provide answers using a slider tool, these participants might also hold accurate beliefs and simply mis-clicked.

environments outside the lab, control will not be so clearly divided into instrumental and non-instrumental. And, even if the control is non-instrumental, people may not treat it as such. I opted to use a clean test with a clear interpretation rather than a more realistic scenario.

To rule out decision seeking or the lure of choice, participants were asked for their guess of the coin toss regardless of whether their or the computer's choice would be used for the payoff. This, too, is likely an artificial experimental feature since the number of choices and the mapping between choices and outcomes likely go hand in hand in most non-experimental conditions.

2.3 Results

I create a sense of control index using the answers from the sense of control statements. I assign numerical values to each level of agreement from 1 indicating the least control to 5 indicating the most and average across the four statements. The mean sense of control index by group is presented in Figure 2.3.1. In this and all other figures, the black capped bars denote 95% confidence intervals.

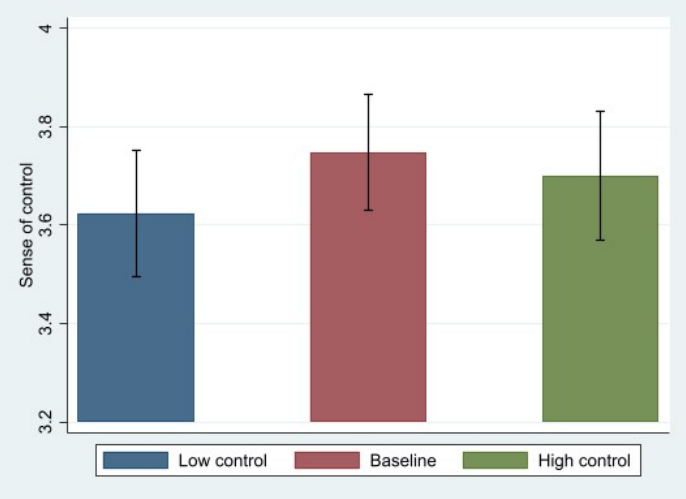


Figure 2.3.1: Mean sense of control by group

The treatment did not have a statistically significant effect on participants' sense of control. I use the same treatment as Kraus et al. (2009) and had a sufficiently large sample size to detect effects of the magnitude they find (in an effort to reduce the load on participants, I asked fewer questions about sense of control than they did, but I attempted to account for the additional variation this would bring), so I expected to find a significant difference, but this was not borne out. The point estimates are somewhat reasonable in that the low control group reported a lower mean sense of control than the high control group, but this is far from conclusive.

I assign a positive willingness to pay for non-instrumental control to participants who preferred to have their payoff based on their own choice of the coin toss and a negative willingness to pay to those who preferred to use the computer's choice. Since the treatment did not significantly manipulate stated sense of control, it is difficult to interpret treatment effects on willingness to pay for non-instrumental control, but these results are presented in Figure 2.3.2.

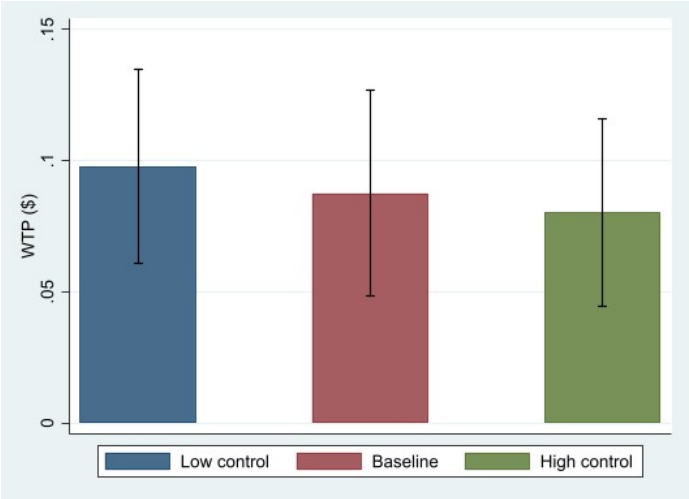


Figure 2.3.2: Mean WTP by group

The differences in mean willingness to pay between groups are not significant. That said, the point estimates are consistent with reactance being the dominant effect: participants in the low control group were willing to pay the most and those in the high control group were

willing to pay the least.

Although the effects are not significant, I estimate the effect of changes in sense of control using a two-stage least squares estimate with the first stage estimating sense of control as a function of the treatment group and the second stage estimating willingness to pay for non-instrumental control as a function of sense of control. Results are reported in Tables 2.3.1 and 2.3.2. The standard deviation of the sense of control index is 0.77 and the standard deviation of willingness to pay is 0.23, so the estimated coefficient of -0.10 on sense of control in the second stage indicates that a one standard deviation increase in sense of control causally leads to a 0.32 standard deviation decrease in willingness to pay for non-instrumental control. However, the standard errors are too large for this to be a reliable estimate.

	Sense of control
Low control	-0.12
	(0.09)
High control	-0.05
	(0.09)
Constant (Baseline)	3.75
	(0.06)
R^2	0.005

Table 2.3.1: First stage regression results

	WTP ($)
$\widehat{\text{Sense of control}}$	-0.10
	(0.23)
Constant	0.44
	(0.84)
R^2	0.005

Table 2.3.2: Second stage regression results

Given the mixed predictions from psychology about the effect of sense of control on

desire for control, one might expect treatment effect heterogeneity. If so, the insignificant mean effects may belie important effects. However, this is no evidence that this is the case. Kolmogorov-Smirnov tests of the equality of the distributions of willingness to pay show that there are no significant differences between the distributions in the low control group versus the baseline group ($p = 0.99$) or the in the high control group versus the baseline group ($p = 0.84$).

Although the main experimental effect was inconclusive, there were some notable results with the willingness to pay. First, the average willingness to pay for non-instrumental control is significantly positive for all three groups and is relatively large: participants were willing to give up an average of \$0.08 out of a \$0.50 prize in order to have their payoff based on their own choice rather than a computer's. This is especially noteworthy given how transparently non-instrumental the control was.

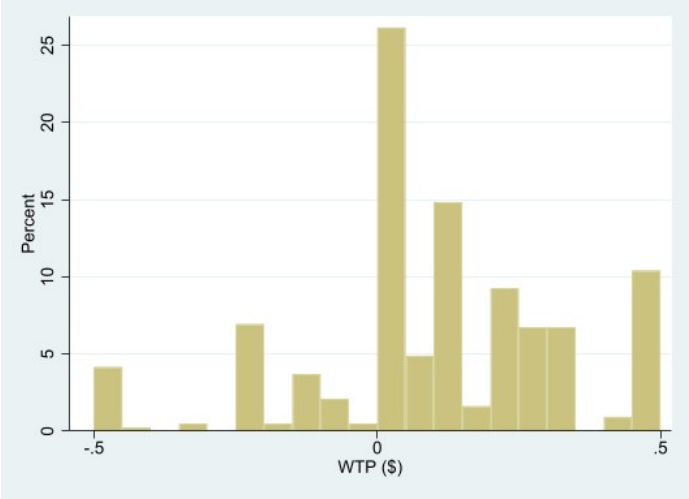


Figure 2.3.3: Distribution of WTP in full sample

Second, there was substantial heterogeneity in participant's willingness to pay. Figure 2.3.3 shows the distribution of willingness to pay in the full sample. Out of all the participants, 25% were willing to pay \$0.23 or more for non-instrumental control, and 19% were willing to pay to avoid it.

Finally, this heterogeneity is somewhat predictable. Willingness to pay and sense of con-

trol are significantly positively correlated ($r = 0.10$). This positive correlational relationship is noteworthy as it goes in the opposite direction of the (statistically insignificant) causal relationship. One possible explanation of the positive correlation is that people who value control a lot (high willingness to pay) seek it out the most, and thus feel the most in control (high sense of control).

To determine if the heterogeneity in willingness to pay is predictable along any other dimensions, I preform a regression of willingness to pay on sense of control and the demographics from Table 2.2.1. The results are presented in Table 2.3.3. Only sense of control is significantly correlated with willingness to pay.

	WTP ($)
Sense of control	0.031
	(0.014)
Age	-0.002
	(0.001)
Female	-0.031
	(0.022)
Employment status:	
Full-time	0.000
	(0.028)
Part-time	-0.003
	(0.036)
Not in paid work	-0.033
	(0.054)
Unemployed	0.041
	(0.031)
Other/unknown	-
Constant	0.030
	(0.066)
R^2	0.029

Table 2.3.3: WTP regression results

2.4 Conclusion

As the sense of control manipulation did not significantly affect participants' stated sense of control or willingness to pay for non-instrumental control, there is little to be said about the central question of this chapter: how do changes in sense of control affect willingness to pay for non-instrumental control? If the results turned out to be statistically meaningful, the magnitudes would be noteworthy. A very weak treatment of writing just 60 words led to a difference in mean sense of control of about 0.10 standard deviations, which suggests that sense of control is somewhat malleable. And a one standard deviation increase in sense of control is estimated to causally decrease willingness to pay for non-instrumental control by 0.32 standard deviations, which could be economically meaningful given the high variation in willingness to pay relative to the size of the reward in this experiment.

Despite the statistical insignificance, it remains possible that there is an effect of sense of control on willingness to pay for non-instrumental control. To more accurately measure this effect, a larger sample size might be useful with the simple writing treatment used here. One might also use a stronger treatment. For example an experiment might have participants actually undergo experiences in which they have a lot or a little control rather than simply recalling and writing about such an experience.

Finally, the heterogeneity and predictability of willingness to pay are promising but preliminary results.

APPENDICES

2.A Experiment text

The full text of the experiment begins on the next page.

Baseline Essay

Describe a building that you have seen recently. Include information about its size and appearance. Do not include identifying information like its name or address.

In order to complete the survey and get full payment, you must write at least 60 words. You may choose to advance with fewer than 60 words. If you do, the survey will end immediately, and you will receive partial payment.

Your current word count is: 0

Control Essay

Please recall a particular incident in which you had [a great deal of control / very little control]. By that, we mean a situation in which you [were / weren't] able to do what you wanted [and / or] get what you wanted. Please describe this situation in which you had [a great deal of control / very little control], including what happened, how you felt, etc.

In order to complete the survey and get full payment, you must write at least 60 words. You may choose to advance with fewer than 60 words. If you do, the survey will end immediately, and you will receive partial payment.

Your current word count is: 0

Sense of Control

Please indicate how much you agree or disagree with the following statements.

	Strongly Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
I can do just about anything I really set my mind to.	○	○	○	○	○
Other people determine most of what I can and cannot do.	○	○	○	○	○
Whether or not I am able to get what I want is in my own hands.	○	○	○	○	○
There is little I can do to change many of the important things in my life.	○	○	○	○	○

To maintain the quality of our data, we need to be sure that participants are thoroughly reading the instructions. As an attention check, please select "Somewhat agree" in the text box below.

Based on these instructions select an option from the choices below.

○ Completely agree

- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Completely disagree

Coin Toss

Thank you for participating in this survey. You will be given $1.25 as payment if you complete this page and the next page.

As a bonus payment, you could receive up to $0.50 more. To do so, either you or a computer must correctly guess the outcome of a fair coin toss. (A "fair coin" means that the outcome is either heads or tails and both are equally likely to happen.)

What is the percent chance that the outcome of the coin toss is "heads"?

0 10 20 30 40 50 60 70 80 90 100

What is your guess for the outcome of the coin toss?

- ○ Heads
- ○ Tails

If you would prefer to have your bonus payment based on your guess for the coin toss, select "My guess" below. If you would prefer to have your payment based on the computer's guess, select "Computer guess" below (the computer will randomly pick "heads" or "tails" with equal probability).

Which option do you prefer?

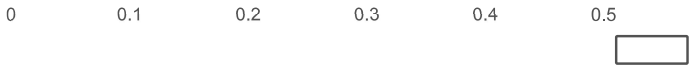
- ○ My guess
- ○ Computer guess

WTP

You said that you prefer to have your bonus payment based on [your / the computer's] guess of the coin toss. We would like to know how much you would be willing to pay to get your choice.

Half of the survey participants will get their preferred option for free, and the other half will be asked to pay for it. If you end up in the half that is asked to pay, we will randomly select a price for you. If it is more than you are willing to pay, your bonus will be based on [the computer's / your] guess of the coin toss, and you will pay nothing. If the price is less than you are willing to pay, your bonus will be based on [your / the computer's] guess of the coin toss, and the price will be deducted from the $0.50 bonus if [your / the computer's] guess is correct (you pay nothing if the guess is incorrect). So, it is in your best interest to honestly say how much you are willing to pay.

How much would you be willing to pay to have your payment be based on [your / the computer's] guess of the coin toss?



Powered by Qualtrics

CHAPTER 3

THE "BLUE WALL OF SILENCE" AND POLICE CAREERS

3.1 Introduction

Many police departments in the United States have an unwritten code "which prohibits disclosing perjury or other misconduct by fellow officers, or even testifying truthfully if the facts would implicate the conduct of a fellow officer" (Chin and Wells, 1997). This "blue wall of silence"[1] has existed for decades in cities throughout the country (Christopher, 1991; Knapp, 1970; Mollen, 1994).

Stories of officers that break the blue wall of silence suggest that it is strictly enforced (Chin and Wells, 1997; Hill, 2021; Johnson, 2006), and survey evidence shows that officers anticipate reprisal for whistle-blowing about other officers (Trautman, 2000). Despite ample anecdotal and survey evidence, little empirical work with observational data has been done on this topic.

In this chapter, I use administrative data on Chicago police officers[2] and the investigation of complaints against them to create a proxy for instances of officers testifying for or against other officers. I validate the proxy using a subset of data for which validation is possible, and test for correlations between the proxy and important career outcomes: leaving the police force, being promoted to higher pay grades, being accused in future complaints, and having those complaints sustained in investigations.

The data and proxy measure are not perfect, and the empirical relationships are correlational, so the results should be read with caution. That said, there are some potentially interesting findings: officers who do not testify against other officers are less likely to leave

1. This code has been called many things including "blue code of silence," "blue curtain," or simply "the code," however "blue wall of silence" seems to be the most commonly used name.

2. Except where explicitly noted, I will used "officers" to refer to members of the Chicago Police Department whose rank is police officer, field training officer, detective, sergeant, lieutenant, or commander.

the force and more likely to be accused in future complaints, and officers who do testify against other officers are more likely to have complaints sustained against them.

3.1.1 Motivation

The blue wall of silence facilitates police misconduct. By its nature, this effect is difficult to quantify, but the Christopher Commission (an investigation in Los Angeles following the beating of Rodney King) found that "officers' unwritten 'code of silence'" is "perhaps the greatest single barrier to the effective investigation and adjudication of complaints [against officers]" (Christopher, 1991). So, while misconduct is rife, the blue wall protects officers that commit infractions as minor as accidentally discharging pepper spray up to those who commit heinous acts of brutality and murder (Christopher, 1991; Department of Justice, 2017; Knapp, 1970; Mollen, 1994). If complaints against officers cannot be effectively investigated and adjudicated and other officers cannot be relied upon to report misconduct, there are very few ways to deter this misconduct or remove incompetent or harmful officers from the force.

Estimating the relationship between the blue wall of silence and police officer careers helps to estimate the magnitude of the problems at hand: both the tendency to cover up misconduct and the preponderance of misconduct itself. In addition, it is informative about the composition of the police force and its administration. Quantitatively measuring this relationship with observational data is a useful complement to existing anecdotal and survey evidence. While numerous stories document the fact that officers who testify against other officers face retaliation in many forms, these stories may be the most extreme examples and don't necessarily represent the typical case. Surveys confirm that officers expect this retaliation if they were to blow the whistle on misconduct in more typical environments, but it's not clear how accurate officer's beliefs are in those surveys or how the stated beliefs affect behavior.

3.1.2 Related Literature

Work in the law and law enforcement literature provides a useful background for the mechanisms which underlie the results in this chapter. First, how can officers avoid breaking the blue wall of silence? One way is literal silence: failing to report wrongdoing they have witnessed. Survey evidence indicates that this is more common with some types of wrongdoing than others; for example officers say they are much more likely to report outright theft from a crime scene than an inappropriate romantic relationship with the victim of a crime (Westmarland and Conway, 2020). Anecdotal evidence suggests that officers are willing to stay silent even about horrible brutality and torture, such as the case of Abner Louima in New York (Skolnick, 2002) in which Louima was subjected to "sadistic assault" by an officer and "proudly showed off" to other officers, but none of the officers that witnessed this immediately came forward and reported the acts. Another way to comply with the blue wall of silence is to lie to cover up wrongdoing. It is common for officers to deny witnessing wrongdoing or claim that apparent transgressions were necessary due to unusual circumstances – even in cases of extreme violence and even when testifying under oath (Chin and Wells, 1997; Koepke, 1999).

And second, what are the consequences of breaking the blue wall of silence? Stories of retaliation against whistle-blowers abound (Chin and Wells, 1997; Hill, 2021; Johnson, 2006). These officers face professional consequences like poor performance reviews and assignment to undesirable posts or duties, and other officers threaten to withhold backup if the whistle-blowing officers are in danger. They also face personal consequences: social exclusion and harassment ranging from insults to death threats. The nature of police work and training leads to strong bonds between officers: comparable to family and close friends (Kleinig, 2001). So even social exclusion – seemingly the least aggressive of these tactics – is a harsh punishment. And, importantly, it's clear from surveys that officers are aware of these retaliatory behaviors and anticipate they would happen if they break the blue wall of silence

(Trautman, 2000).

A number of other studies in economics have used the Chicago police complaints data (Ba, 2020; Ba and Rivera, 2022; Rozema and Schanzenbach, 2001). The most relevant finds that complaints against officers are a strong predictor of civil rights litigation against those officers (Rozema and Schanzenbach, ming). Importantly, they find that even allegations that are not sustained by the investigations are predictive of civil rights litigation. This may be evidence that unsustained allegations are in fact true and the officers were protected in part by the blue wall of silence.

3.2 Background

During the study time period (2007 to 2017), the CPD had a standardized, regulated system by which complaints against officers were supposed to be handled.[3] The process began when a civilian or another officer reported wrongdoing to the police department or a city government oversight agency called the Independent Police Review Authority (IPRA).[4] A single complaint could contain multiple allegations of wrongdoing and multiple accused officers. The police department was required to report any complaints it received to IPRA, so it managed all complaints.

After receiving the complaint, IPRA determined which agency would investigate the complaint based on the nature of the alleged misconduct. City ordinance dictated that IPRA had jurisdiction over some types of misconduct,[5] and all other types were investigated

3. Only the most relevant details of the complaint process are mentioned here. A fuller explanation can be found in *Independent Police Review Authority ("IPRA") Rules* (`http://www.chicagocopa.org/wp-content/uploads/2016/08/Final-IPRA-Rules-Regulations.pdf`). Direct quotes in this section are taken from that document.

4. Prior to 2007 the oversight agency was a civilian-run branch of the police department called the Office of Professional Standards. In 2007, the agency became independent of the police department and was named the Independent Police Review Authority. In 2017 IPRA was replaced by the Civilian Office of Police Accountability.

5. These are excessive force, domestic violence, coercion, and bias-based verbal abuse.

by the CPD Bureau of Internal Affairs.[6]

The investigating agency performed the investigation by reviewing the available evidence and conducting interviews with the complainant, witnesses, and accused officers. Sometimes CPD officers acted as witnesses in the investigation process. IPRA defined witness officers as officers "against whom formal action [was] not reasonably contemplated (including disciplinary, administrative, or criminal sanctions)." IPRA could issue subpoenas to officers, and the officers could be compelled to testify under oath.

When the investigation was complete, the investigating agency considered the preponderance of evidence and assigned each allegation against each accused officer to one of four categories: sustained (the allegation is supported by substantial evidence), not sustained (there is insufficient evidence to either prove or disprove the allegation), unfounded (the allegation is false or not factual), or exonerated (the incident occurred but the actions of the accused were lawful and proper). If the allegations were sustained, the investigating agency recommended discipline for the officer. The CPD reviewed the findings and recommendations and might approve or amend them with its own processes.

3.2.1 Comments

Two features of this process are especially important in interpreting the empirical results below, especially the proxy for testimony defined in Section 3.4.2. First, acting as a witness was not voluntary for officers; the investigating agencies chose which officers to question and were able to compel officers to testify. So, simply appearing as a witness does not indicate any breaking of the blue wall of silence. Second, witness officers were specifically not being considered for disciplinary action. So, for example, acting as a witness in an investigation that finds wrongdoing on the part of the accused officers is not an indication that the witness committed any misconduct.

6. This includes criminal misconduct, bribery and other forms of corruption, drug or other substance abuse, driving under the influence, as well as all operational and other violations of CPD rules.

It should be noted that the process described above was the *de jure* system, but it does not fully represent the *de facto* system. A 2017 report by the United States Department of Justice documents a litany of ways the process was interfered with or performed inadequately by CPD officers and administration (Department of Justice, 2017). The section titles alone are telling: "The City Has Put in Place Policies and Practices that Impede the Investigation of Officer Misconduct," "Investigations That CPD Does Conduct Are Neither Complete Nor Fair," "Insufficient Staffing Contributes to IPRA's Investigative Deficiencies," and "Investigations Lack Timely Resolutions, Undermining the Quality of Investigations and Credibility of the Process." Specific examples of interference include tampering with video and audio evidence and intimidating complainants and witnesses. The results of this chapter should be understood with these institutional problems in mind.

3.3 Data

The data for this chapter come from the Invisible Institute's Chicago Police Data made available through the Citizens Police Data Project.[7] The Invisible Institute is a journalism production company based in Chicago. The data were obtained by the Invisible Institute through litigation and Freedom of Information Act requests. I restrict attention the complete years in the data 2007 to 2017.

3.3.1 Officer Data

The police officer data were initially collected by CPD. I include all officers who were on the force for at least a full year in the sample period. The data cover officer race, gender, and date of birth as well as the date the officer was appointed and (if the officer has left the force as of 2017) the date the officer left the force (the data do not specify if the officer was terminated or left the force voluntarily). It also includes panel data for each officer-year

7. `https://cpdp.co/`

of the officer's salary and rank. In total, there are 16,463 unique officers represented (with 12,752 officers active in a given year on average).

Age	43.0
	(8.9)
Tenure (years)	14.3
	(8.0)
Salary ($)	80,592
	(12,814)
Female	0.23
	(0.42)
Race:	
Asian/Pacific Islander	0.03
	(0.16)
Black	0.24
	(0.43)
Hispanic	0.20
	(0.40)
Native American/Alaskan Native	0.00
	(0.06)
White	0.52
	(0.50)

Table 3.3.1: Demographics (Officer-year)

3.3.2 Complaints Data

The complaints against officers data were initially collected by IPRA. I include all complaints whose investigations were closed during the sample period and whose findings appear in the data. Each complaint has data on which officers were accused, the alleged misconduct, which officers acted as witnesses, the date on which the investigation was completed, and the findings of the investigation for each allegation and accused officer (sustained, not sustained, unfounded, or exonerated). I reduce the four possible findings to two: sustained and not

sustained (which includes not sustained, unfounded, and exonerated). Two findings are recorded in the data: the finding recommended by the investigating agency and the final finding as decided by the CPD. The recommended finding is missing from a small number of allegations (about 1%). For those allegations, I assume that the recommended finding is the same as the final finding (this is true for 99% of the allegations for which both are available). In a small number of complaints (about 0.5%) an officer is included both as an accused officer and a witness officer; this is likely an error given the institutional rules, so these officers are included only as accused officers in the analysis. In total, there are 34,165 complaints of which 3,850 have at least one allegation sustained as recommended by the investigating agency and 3,633 have at least one allegation sustained as the final decision by CPD.

3.4 Analysis

3.4.1 Descriptive Statistics

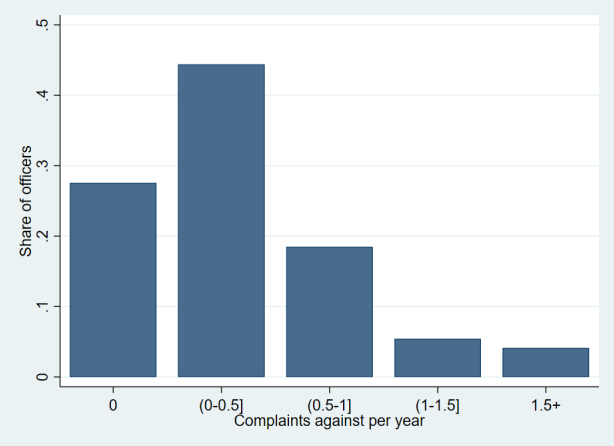


Figure 3.4.1: Complaints against officers per year

The career outcomes of interest are leaving the force, receiving a promotion (defined as an increase in pay grade), receiving complaints, and having those complaints sustained in

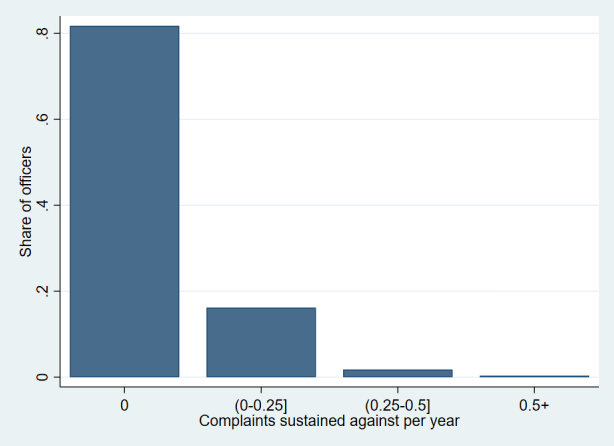


Figure 3.4.2: Complaints sustained against officers per year

the recommendation by the investigating agency (i.e. complaints in which any allegation is sustained against the officer).

Of the 16,463 officers in the sample, 3,689 leave the force during the sample period. On average, officers leaving the force are 55.3 years old (s.d. = 7.9) and have 25.1 years of tenure as an officer for the CPD (s.d. = 8.3).

And 2,241 officers receive at least one promotion (with a total of 2,471 promotions). On average, officers receiving a promotion receive a 13.9% increase in salary (s.d. = 8.4) for an average of $11,436 per year (s.d. = 6,748).

In total, 11,924 officers receive at least one complaint against them and 3,002 officers have at least one complaint sustained against them. The distribution of complaints received is shown in Figure 3.4.1, and the distribution of complaints sustained is shown in Figure 3.4.2. Because the data are an unbalanced panel, these numbers are normalized to a per officer-year basis. The most notable feature of the complaints is the right skew: 10% of officers account for 40% of all officer-complaints and 10% of officers account for 74% of all sustained officer-complaints. This is a well-known phenomenon and is true in many cities around the country (Terrill and Ingram, 2016).

The inputs of interest are the findings of investigations in which an officer acts as a

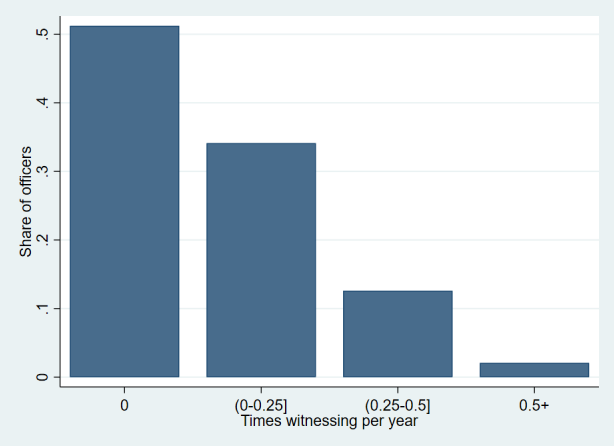


Figure 3.4.3: Officers acting as witness per year

	Has officer witness?	
	Yes	No
Any sustained	5,068	25,247
None sustained	1,514	2,336

Table 3.4.1: Investigation recommended finding and presence of officer witness

witness. The distribution of number of times witnessing is shown in Figure 3.4.3. This distribution is similarly right-skewed: 10% of officers account for 43% of officer-witnesses. Table 3.4.1 summarizes the recommended findings of investigations with and without officer witnesses; "any sustained" refers to investigations in which at least one allegation was sustained against any of the accused officers and "none sustained" refers to investigations in which none of the allegations against any of the accused officers were sustained. A higher share of investigations have any allegation sustained when an officer witness is present. Table 3.4.2 shows the distribution of investigation findings at the officer level. While most officers witness in investigations where none of the allegations are sustained, there is some variation: among officers who witnessed in 3 or more investigations, 10% had more than 66% of those investigations end with at least one allegation sustained.

There is substantial variation in witnessing by officer rank. Table 3.4.3 shows the number

Share "any sustained"	Number of times witnessed		
	1	2	3+
0	0.74	0.61	0.45
(0, 1/3)			0.18
[1/3, 2/3]		0.33	0.27
(2/3, 1)			0.08
1	0.26	0.06	0.02
N officers	3,854	2,077	2,099

Table 3.4.2: Share of complaints with any allegation sustained with officer as witness

Rank	N officers
Police officer	13,689
Field training officer	394
Detective	1,758
Sergeant	2,119
Lieutenant	460
Commander	141

Table 3.4.3: Number of officers ever holding rank

of officers who ever held each rank during the study time period (the sum is higher than the number of officers as some officers hold different ranks at different times in the study time period). Figure 3.4.4 shows the number of times witnessing by rank and Figure 3.4.5 shows the share of "any sustained" investigations by witness rank. In these and all other figures, the black capped bars denote 95% confidence intervals. Both generally increase with rank. The increase in number of times witnessing makes sense as an accused officer's superior officer is often called to witness – even if they were not present for the alleged wrongdoing. (Detectives are an outlier rank in this pattern, but this also makes sense as detectives are not supervising officers and are typically not present on the scene for the sorts of civilian interactions that lead to complaints.) The increase in share of "any sustained" investigations

could be a number of factors; possibly the investigating agency is more likely to bring in a superior officer as a witness if they anticipate sustaining allegations.

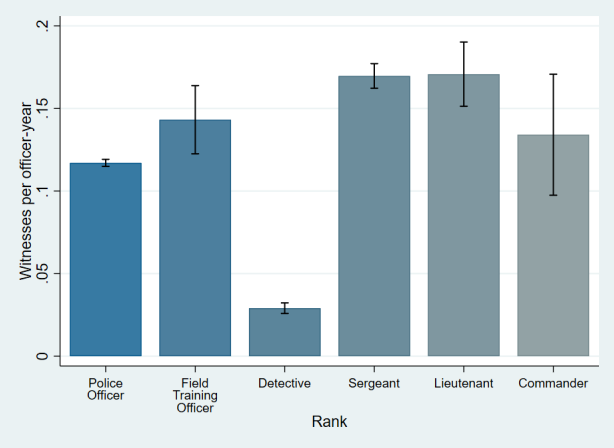


Figure 3.4.4: Number of times officer acts as witness per year by rank

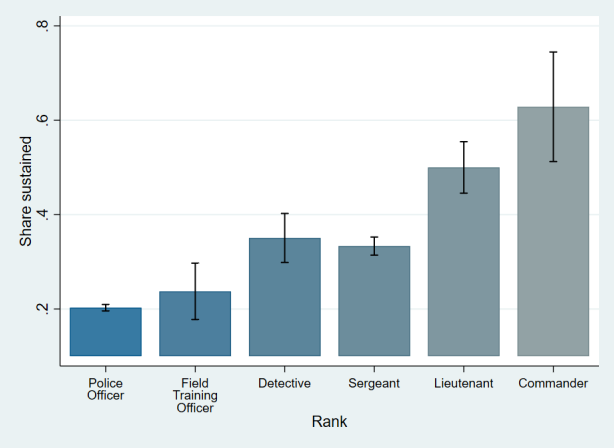


Figure 3.4.5: Share of complaints with any allegation sustained with officer as witness by rank

3.4.2 Proxy for Testimony

For most investigations, the content of witness officer testimony is not directly observable. So, I must create a proxy for whether the witness officers testified for or against the accused officers. To do so, I use the finding of the investigation as recommended by the investigating

agency: if the any of the allegations in the complaint were sustained, this is evidence that the witness officers testified against the accused officers, and if none of the allegations in the complaint were sustained, this is evidence that the witness officers testified for the accused officers. This is motivated by the fact that officers are seen as more credible than civilians (Chin and Wells, 1997). So, if the complaint is sustained, it is likely that there was some other form of evidence than the testimony of civilian complainants or witnesses. This additional evidence could be the testimony of a witness officer. I use the finding recommended by the investigating agency as this seems more likely than the final finding to hinge on the content of officer testimony.

To validate this proxy, I use the subset of cases for which the investigation summary reports are publicly available (due to e.g. statute or Freedom of Information Act request). In these reports, the content of the witness officers' testimony is not available verbatim, but is summarized in sufficient detail to determine its contribution to the investigation. I randomly sampled 10 investigations with at least one witness officer and coded the testimony as against the accused officers if the testimony provided new information or corroborated existing information that indicates wrongdoing on the part of the accused officers. If the testimony indicated no wrongdoing or was uninformative (for instance if the witness officer said they did not remember the event or any important details), I coded the testimony as for the accused officer. The results of this exercise are summarized in Table 3.4.4 (the sum of the cells in the table is greater than 10 since investigations can have multiple witness officers).

	For Accused	Against Accused
Any sustained	10	9
None sustained	21	0

Table 3.4.4: Recommended finding and content of officer testimony

For all of the sampled investigations in which none of the allegations were sustained, all of the witness officers testified for the accused officers. For the investigations in which at least one of the allegations was sustained, about half of the witness officers testified for the accused and and half against. So, let any_{it} and $none_{it}$ be the number of investigations in which officer i was a witness in year t for which the finding was any sustained or none sustained respectively. I create the proxies for testifying for and against the accused officers as follows:

$$\widehat{for}_{it} = none_{it} + \frac{any_{it}}{2}, \text{ and}$$

$$\widehat{against}_{it} = \frac{any_{it}}{2}.$$

In the analysis below, I will show regression results for both the findings (*any* and *none*) and the proxies ($\widehat{for}$ and $\widehat{against}$). Note that these are simply linear transformations of each other, so the coefficients and their standard errors can be computed easily from each other.[8]

One concern with this proxy is that officers who are witnesses in cases where allegation are sustained are themselves prone to misbehavior (since they were in the presence of others' misbehavior). If so, the proxy would capture misbehavior rather than the content of officer testimony. This is unlikely for three reasons. First, as mentioned above, witness officers are defined as being out of consideration for disciplinary action. Second, the complaint process takes a broad view of which officers are considered accused; from the review of publicly available investigation summary reports, it is clear that any officer who possibly committed wrongdoing is included as an accused officer and thus not a witness officer (for example if the complainant is uncertain which officer committed the behavior, the investigation will generally include all officers present as accused officers). Finally, as shown in Table 3.4.5,

8. Let β_i and σ_i be the coefficient and corresponding standard error on i. Then $\beta_{\widehat{for}} = \beta_{none}$, $\sigma_{\widehat{for}} = \sigma_{none}$, $\beta_{\widehat{against}} = 2\beta_{any} - \beta_{none}$, and $\sigma_{\widehat{against}} = \sqrt{4\sigma^2_{any} + \sigma^2_{none} - 4\sigma_{any,none}}$.

officers who witness when allegations are sustained receive fewer complaints than officers who witness when allegations are not sustained, indicating less misbehavior as perceived by the public.

	As witness		
	More "Any sustained"	Equal	More "None sustained"
Complaints against	3.6	5.7	6.7
	(4.3)	(5.5)	(6.6)
N officers	1,379	783	5,870

Table 3.4.5: Complaints against officers by witness investigation outcomes

3.4.3 Career Outcomes

To measure the relationship between witnessing behavior and career outcomes, I perform logistic regressions of the outcomes in each officer-year on the cumulative sums of the number of any sustained and none sustained findings in which the officer was a witness or the proxies of testifying against and for the accused officer. Since terms for both any and none sustained or terms for both proxies of testifying against and for other officers are included in the regression, the baseline of comparison is officers who have not testified at all. I include controls for year (fixed effect for each year), age (fixed effects for 5-year bins), tenure on the force (fixed effects for 5-year bins), race, gender, rank, salary (fixed effects for $5,000 dollar bins), and cumulative sum of complaints against and complaints sustained against. In the promotions regression, I use lagged rank and salary, and in the complaints regressions, I use lagged complaints against and complaints sustained against. Specifically, the regressions are of the form

$$\text{logit}(p_{it}) = \alpha + \beta X_{it} + \gamma \sum_{t'=2007}^{t} v_{it} + \delta \sum_{t'=2007}^{t} w_{it} + \epsilon_{it}$$

for officer i in year t where p are the probabilities of the outcomes of interest, X are the controls, and v and w are the witness inputs (either *any* and *none* or $\widehat{against}$ and $\widehat{for}$). For all the regressions, standard errors are clustered at the officer level. The number of officer-years varies by regression since some covariates perfectly predict outcomes, and the perfectly predicted observations are dropped. Tables 3.4.6 and 3.4.7 summarize the results.

	Leaving the force	Leaving the force	Receiving Promotion	Receiving Promotion
Cumulative sum, as witness:				
Any sustained (*any*)	-0.023		0.046	
	(0.036)		(0.041)	
None sustained (*none*)	-0.069		-0.038	
	(0.029)		(0.023)	
Cumulative sum, proxy:				
Against accused ($\widehat{against}$)		0.022		0.129
		(0.082)		(0.088)
For accused ($\widehat{for}$)		-0.069		-0.038
		(0.029)		(0.023)
Controls	X	X	X	X
N officer-years	113,750		116,414	
Pseudo-R^2	0.35		0.15	

Table 3.4.6: Logistic regression results - Personnel changes

Officers who testify in favor of other officers are less likely to leave the force than both officers who testify against other officers and officers who don't testify at all. The parameter estimate of -0.069 implies an average marginal effect of not testifying against another officer (according to the proxy) is a 0.16 percentage point decrease in the probability of leaving the force each year; since the baseline probability is 2.6%, this is a 6.1% decrease. This may reflect career benefits for maintaining the blue wall of silence. However, officers who testify against other officers show no significant difference in leaving the force relative to officers who don't testify at all.

There are no significant relationships between testimony and promotions, which is not surprising as promotions are primarily determined by test scores.[9] That said, it may be that the officers who are more likely to remain on the force after maintaining the blue wall of silence are less likely to become officers (if e.g. they stayed on the force because of camaraderie with other officers rather than a desire to advance up the ranks). If so, there could be a causal effect that does not appear in the regression as a result of selection.

	Against officer			
	Any complaint	Any complaint	Any complaint sustained	Any complaint sustained
Cumulative sum, as witness:				
Any sustained (*any*)	0.034		0.097	
	(0.018)		(0.031)	
None sustained (*none*)	0.121		0.008	
	(0.009)		(0.019)	
Cumulative sum, proxy:				
Against accused ($\widehat{against}$)		-0.053		0.186
		(0.037)		(0.068)
For accused ($\widehat{for}$)		0.121		0.008
		(0.009)		(0.019)
Controls	X	X	X	X
N officer-years	122,186		121,905	
Pseudo-R^2	0.15		0.06	

Table 3.4.7: Logistic regression results - Complaints

Interestingly, officers who testify in favor of other officers are significantly more likely to receive complaints against them (with an average marginal effect of 2.10 percentage points or 7.3%), but officers who testify against other officers are significantly more likely to have complaints sustained against them (with an average marginal effect of 0.47 percentage points or 17.3%). So, conditional on having a complaint made against them, officers who testify

9. *Chicago Police Department Hiring Plan For Sworn Titles* (`https://www.chicago.gov/content/dam/city/depts/dhr/supp_info/HRpolicies/CPD_Hiring_Plan_with_apx.pdf`)

against other officers are more likely to have the complaint sustained.

These results are consistent with reciprocity in witnessing. That is, officers may be more likely to break the blue wall of silence against officers who have broken the blue wall of silence in the past (thus increasing the rate at which complaints against them are sustained) and less likely against those who haven't. If so, officers who have maintained the blue wall of silence in the past may feel safe committing misconduct knowing that other officers will not testify against them – thus increasing the rate at which they receive complaints. If this is the case, the regression results are a biased estimate of the causal effect of testifying for or against other officers on leaving the force and receiving a promotion since I control for the total number of complaints against and number of complaints sustained against officers; if those investigation findings are endogenous, the parameter estimates will be biased. I have no direct evidence of this sort of reciprocity, so this is merely one possible explanation of the patterns in the data.

3.5 Conclusion

The effects of the blue wall of silence are inherently difficult to quantify. With this chapter, I present an initial attempt to do so by investigating the relationships between officers witnessing in complaint investigations and their career outcomes.

The most important weakness of this evidence is that it is purely correlational and thus may suffer from major endogeneity bias. There are a number of ways this could manifest. The most obvious is that officers likely have information that is unobserved in the data which relates to the consequences of testifying, and they may choose how to testify based on this information. This would likely lead to overestimating the benefits and underestimating the costs of testifying either for or against another officer. And, as mentioned above, the various inputs and outputs could feed into each other in ways that are not picked up with these estimates.

The other major weakness is that I must use a proxy for the content of testimony rather than directly observing the testimony itself. Given that I validate the proxy and find it reasonably convincing, I believe this is likely to add noise rather than be systematically incorrect. That said, the noise could be substantial. In principle, one could attempt to retrieve all of the reports of the investigations (e.g. through Freedom of Information Act requests) and code all of the officer's testimony. I leave this task to future researchers.

With these weaknesses in mind, the key findings are that officers who maintain the blue wall of silence by not testifying against other officers (according to the proxy) are less likely to leave the force and more likely to be accused in future complaints. Officers who break the blue wall of silence are more likely to have complaints sustained against them. All told, these results are consistent with career benefits from choosing not to testify against other officers and possible retaliation against officers who do testify against other officers, however this is not conclusive.

REFERENCES

Ba, B. (2020). Going the extra mile: the cost of complaint filing, accountability, and law enforcement outcomes in chicago. Working paper.

Ba, B. and R. G. Rivera (2022). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. Working paper.

Bandiera, O., N. Parekh, B. Petrongolo, and M. Rao (2022). Men are from Mars, and women too: A Bayesian meta-analysis of overconfidence experiments. *Economica 89*(S1), S38–S70.

Banerjee, A. and S. Mullainathan (2010). The shape of temptation: Implications for the economic lives of the poor. Working paper, National Bureau of Economics Research.

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

Bartling, B., E. Fehr, and H. Herz (2014). The intrinsic value of decision rights. *Econometrica 82*(6), 2005–2039.

Bartling, B. and U. Fischbacher (2012). Shifting the blame: On delegation and responsibility. *The Review of Economic Studies 79*(1), 67–87.

Baum, A., R. Fleming, and D. M. Reddy (1986). Unemployment stress: Loss of control, reactance and learned helplessness. *Social Science & Medicine 22*(5), 509–516.

Beattie, J., J. Baron, J. C. Hershey, and M. D. Spranca (1994). Psychological determinants of decision attitude. *Journal of Behavioral Decision Making 7*(2), 129–144.

Bengtsson, C., M. Persson, and P. Willenhag (2005). Gender and overconfidence. *Economics Letters 86*(2), 199–203.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). Social identity and preferences. *American Economic Review 100*(4), 1913–28.

Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones (2012). What do you think would make you happier? What do you think you would choose? *American Economic Review 102*(5), 2083–2110.

Bertrand, M. (2011). Chapter 17 - New perspectives on gender. Volume 4 of *Handbook of Labor Economics*, pp. 1543–1590. Elsevier.

Bodadilla-Suarez, S., C. Sunstein, and T. Sharot (2017). The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses. *Journal of Risk and Uncertainty 54*, 187–202.

Bohren, J. A., A. Imas, and M. Rosenberg (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review 109*(10), 3395–3436.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019). Beliefs about gender. *American Economic Review 109*(3), 739–73.

Bown, N. J., D. Read, and B. Summers (2003). The lure of choice. *Journal of Behavioral Decision Making 16*(4), 297–308.

Brehm, J. W. (1966). A theory of psychological reactance.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review 78*(1), 1 – 3.

Burger, J. (1992). *Desire for Control: Personality, Social and Clinical Perspectives*. The Springer Series in Social Clinical Psychology. Springer US.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017). 'Acting wife': Marriage market incentives and labor market investments. *American Economic Review 107*(11), 3288–3319.

Byrnes, J. P., D. C. Miller, and W. D. Schafer (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin 125*(3), 367–383.

Charness, G. and U. Gneezy (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry 48*(1), 133–146.

Chin, G. J. and S. C. Wells (1997). The blue wall of silence as evidence of bias and motive to lie: A new approach to police perjury. *University of Pittsburgh Law Review 59*, 233–299.

Christopher, W. (1991). Report of the Independent Commission on the Los Angeles Police Department. Report.

Coffman, K. B., C. L. Exley, and M. Niederle (2021). The role of beliefs in driving gender discrimination. *Management Science 67*(6), 3551–3569.

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

Cueva, C., R. E. Roberts, T. Spencer, N. Rani, M. Tempest, P. N. Tobler, J. Herbert, and A. Rustichini (2015). Cortisol and testosterone increase financial risk taking and may destabilize markets. *Scientific Reports 5*.

Danz, D., L. Vesterlund, and A. J. Wilson (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review 112*(9), 2851–83.

Davidson, D. J. and W. R. Freudenburg (1996). Gender and environmental risk concerns: A review and analysis of available research. *Environment and Behavior 28*(3), 302–339.

Department of Justice (2017). Investigation of the Chicago Police Department. Report.

Drew, R. B. (2014). Believing in homeownership: Behavioral drivers of housing tenure decisions. Working paper, Joint Center for Housing Studies Harvard University.

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

Eckel, C. C. and P. J. Grossman (2008b). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization 68*(1), 1–17.

Egan, M., G. Matvos, and A. Seru (2022). When Harry fired Sally: The double standard in punishing misconduct. *Journal of Political Economy 130*(5), 1184–1248.

Evans, G. W., D. H. Shapiro, and M. A. Lewis (1993). Specifying dysfunctional mismatches between different control dimensions. *British Journal of Psychology 84*(2), 255–273.

Exley, C. L., M. Niederle, and L. Vesterlund (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy 128*(3), 816–854.

Flynn, J., P. Slovic, and C. K. Mertz (1994). Gender, race, and perception of environmental health risks. *Risk Analysis 14*(6), 1101–1108.

Gagné, J.-P. and A. S. Radomsky (2017). Manipulating beliefs about losing control causes checking behaviour. *Journal of Obsessive-Compulsive and Related Disorders 15*, 34–42.

Gallupe, O. and S. W. Baron (2014). Morality, self-control, deterrence, and drug use: Street youths and situational action theory. *Crime & Delinquency 60*(2), 284–305.

Goldsmith, A. H., J. R. Veum, and W. Darity (1996). The psychological impact of unemployment and joblessness. *The Journal of Socio-Economics 25*(3), 333–358.

Goldstein, D. and D. Rothschild (2014). Lay understanding of probability distributions. *Judgment and Decision Making 9*(1), 1–14.

Greenberger, D. B. and S. Strasser (1986). Development and application of a model of personal control in organizations. *The Academy of Management Review 11*(1), 164–177.

Guiso, L. and M. Paiella (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association 6*(6), 1109–1150.

Gupta, V. K., S. C. Mortal, S. Silveri, M. Sun, and D. B. Turban (2020). You're fired! Gender disparities in ceo dismissal. *Journal of Management 46*(4), 560–582.

Gustafson, P. E. (1998). Gender differences in risk perception: Theoretical and methodological perspectives. *Risk Analysis 18*(6), 805–811.

Hengel, E. (2022). Publishing while female: Are women held to higher standards? Evidence from peer review. *The Economic Journal 132*(648), 2951–2991.

Hill, C. (2021). Police officers are testifying against Chauvin, but the 'blue wall of silence' still stands, former cops say. *Yahoo News*.

Hossain, T. and R. Okui (2013). The Binarized Scoring Rule. *The Review of Economic Studies 80*(3), 984–1001.

Huang, J. and D. J. Kisgen (2013). Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of Financial Economics 108*(3), 822–839.

Hurst, E. and B. W. Pugsley (2011). What do small businesses do? *Brookings Papers on Economic Activity*, 73–142.

Jamieson, J. P., K. Koslov, M. K. Nock, and W. B. Mendes (2013). Experiencing discrimination increases risk taking. *Psychological Science 24*(2), 131–139.

Johnson, R. A. (2006). Whistleblowing and the police. *Rutgers Journal of Law and Urban Policy 3*, 74–83.

Kleinig, J. (2001). The blue wall of silence: An ethical analysis. *International Journal of Applied Philosophy 15*(1), 1–23.

Knapp, W. (1970). The Knapp Commission report on police corruption. Report.

Koepke, J. E. (1999). The failure to breach the blue wall of silence: The circling of the wagons to protect police perjury. *Washburn Law Journal 39*, 211–242.

Kraus, M. W., P. K. Piff, and D. Keltner (2009). Social class, sense of control, and social explanation. *Journal of personality and social psychology 97*(6), 992.

Lachman, M. E. and S. L. Weaver (1998). The sense of control as a moderator of social class differences in health and well-being. *Journal of personality and social psychology 74*(3), 763.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology 32*(2), 311–328.

Lindquist, G. S. and J. Säve-Söderbergh (2011). "Girls will be girls", especially among boys: Risk-taking in the "daily double" on jeopardy. *Economics Letters 112*(2), 158–160.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch (2001). Risk as feelings. *Psychological bulletin 127*(2), 267.

Loomes, G. and R. Sugden (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal 92*(368), 805–824.

McEwan, P. J., S. Rogers, and A. Weerapana (2021). Grade sensitivity and the economics major at a women's college. *AEA Papers and Proceedings 111*, 102–06.

Mirowsky, J. and C. E. Ross (1990). Control or defense? depression and the sense of control over good and bad outcomes. *Journal of Health and Social Behavior 31*(1), 71–86.

Mollen, M. (1994). The City of New York Commission to Investigate Allegations of Police Corruption and the Anti-Corruption Procedures of the Police Department commission report. Report.

Moro, A. and P. Norman (2004). A general equilibrium model of statistical discrimination. *Journal of Economic Theory 114*(1), 1–30.

Moskowitz, T. J. and A. Vissing-Jørgensen (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review 92*(4), 745–778.

Nelson, J. A. (2014). The power of stereotyping and confirmation bias to overwhelm accurate assessment: the case of economics, gender, and risk aversion. *Journal of Economic Methodology 21*(3), 211–231.

Nelson, J. A. (2015). Are women really more risk-averse than men? A re-analysis of the literature using expanded methods. *Journal of Economic Surveys 29*(3), 566–585.

Newcomb, M. D. and L. L. Harlow (1986). Life events and substance use among adolescents: mediating effects of perceived loss of control and meaninglessness in life. *Journal of personality and social psychology 51*(3), 564.

Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics 122*(3), 1067–1101.

Owens, D., Z. Grossman, and R. Fackler (2014). The control premium: A preference for payoff autonomy. *American Economic Journal: Microeconomics 6*(4), 138–61.

Rask, K. and J. Tiefenthaler (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review 27*(6), 676–687.

Rothbaum, F., J. R. Weisz, and S. S. Snyder (1982). Changing the world and changing the self: A two-process model of perceived control. *Journal of personality and social psychology 42*(1), 5.

Rozema, K. and M. Schanzenbach (2001). Does discipline decrease police misconduct? Evidence from chicago civilian allegations. *International Journal of Applied Philosophy 15*(1), 1–23.

Rozema, K. and M. Schanzenbach (Forthcoming). Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Applied Economics*.

Ryon, H. S. and M. E. Gleason (2014). The role of locus of control in daily life. *Personality and Social Psychology Bulletin 40*(1), 121–131.

Sapienza, P., L. Zingales, and D. Maestripieri (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences 106*(36), 15268–15273.

Sarsons, H. (2017). Interpreting signals in the labor market: Evidence from medical referrals. Working paper.

Sarsons, H., K. Gërxhani, E. Reuben, and A. Schram (2021). Gender differences in recognition for group work. *Journal of Political Economy 129*(1), 101–147.

Schubert, R., M. Brown, M. Gysler, and H. W. Brachinger (1999). Financial decision-making: Are women really more risk-averse? *The American Economic Review 89*(2), 381–385.

Simon, C. A. (2019). The effect of cash-based interventions on gender outcomes in development and humanitarian settings. Discussion paper, UN Women.

Skinner, E. (1996). A guide to constructs of control. *Journal of personality and social psychology 71*(3), 549–70.

Skinner, E., M. Chapman, and P. Baltes (1988). Control, means-ends, and agency beliefs: A new conceptualization and its measurement during childhood. *Journal of Personality and Social Psychology 54*(1), 117–133.

Skolnick, J. (2002). Corruption and the blue code of silence. *Police Practice and Research 3*(1), 7–19.

Skovdal, M., P. Mushati, L. Robertson, S. Munyati, L. Sherr, C. Nyamukapa, and S. Gregson (2013). Social acceptability and perceived impact of a community-led cash transfer programme in zimbabwe. *BMC Public Health 13*(1), 1–13.

Slovic, P. (1999). Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. *Risk Analysis 19*(4), 689–701.

Sneath, J. Z., R. Lacey, and P. A. Kennett-Hensel (2009). Coping with a natural disaster: Losses, emotions, and impulsive and compulsive buying. *Marketing Letters 20*(1), 45–60.

Surgenor, L. J., J. Horn, E. W. Plumridge, and S. M. Hudson (2002). Anorexia nervosa and psychological control: a reexamination of selected theoretical accounts. *European Eating Disorders Review 10*(2), 85–101.

Terrill, W. and J. R. Ingram (2016). Citizen complaints against the police: An eight city examination. *Police Quarterly 19*(2), 150–179.

Trautman, N. (2000). Police code of silence facts revealed. Conference paper, Americans for Effective Law Enforcement, Inc.

Weisz, J. (1986). Understanding the development of the understanding of control. Volume 18 of *Minnesota Symposium on Child Psychology*, pp. 219–278. Academic Press.

Westmarland, L. and S. Conway (2020). Police ethics and integrity: Keeping the 'blue code'of silence. *International Journal of Police Science & Management 22*(4), 378–392.

Wicklund, R. A. (1974). *Freedom and reactance.* Lawrence Erlbaum.

Wikström, P.-O. H. (2010). Explaining crime as moral actions. In S. Hitlin and S. Vaisey (Eds.), *Handbook of the Sociology of Morality*, pp. 211–239. New York, NY: Springer New York.

Williams, J. and G. P. Koocher (1998). Addressing loss of control in chronic illness: Theory and practice. *Psychotherapy: Theory, Research, Practice, Training 35*(3), 325.

Wortman, C. B. and J. W. Brehm (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. In *Advances in experimental social psychology*, Volume 8, pp. 277–336. Elsevier.

Zethraeus, N., L. Kocoska-Maras, T. Ellingsen, B. v. Schoultz, A. L. Hirschberg, and M. Johannesson (2009). A randomized trial of the effect of estrogen and testosterone on economic behavior. *PNAS Proceedings of the National Academy of Sciences of the United States of America 106*(16), 6535–6538.