8-9-2016

Three Essays on Social Issues in Experimental Economics

Daniel Lee

ABSTRACT

THREE ESSAYS ON SOCIAL ISSUES IN EXPERIMENTAL ECONOMICS

By

DANIEL J. LEE

AUGUST, 2016

Committee Chair: Dr. Susan K. Laury

Major Department: Economics

This dissertation consists of three essays, all of which use the toolbox of experimental methods to explore behavioral issues that fall out of the concepts of human capital and public economics. The essays examine how an individual alters her behaviors in response to changes in price, information, and social pressures. Understanding these behavioral changes can help us to better explore the pathways that can then inform optimal policy design.

The first essay, *Racial Bias and the Validity of the Implicit Association Test*, examines Implicit Bias from an economic standpoint. Implicit associations and biases are carried without awareness of conscious direction. In this paper, I develop a model to study giving behaviors under conditions of implicit bias. I test this model by implementing a novel laboratory experiment—a Dictator Game with sorting to study both these giving behaviors, as well as a subject's willingness to be exposed to a giving environment. In doing so, I adapt the Implicit Association Test (IAT), commonplace in other social sciences, for use in economics experiments. I then compare IAT score to dictator giving and sorting as a necessary test of its validity. I find that the presence of sorting environments identify a reluctance to share and negatively predict giving. However, despite the IAT's ever-growing popularity, it fails to predict even simple economic behaviors such as dictator giving. These results are indicative that implicit bias fails to overcome selfish interests and thus the IAT lacks external validity.

In the second essay, *Will Girls be Girls? Risk Taking and Competition in an All-Girls School* my coauthors and I conduct an experiment that tests the effect that all-girl schooling

has on risk taking and competitive behavior. In it, we compare decisions made by students in an all-girls school to those made by students in a closely matched co-educational school. We further investigate the developmental nature of this behavior by comparing choices made by younger students (Grades 7-8) with those of older students (Grades 11-12). By focusing on the structural differences between those who select into the all-girls' school, we find that although girls educated in a single-gender environment are the most risk averse, they are also among the most competitive. These results lend support to the hypothesis that "nurture matters" in the gender differences debate.

Finally, I discuss an essay on charitable giving, entitled *The Richness of Giving: Charity Selection and Charitable Gifts in a Large Field Experiment*. It builds on previous work in the charitable giving literature by examining not only how much subjects give to charity, but also which charities subjects prefer. This choice is operationalized in an artefactual field experiment with a representative sample of respondents. These data are then used to structurally model motives for giving. The novelty of this design allows me to ask several interesting questions regarding the choices one undertakes when deciding both whether and how much to give to charity. Further, I ask these questions in the context of a standard utility framework. Given the unique set up of this experiment, I also explore how these distributional preference parameters differ by charity choice and from what we have observed in the past. I find that there is more variation within demographics and charity types than across distributions.

I close with a brief summary and personal reflection.

THREE ESSAYS ON SOCIAL ISSUES IN EXPERIMENTAL ECONOMICS

BY
DANIEL J. LEE

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy in Economics
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY
2016

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Dr. Susan K. Laury

Committee: Dr. Charles Courtemanche
Dr. John A. List
Dr. Michael K. Price

Electronic Version Approved:

Mary Beth Walker, Dean
Andrew Young School of Policy Studies
Georgia State University August, 2016

Contents

List of Tables

List of Figures

Chapter I

Racial Bias and the Validity of the Implicit Association Test

I.1 Introduction

The Center for American Progress estimates the costs of discrimination at $64 Billion *per year* or roughly 2 million annually displaced American workers (Burns, 2012). Discrimination is clearly costly. It is, almost universally, a unique and puzzling issue. And yet, though its existence is widely acknowledged, it is rarely discussed publicly. In particular, in Becker's (1957) model of taste-based discrimination, animus is not only morally reprehensible, but also damaging to both social welfare and efficiency as animus necessarily burns money. However, evidence of animus is rarely observed in either naturally occurring data or field and laboratory experiments. This is perhaps due to the nature of such experiments, which tend to focus on non-visceral or unaroused decision making (*cold-phase*) when intuition dictates that personal distaste is more likely to be expressed in *hotter-phase* decisions.

This paper speaks to a recent trend in the social social sciences–the claim that discrimination from animus stems from *implicit* biases and associations. The concept of implicit bias suggests that subtle cognitive processes govern our behavior. As a result, implicit biases are those that we carry without awareness of conscious direction (Kang, 2009). The development of the Implicit Association Test (IAT henceforth, discussed further below) has lent support to these claims by introducing a tractable measure of these implicit biases without having to rely on self-reporting mechanisms, which are known to be unreliable. The IAT is essentially several timed sorting tasks. In it, subjects match features, such as faces, to highly and lowly associated attributes, such as good or bad words. Allegedly, it is easier for an experimental subject to sort any feature with its more closely associated attributes. For instance, a picture of a chair is more closely associated with a word "furniture" than a word "food", and hence more likely to be sorted faster as such. Thus, it is through this primitive of differential timing that one reveals his or her implicit biases.

There is some common-sense validation to this argument. Frequently cited examples of these biases in decision making are men being more associated with management or white faces being more associated with pleasant words and feelings. As economists, we can think of these biases as coming through on the *hot-phase* of our decision-making process. However, to act on these biases in an IAT is costless, and can be thought of as a cheap-talk action. Furthermore, there has yet to be an in-depth economics experiment to test the validity of the IAT.

Regardless, meta-analyses seem to illustrate that these biases persist, (Bertrand et al., 2005; Greenwald et al., 2009) but should we care, and if so, to what extent? The relevant question isnt merely one of existence, but whether an individual is both willing and able to act on these biases (e.g. in the case of a giving decision). To quote Dr. James Heckman, "The authors of these [discrimination] papers focus on the question of whether society is color blind, not on the specific question of whether there is market discrimination in realized transactions" (Heckman, 1998).

Given this critique and the damaging effects of bias, we want to know whether a well-functioning market can overcome implicit bias, or if it is robust to market interaction. Unfortunately, there appears to be some evidence that implicit bias is robust. For instance, Price and Wolfers (2010) claim that, due to the split-second nature of the occupation, implicit biases can explain their findings of discriminatory behaviors in NBA referees. This behavior and similar ones suggest a role for the IAT in economic research. What we first need, therefore, is a clean experimental test to see if implicit bias can predict economic behaviors. That is, is the IAT measuring the bias it claims to, and if so, does that bias influence behavior?

In this paper, I take a necessary first step in this line of research by writing a model of giving under implicit bias. I then conduct a laboratory experiment that examines the extent to which these IAT scores co-move with pro-social (giving) behaviors. Additionally, I allow subjects to sort in and out of giving environments to better identify the biases of different sharers and how they manifest in the market. I focus on giving behaviors because of a growing body of work in the social sciences discusses the relationship between bias and giving behaviors (Triplett, 2012). Furthermore giving behaviors are both non-strategic and non-spontaneous, and therefore easily controlled by the subject.

This chapter proceeds as follows: The next section provides background on the IAT and relevant literature. Section 3 describes my model. Section 4 outlines the experiment and describes the data. Sections 5 through 7 present and discuss the results. A final section concludes.

I.2 Background

I.2.1 The IAT

Bias cannot be randomly assigned, so the question remains, how can we measure it, particularly when we may be unaware of the biases we hold? Describing implicit bias as automatic, and analogizing the mechanics of it to those of a reflex, social psychologists Greenwald, McGhee, and Schwartz first claimed to be able to test for it using their Implicit

Association Test, introduced in 1998. The test is explained in their seminal paper as follows[1]:

> An implicit association test (IAT) measures differential association of 2 target concepts with an attribute. The 2 concepts appear in a 2-choice task (e.g., flower vs. insect names), and the attribute in a 2nd task (e.g., pleasant vs. unpleasant words for an evaluation attribute). When instructions oblige highly associated categories (e.g., flower + pleasant) to share a response key, performance is faster than when less associated categories (e.g., insect + pleasant) share a key. (Greenwald et al., 1998)

Though screenshots of the IAT tasks are presented in Appendix A, this description merits further discussion. The reader will note that at its core, the IAT is essentially four (timed) sorting tasks. The first two tasks, also known as the "2-choice" tasks, are relatively simple, requiring the subject to sort either concepts or evaluation attributes. In this paper, I utilize a race (Black-White) IAT which has yet to be used in economics literature, despite the fact that black-white relations remain one of society's most divisive issues.

Here, the measure of interest is implicit racial bias. Concepts in the 2-choice task are pictures of (black and white) faces, while the attributes are are (good and bad) words[2]. To further illustrate, in the 2-choice tasks a subject may be asked to sort black faces on the left and white faces on the right. Similarly, another 2-choice task would be sorting associated attributes, in this case sorting good words on one side, and bad words on the other.

The other two stages combine these two sorting tasks in a "shared-response task". Here the IAT might say good words AND white faces on the left. In this case *either* a face or a word will show up and you sort it accordingly. Then the task flips the association to say good words AND black faces on the left. This is a key distinction because the test is not eliciting a matching or opinion from subjects. Rather, this is simply a *joint* sorting task, designed to

[1]It may be difficult to visualize the assessment from this description alone, for further understanding I recommend visiting *Project Implicit*® at http://implicit.harvard.edu.

[2]examples of good words: *Joy, Love, Peace, Wonderful, Pleasure, Glorious, Laughter, Happy*; examples of bad words: *Agony, Terrible, Horrible, Nasty, Evil, Awful, Failure, Hurt.*

measure the strength of the association between concept and attribute. While this concept of associations may seem foreign to economists, it actually finds its roots in early utilitarian philosophies, wherein people seek not pleasure itself, but rather the objects associated with those pleasures (Mill, 1869). Table I.1 shows the the progression of IAT tasks.

Table I.1: Progression of IAT Tasks

Stage	**Name**	**Description**
Stage 1	Image Stimulus Learning Trial	In this trial, the custom stimulus (either images, when present, or custom words) will be presented and paired with the response to either the 'e' or 'i' key.
Stage 2	Word Stimulus Learning Trial	Most IATs that assess preference or stereotypes use positive or negative words as the associative stimuli. In this second trial, these words are presented.
Stage 3	Paired Test Trial #1	Stage 3 pairs the associations learned in Stages 1 and 2 and randomly presents a stimulus sampled from either of those sets of stimuli.
Stage 4	Reverse Image or Word Stimulus Learning Trial	Stage 4 is identical to Stage 1, except that the associations are learned with the opposite hand.
Stage 5	Paired Test Trial #2	Stage 5 combines the associations learned in Stages 2 and 4.

Source: Meade (2009)

These IAT tasks are conducted at a computer terminal where responses are measured by keystroke (e.g. E for left, I for right). The testing experience is comparable to a human player interacting with a computer game. Empirical evidence has shown that the latent sorting time of black faces and good words in the same column is longer than it is with white faces and good words.

The standard scoring metric for the IAT is known as the *D-Score* (Greenwald et al., 2003). It is similar to Cohen's measure for effect size, d, and is calculated as the difference mean latencies within test blocks divided by the standard deviation of latencies across test blocks. For the purposes of IAT scoring only paired trials (stages 3 and 5 in this experiment) are considered *test blocks*[3]. The equation for this *D-Score* is illustrated in equation I.1 below:

$$D = \frac{\overline{x_3} - \overline{x_5}}{SD_{3\&5}} \tag{I.1}$$

Accordingly, the D-score can be either positive or negative. In the case of a Black-White IAT, a positive score indicates a positive (automatic) preference for whites, and vice-versa for a negative score[4]. A score of zero indicates little or no preference. The authors further classify and interpret D-scores using the conventional measures for effect size (Cohen, 2013), with break points at $\pm 0.15, 0.35, 0.65$ for 'slight', 'moderate', and 'strong' associations, respectively.

These authors and others have then used the IAT to make several claims. For instance, these differences in latent sorting speed (and resulting D-scores) are measurements of implicit bias (i.e. the strengths of the associations we hold) and these biases are persistent–with extant anti-black biases even among minority groups (Nosek et al., 2002)! These are interesting claims, with the benefit that if true, we can observe a personal bias which people may not know, or may be unwilling to divulge. However, the claims are also dubious. While it is understandable why the IAT and similar tests use the primitive of timing, we need something stronger and more applicable in order to draw economic conclusions. As such, I ask what is the IAT actually measuring? For instance, Norton et al. (2012) suggest not wanting to appear biased (or wanting to appear race neutral) can cause a "race-paralysis" in this sort of task.

My critique is twofold, in that questions of both internal and external validity remain unanswered. Internally, consider the case of someone who may be particularly biased, but also finds sorting tasks enjoyable. Inversely, consider a subject who is unbiased but maladroit at sorting. Do we expect this sorting ability (or lack thereof) to offset the time differential?

[3]See e.g. table I.1.

[4]My critique notwithstanding, in this paper I will continue to use the terminology of preferences so as to remain consistent with the literature.

I formalize this aspect of my critique by adapting the notation of Borghans et al. (2008):

$$T_{i,IAT} = h_i(f_i, V_i) \tag{I.2}$$

Let $T_{i,IAT}$ denote person i's performance on the IAT task. Output in this task is generated by an individuals implicit associations, f_i, as well as V_i, a vector of other determinants of task productivity, such as sorting ability.

Now consider, without loss of generality, the case of two individuals, i and j, with equivalent biases and productivity functions, yet one is better at sorting. That is $f_i = f_j$, and $V_i > V_j$, implying differential task performance $h_i(f_i, V_i) > h_j(f_i, V_j) \Rightarrow T_{i,IAT} > T_{j,IAT}$. When we allow for heterogeneity in either the bias or the production function (or both) it becomes evident that implicit bias remains unidentified.

Further, some may be quick to point out that these subjects are unmotivated. As Grether and Plott (1979) note, this lack of motivation can be a true cause for concern in the validity of psychology experiments. However, when the outcome of interest is cheap talk (as in the IAT), unmotivated subjects may still be valid. The question of interest, which remains to be answered, is whether or not this IAT cheap talk predicts bias in marketplace behaviors. That is, is the IAT mapping into economically relevant decisions—is it externally valid? If so, what are the dosage implications? That is how much more do severe levels of implicit bias map into these decisions as opposed to moderate or even slight bias?

I.2.2 Literature Review

As such, the true rub lies within the application of this test. Several studies suggest we should be interested in implicit bias by claiming that it has an effect on economic decision-making[5]. Again, Price and Wolfers (2010) argue implicit bias explains discriminatory behavior amongst NBA referees, although they do not use an IAT explicitly. Select few studies in

[5]For a more thorough review of the recent psychology and management literature regarding the IAT see Jost et al. (2009).

economics do. Of them, Lowes et al. (2015) find evidence of ethnic homophily while Reuben et al. (2014) and Rooth (2010) find predictive evidence of negative hiring conditions. The former uses an experimental labor market for women in STEM fields, and the latter uses a correspondence study with an IAT follow-up. It finds that implicitly associated stereotypes (e.g. Arabs are lazy) forecast interview callbacks in Sweden. However, none of these papers use a Race IAT, which is the standard and most common. The alleged interaction between implicit bias and labor market decisions suggests a role for further economic analysis in other areas of decision-making, such as pro-social behavior.

Thus far, the economic study of bias has primarily dealt with competitive models–those in which individuals optimize their own behavior. These models date back to Becker (1957) as well as Phelps and Arrow (1972; 1973, respectively) who developed models relating taste-based (preferential) and statistical (informational) bias, respectively. Since these two models have different policy implications it is particularly important to properly identify the channels of bias. Briefly, in Becker's model employers may experience a disutility from hiring minority workers. Consequently, these workers may have to accept lower wages or similarly increase productivity to 'compensate' employers in-kind for this bias[6]. In Arrow's model firms have limited information about potential employees and are forced to infer productivity information from primitive observables. In the following discussion, I will talk about discrimination as resulting from these biases.

In this vein of "primitive observables" our natural inclination as economists to identify the effects of bias is to plug some outcome of interest (e.g. wages, employment) into a regression with some likely covariates (e.g. sex, race), control for as many factors as possible, and interpret the results, or relegate bias to a residual. Comprehensive works by both Yinger (1998) and Altonji and Blank (1999) review this regression model of identification. The consensus is that there are markets in which discrimination both exists and is prevalent. The empirical challenges in these studies, however, are twofold. First, with these reduced-

[6]This does not necessarily imply an absence of discrimination, although with enough unbiased employers, discrimination can be competed away.

form models, we cannot identify the causal pathways for this discrimination (such as implicit cognition). Secondly, the observed outcomes may be severely biased due to missing data. Charles and Guryan (2011) further critique this regression approach by asking what is the ideal experiment the regressions are mimicking.

In this sense, a natural solution to these shortcomings is to run experiments. A popular method in this research has been fictitious tests in the form of either audit or correspondence studies. Both audit studies, which use trained testers (e.g. Gneezy et al., 2012) and correspondence studies, which use fabricated paper applications (e.g. Bertrand and Mullainathan, 2004; Hanson and Hawley, 2011; Hanson et al., 2011) provide further evidence of the existence of discrimination, though they are largely silent on the magnitude of the effect. In this paper, I help to identify the magnitude (or lack thereof) of any differential treatment. Furthermore, both audit and correspondence studies have the potential to produce spurious evidence of discrimination (Neumark, 2012), and are subject to the Heckman (1998) critique of auditor influence and inferences drawn that are based on otherwise unobservable factors.

These critiques suggest a role for other field and laboratory experiments. Evidence for bias is consistently found in the field. Furthermore, this evidence is persistent across a wide-variety of circumstances and domains, from excessive in-group cooperation amongst kibbutz members, when compared to Israeli city-dwellers (Ruffle and Sosis, 2006) to Pigouvian price discrimination amongst sports card traders (List, 2004).

However, laboratory experiments have not yet found significant consensus regarding the presence of bias, and an open question is the role of *implicit* bias. Several of these studies use the methodology of a Voluntary Contribution Mechanism (VCM henceforth) public goods game (Brown-Kruse and Hummels, 1993; Cadsby and Maynes, 1998; Solow and Kirkwood, 2002; Castillo and Petrie, 2010). An outstanding issue is that VCM games study group behavior and are not reflective of the one-on-one interactions of the audit and correspondence studies described above. Furthermore, laboratory experiments should be more reflective of the discriminatory practices that we view to be most damaging to society and welfare. In this vein we consider experiments that incorporate power asymmetries that a standard VCM game lacks, to mimic realms where bias is most present.

In response, several studies of note that have used 2-player games to measure discrimination. To study discrimination in culture, Ferraro and Cummings (2007) use the standard ultimatum game with Hispanic and Navajo subjects in New Mexico. They find significantly different behavior between the two groups. Furthermore, by eliciting subjective beliefs they claim these different behaviors are indicative of statistical discrimination. Similarly, Fershtman and Gneezy (2001) use a paired design to test for and disentangle channels of discrimination in Israeli society. In their experiment, significantly less money was passed to male Jews of Eastern origin in a trust game. However, this result was not replicated with a dictator game, indicating statistical discrimination.

Slonim and Guillen (2010) use the design of a trust game to detect gender discrimination. Further, to disentangle possible effects they include a treatment that allows for partner selection. They find (almost) no discrimination without selection but significant taste-based discrimination with selection. Finally, Eckel and Petrie (2011) use a trust game with a costly option to see your partner's picture and find both a demand for pictures, and increased first-mover earnings under pictures.

These 2-player designs allow for much cleaner identification than the group play of a VCM design, particularly when sorting or selection is used as a treatment cell[7]. The problem is this set of games still involves strategic interactions. Thus, instead of trust or ultimatum, I find a dictator game (the unique elements of which are described below) to be more appropriate to studying bias in pro-social behavior. Here, since the second player is passive, any giving is non-strategic and differences in giving can only be due to discrimination. This is discussed further in the section on experimental design below. In his review of the dictator game literature, Camerer (2003) notes that we tend to observe 10-30% of passed endowments. These rates are problematic if they are only artefactual of the lab, and could be indicative

[7]Sorting refers to opting out of playing, whereas selection refers to picking ones partner.

of experimenter demand effects or privacy concerns.

One check on the observed rates is to allow subjects to sort out of dictator giving, that is offering dictators a potentially different payoff, $\$w'$, to not play the dictator game (i.e. allocate $\$w$). There are three notable papers that address sorting, and thereby motives for giving. In Dana et al. (2006) one third of subjects opted to take a private \$9 payoff instead of playing a \$10 dictator game. Broberg et al. (2007) extend this design by eliciting a subject's willingness to pay to exit using a BDM mechanism. They find more subjects are willing to exit, and for higher prices. Finally, and serving as the inspiration for this design, Lazear et al. (2012) (LMW henceforth) examine both costly (exit) and subsidized (entrance) sorting. Using a framework of social preferences, they find that sorting not only affects *how many* people share, but also what *kinds* of people share.

Though not yet used to examine bias, the motives for giving argument and the particulars of a sorting design apply nicely to this field of inquiry. It is my intent to describe these *kinds* of people not only by their giving behaviors, but potentially by their implicit biases as well. Further, I examine how these biases affect their decisions. This naturally follows from the behavioral finding that subjects are more likely to opt out of cross-race environments necessitating a judgment of racial characteristics relevant to common stereotypes (Norton et al., 2012).

In addition to addressing the above problems, this paper contributes to the literature in several novel ways. It is the first to examine the psychological pathways of bias by using the IAT. This is important because as stated above different pathways may have different economic implications for behavior. This paper is unique in providing racial information of the recipients and allowing a sorting option with varying property rights in a dictator game. By comparing the observed rates giving and differential exits to IAT scores, this paper investigates validity of the IAT in a way the research was previously lacking. Accordingly the extent of racial bias and the external implications of the test are thereby assessed.

I.3 Model Description

Assuming the IAT actually measures bias, it should also be able to predict economic decisions reflecting that bias, such as giving and sorting behaviors. However, the directions and theory underlying these decisions have not been fully explored. To that end I formalize a model of giving under implicit bias.

LMW note that different kinds of sharers exist, and introducing a sorting environment allows us to distinguish between these types, described as follows: Willing Sharers, who prefer to share and enter into sharing environments; Reluctant Sharers, who prefer not to share but do so to comply with social pressures, norms, or mores; and Non-Sharers who simply do not share[8].

In this vein, LMW aim to detect a reluctance to share. I revisit this analysis and extend the definitions further by examining one potential pathway of this reluctance–*conditional* on bias.

For the purposes of this experiment, consider a utility maximizing individual, henceforth referred to as the dictator. The dictator is indexed by her level of bias, i, which I assume manifests as animus and perfectly correlates to the dictator's IAT D-score [9]. The D-score which is drawn from a standard normal distribution, that is $i \sim N(0,1)$[10]. In this model, the dictator may be in an economic environment that allows sorting, and may also have photographic information on her receiver. If the former, the dictator can take take up to two possible actions. First, the sorting decision, that is the decision between allocating an amount w (sorting in) or receiving an amount w' (exiting out). Conditional on sorting in, she must now make the decision of how much to give, that is how to split the endowment w between herself, x, and the recipient, y, such that $x + y = w$.

I further hypothesize that individuals also sort based on who they are sharing with, and

[8]Formal definitions for these types can be found in Appendix B.

[9]This assumption that the test is registering bias as opposed to cultural knowledge of stereotypes is consistent with accepted interpretations of the IAT (Nosek and Hansen, 2008).

[10]The actual IAT D-score is truncated at -2 and 2, but this does not affect model predictions.

this sorting also manifests itself as animus. As such, I also allow the dictator to consider the race of the recipient r. This consideration only occurs if the dictator has photographic information. Thus, the dictator has preferences over her environment D, her payoff, x, the payoff to the recipient y and the similarity of the race of the recipient r. It is these preferences that determine sorting or not sorting, and potentially the giving decision:

$$U_i = U_i(D, x, y, r) \tag{I.3}$$

where D is an indicator variable such that $D = 1$ if the environment has sorting and 0 otherwise; and r is an indicator variable such that $r = 1$ if the dictator has photographic information and is the same race as the individual, and 0 otherwise.

Within subjects, the theory of animus dictates that not only is an individual's utility greater for an equal amount given to the preferred race[11]:

$$i \geq 0.15 \Rightarrow U_i(D, \bar{x}, w - \bar{x}, 1) \geq U_i(D, \bar{x}, w - \bar{x}, 0) \tag{I.4}$$

but also that a person is willing to take a utility hit to express his or her distaste. Here, that means a willingness to sort out (even if the sort is costly) for the sole purpose of *not sharing*:

$$i \geq 0.15 \wedge (w < w') \Rightarrow U_i(1, w', 0, 0) > U_i(0, w, 0, 0) \tag{I.5}$$

This unwillingness to interact is a core concept of animus. As such, across subjects, the model of animus predictions that greater bias should have more costly sorting, in addition to less sharing across races. That is:

$$U_i(D, \bar{x}, w - \bar{x}, 0) > U_j(D, \bar{x}, w - \bar{x}, 0) \Rightarrow i < j, \forall \bar{x} < w \tag{I.6}$$

$$w < w' \Rightarrow U_i(1, w', 0, 0) > U_j(1, w', 0, 0), \forall i > j \tag{I.7}$$

[11]I have written this model as biased against people of color, but it is trivial to generalize to all racial bias.

In this experiment, I restrict my focus to the across subject design. Broadly speaking, I ask two initial empirical questions based on this model. If the answer to either of these first two questions is yes, it suggests that there is a clear pathway from the the *hot-phase* IAT task to some of the *cold-phase* decisions it has been used to explain. Absent evidence of this pathway, I ask a final question concerning meta-awareness of bias:

1. Does the IAT predict giving behavior?

2. Does the IAT predict sorting out of giving environments?

3. Do biased givers attempt to mitigate their bias with small gifts?

I.4 Experiment

I.4.1 Procedures

Given that previous lab experiments have demonstrated that these different types of individuals exist, I ask what are the IAT's implications for both laboratory and naturally occurring behavior. I use the toolbox of experimental economics to see if IAT performance is related to differential treatment of receivers and if so, to what extent. In doing so, I examine the IAT as a predictor of pro-social behavior in an experimental market. This behavior includes giving as well as sorting out of potential giving environments.

To properly ask (and answer) these questions this experiment necessarily progresses in two stages: first the dictator game (potentially with a sorting option), and second with the IAT. Upon arriving at the lab subjects are randomly split into receivers and dictators. I will now explain the two roles in turn.

In a standard dictator game, a first mover is given $10 and asked how much she would like to give to a paired (and passive) player; her choice ends the game. Thus, giving in this game is non-strategic. I begin with this standard (no information) treatment to gauge dictator giving without information on the race of the recipient.

From here, I differ from a standard game in that some treatments employ a sorting environment. Specifically, I offer some dictators an exit option as in LMW. In other words, dictators are given a chance to leave the game in such a way that the passive player never knows he or she was playing a dictator game. In doing so, I aim to disentangle social pressure as a motive for giving. This opportunity (choice) can be either costly or free. The costly option is necessarily payoff dominated by at least one dictator game choice.

Finally, these treatments are run in two types of sessions: Ones with no information (anonymous), and pictures sessions, where dictators can see who they are passing to, and use that picture as a proxy for race. For the most part, we are concerned with outcomes in the "Pictures" sessions. However, the anonymous treatments serve as an interesting comparison and are necessary for commenting on the social closeness afforded by a picture. Further, the cross between pictures sessions and sorting treatments allows us to see whether implicit bias is affecting behavior on either the extensive or intensive margin. That is, the decision *to engage* in giving as well as *how much* to give.

Table I.2: Dictators by Treatment

		Sorting	
	Baseline	Costly	Free
No Information	20	13	20
Pictures	48	68	59
Total	68	81	79

Source: Author's calculation

As such, this experiment necessitates a 2x3 design. The treatment cells are as follows: A standard (baseline) dictator game, and two dictator games with sorting: costless and costly. In costless sorting, the dictator receives the same amount in entry and exit ($10). In costly sorting, the dictator receives $9 upon exit. These dictator games are all played across both anonymous and pictured sessions. The treatment cells and number of dictators that participated in each treatment are described further in table I.2 as well as in the data section below.

After roles are assigned, the dictators are randomly paired with a receiver, and in the *Pictures* treatments shown a picture of that receiver's face. The photos serve as a proxy for race. In the *No Information* treatments dictators are not informed about their receivers. In both versions, dictators are then explained the rules of the dictator game. In all but the baseline treatment, they are asked whether or not they choose to participate. In the event that a dictator elects to not participate (takes the exit option), their receiver is not given any information about allocation task, and the dictators are given their exit fee ($9 or $10, depending on treatment). Otherwise, dictators decide how to allocate a sum of $10 between themselves and their receiver.

Meanwhile, the receivers are passive in their role. They have their pictures taken, are guaranteed a show-up fee, and asked to participate in a different task. In this case, that task is a real-money, $1x$ risk-preference elicitation (Holt and Laury, 2002), the results of which I discuss in a companion paper (Lee, *in-progress*). The receiver task is constant across treatments.

The next task in the experiment is a race IAT (as described above) on all subjects. I run this task second because an IAT can possibly influence amounts passed. However, knowing they have just participated in a dictator game should not influence IAT score, as evidence shows it is difficult to fake or otherwise manipulate (Fiedler and Bluemke, 2005). I then close by collecting demographic data in the form of a survey, and pay subjects privately. Complete subject instructions and survey questions can be found in Appendices C and D, respectively.

I.4.2 Data

These experiments were conducted during the summer and fall of 2015 at the Center for Experimental Economics at Georgia State University (ExCEN). Subjects were recruited via email using the center's recruiter. While I strove for sessions to be racially balanced, this was not possible given the makeup of the subject pool. However, I believe this to be non-

Table I.3: Dictator Summary Statistics

Variable	**Mean**	**Std. Dev.**	**N**
Male	0.399	0.491	228
Black	0.724	0.448	228
Catholic	0.092	0.29	228
Previous Experience	0.794	0.405	228
Business or Econ Major	0.268	0.444	228
Age	21.775	4.838	227
Year in School	3.149	1.07	221
GPA	3.302	0.448	189

Source: Author's calculation

problematic given the experimental design, as well as the evidence cited above on implicit attitudes and minority groups.

Table I.4: Baseline Comparison of Roles in the Experiment

Panel A: χ^2 Tests			
Covariate	**Dictator %**	**Receiver %**	**p-Value**
Male	40.08	51.54	0.014*
Black	72.24	67.84	0.306
Catholic	9.257	9.69	0.873
Previous Experience	79.29	76.21	0.430
Business or Economics Major	26.87	21.59	0.189
Panel B: *Rank-Sum* Tests			
Covariate	**Dictator Mean**	**Receiver Mean**	**p-Value**
Age	21.78	21.15	0.216
Year in School	3.15	3.10	0.835
GPA	3.30	3.26	0.563

Notes: * Significant at the 5% level

Source: Author's calculation

Overall, I ran 17 experimental sessions across the 6 treatments, with a roughly equal balance of subjects across treatment rows[12]. In total, 227 dictators (i.e. 454 subjects) participated in the experiment. Table I.3 describes the demographic breakdown of the dictators. Dictators in this experiment are (on average) 22, with a 3.3 GPA. Roughly 72% are Black and 40% are Male. Most have previous experience in economics experiments, and the modal

[12]Given my power analysis and the fact that receiving in the *No-Information* treatments is anonymous, I didn't require as many subjects.

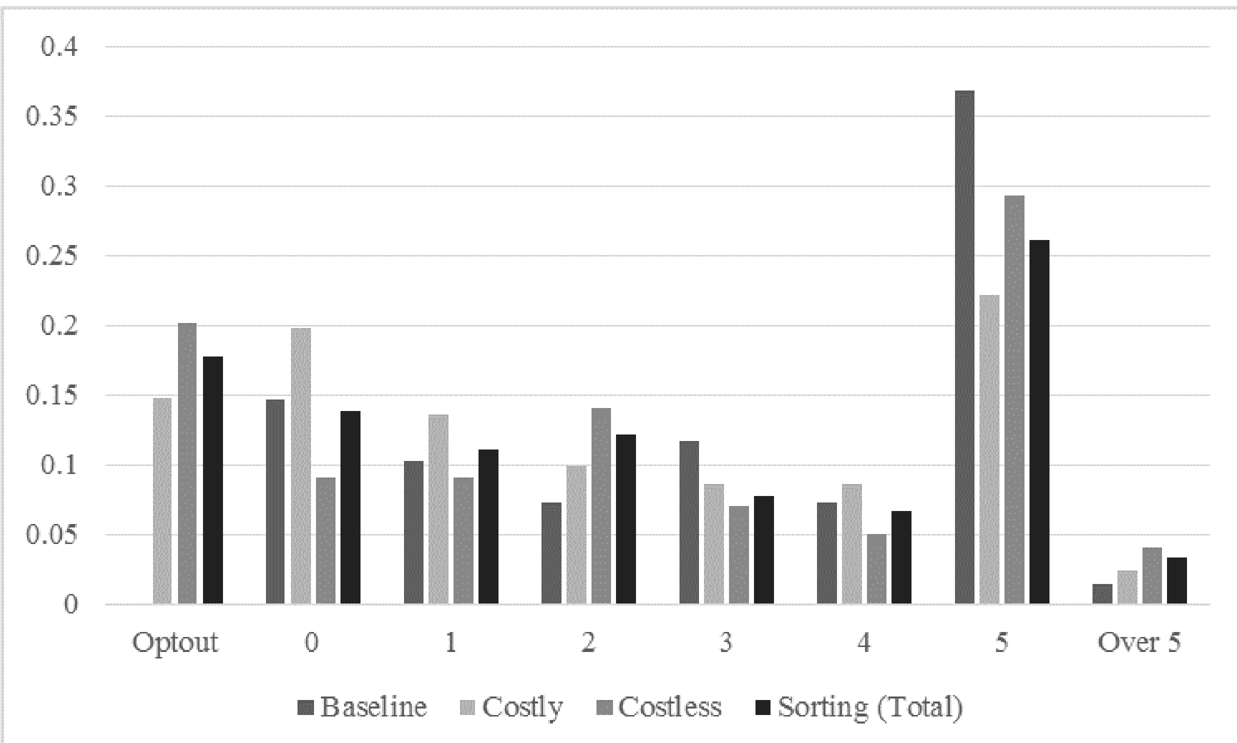
Table I.5: Experimental Summary Statistics

Variable	Mean	Std. Dev.	N
Passed to Male	0.513	0.501	228
Passed to Black	0.675	0.469	228
Amount Passed	2.692	2.238	228
Opted Out (Total)	0.186	0.389	167
Opted Out (Costly)	.148	0.357	81
Opted Out (Costless)	0.202	0.404	99
IAT D-score	0.054	0.495	225

Source: Author's calculation

year in school is senior[13].

Figure I.1: Distribution of Amounts Passed



Source: Author's illustration

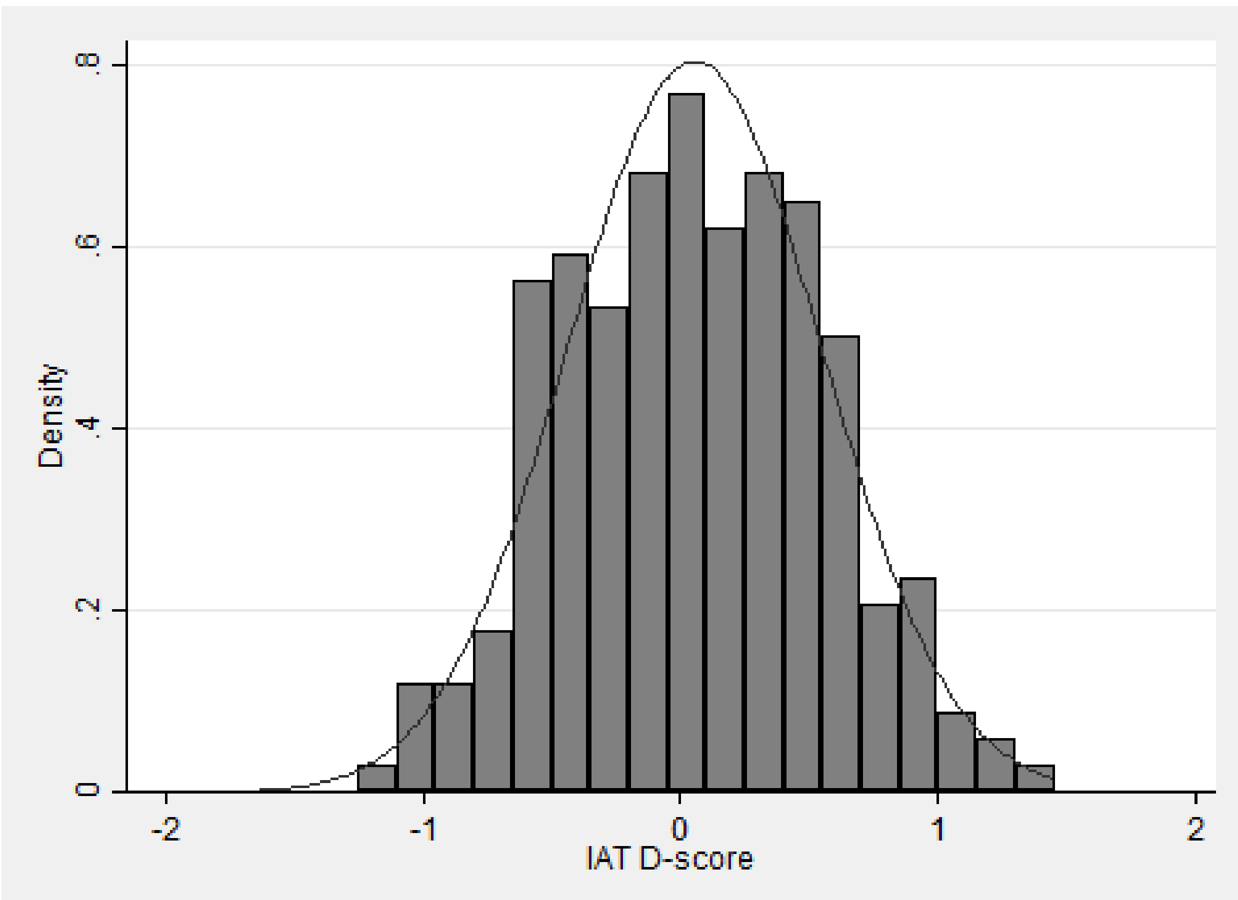
Non-parametric analyses in the form of χ^2 and Rank-Sum tests examine covariate balance between roles. For the most part, I find no significant difference across them, and conclude

[13]This is perhaps an artifact of running a summer experiment, where both former juniors and recent grads identify as seniors.

that the sample is balanced[14]. These results are reported in full in table I.4.

Finally, table I.5 provides a brief description of dictator choices and performance in the experiment. On average, 27% of the endowment was passed, and a little more than 18% of those offered an exit option opted out, with more people exiting when it is costless. Rank Sum tests show sorting significantly decreases sharing, even when sorting is costly (Sorting: $z = 2.146$, $p < 0.05$; Costly Sorting: $z = 2.370$, $p < 0.05$). These numbers are roughly similar to previous findings. Full distributions of amounts passed are illustrated further in figure I.1.

Figure I.2: Distribution of IAT D-scores



Source: Author's illustration

[14]There were significantly more males in the receiver role (117 as opposed to 91), but this is likely a byproduct of hypothesis testing across several covariates.

Regarding the IAT, the average D-score was 0.05, suggesting little to no automatic bias. I depict these scores in figure I.2 for further exploration. The scores follow a fairly normal distribution, consistent with both model assumptions and extant results across a variety of subject pools. The modal score is in the bin 0-0.15 (no automatic bias). However, there is significant implicit bias in the sample. Over 44% of dictators have an IAT D-score greater than or equal to 0.15, indicating a pro-white implicit bias. Consonant with the above evidence, this bias is present and perhaps stronger in subjects identifying as black, with a mean IAT score of 0.162.

I.5 Discrete Results

We have seen descriptively that sorting environments affect giving behaviors. However, given the empirical questions asked above, I now turn my focus to the role of the IAT in making these economic decisions. I first explore this role by simply looking at average amounts passed, broken up by the dictator's bias. Specifically, table I.6 shows the mean pass broken down by both the strength of the association, and the recipient. First of all, these differences are not significant. Secondly, if implicit bias had a one-to-one mapping into giving behaviors, we would expect passes to black subjects would get smaller as we move down the table (strengthen the bias towards whites), and the opposite pattern for whites. However, these directional patterns do not emerge, particularly in the black recipient column. Here, those who have dictators biased against them end up earning *more* on average.

Next, I take what we learned in table I.6 and discretize IAT score into the blunt question of "do I (implicitly) like or dislike my recipient?"

I express this question in equation I.8:

$$Outcome_i = \alpha_0 + \beta_1(LikeReceiver_j) + \beta_2(DislikeReceiver_j) \tag{I.8}$$

Here, I regress an outcome variable on two variables *Like* and *Dislike*. The outcome takes

Table I.6: Average Amount Passed by IAT score and Race of Receiver

	Passed to:		
Strength of Implicit Bias	**Black**	**White**	**Anonymous**
Strong for Blacks	2.07	1.67	2.5
Moderate for Blacks	2.28	2.33	2.75
Slight for Blacks	3.43	1.67	2.83
Little to None	2.18	2.55	2.33
Slight for Whites	3.13	3.5	3.33
Moderate for Whites	2.47	2.33	2.54
Strong for Whites	2.68	3.33	2.4

Source: Author's calculation

the form of either a continuous variable representing the percent of endowment shared, or a binary variable indicating whether a dictator took an exit option. The two variables *Like* and *Dislike* are essentially binary interaction terms defined formally as follows in equation I.9:

$$
\begin{aligned}
Like &= \begin{cases} 1 & \text{when } IAT \geq 0.15 \text{ and Receiver is White} \\ 1 & \text{when } IAT \leq -0.15 \text{ and Receiver is Black} \\ 0 & \text{otherwise} \end{cases} \\
Dislike &= \begin{cases} 1 & \text{when } IAT \leq -0.15 \text{ and Receiver is White} \\ 1 & \text{when } IAT \geq 0.15 \text{ and Receiver is Black} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{I.9}
$$

That is to "like" your receiver means to either hold a pro-white bias and pass to a white receiver, or hold a pro-black bias and pass to a black receiver. I later decompose the variable into these two components (pro-white, white receiver and pro-black, black receiver). Similarly, to "dislike" means to have the one of same IAT scores as above, but with the race of your receiver flipped. Accordingly, the intercept term, α, represents those dictators who hold little to no implicit bias ($-0.15 < IAT < 0.15$).

Results from these discrete estimations are presented in table I.7. In the first column we

Table I.7: Discrete IAT Estimations

VARIABLES	OLS–Percent Shared (1)	(2)	Probit–Opted Out (3)	(4)
Like Receiver	0.0696		0.0783	
	(0.0469)		(0.336)	
Pro-White, White Receiver		0.0529		-0.578
		(0.0593)		(0.568)
Pro-Black, Black Receiver		0.0774		0.292
		(0.0507)		(0.358)
Dislike Receiver	0.0291		0.166	
	(0.0492)		(0.322)	
Pro-White, Black Receiver		0.0423		0.0955
		(0.0498)		(0.337)
Pro-Black, White Receiver		-0.0289		0.456
		(0.101)		(0.504)
Constant	0.229***	0.229***	-0.887***	-0.887***
	(0.0391)	(0.0391)	(0.257)	(0.257)
Observations	172	172	126	126

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

find that unbiased givers share about 23% of their endowment, and being biased against (or in favor of) your receiver does not significantly alter this giving pattern. Furthermore, both directions of bias remain insignificant when decomposing the *Like* and *Dislike* variables into their respective components in column 2.

Similarly, table I.7, columns 3 and 4 look at how bias influences the probability of opting out. In both the blunt (column 3) and decomposed (column 4) measures, neither liking nor disliking one's receiver has any significant impact on giving.

These results indicate that bias does not affect the decision giving *on average*. However, a relevant question is do dictators who are biased against black (white) receivers behave differently than the average dictator with a black (white) receiver. To answer this question, I conduct an exercise similar to the one outlined in equation I.8, but restrict the sample based on race of receiver and only regress on the *Dislike* variable.

Table I.8: Discrete Estimations, Conditional on Race of Receiver

	Black Receiver		White Receiver	
VARIABLES	**(OLS)**	**(Probit)**	**(OLS)**	**(Probit)**
Dislike Receiver	-0.00310	-0.132	-0.0727	0.952*
	(0.0407)	(0.290)	(0.100)	(0.576)
Constant	0.274***	-0.659***	0.273***	-1.383***
	(0.0267)	(0.191)	(0.0381)	(0.374)
Observations	127	93	45	33

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

Even when we isolate the sample by race of receiver, biased dictators are not behaving in ways that are not statistically different than the average dictator, nor is this difference *economically* significant. While the above results are indicative that implicit bias fails to overcome selfish concerns, they have mostly examined the effect of IAT score on economic behaviors. Another way of looking at the question is to treat the data as observational, and ask (with some abuse of notation) what is the treatment effect of being paired with someone you hold a bias against?

To answer this question I exploit the random assignment of roles and partners and implement propensity score matching. Here, I treat each person as having a particular bias strength and direction, ranging from strongly pro-black to strongly pro-white (e.g. see table I.6). I match on the strength and direction of this bias as well as covariates describing the dictator's age, race, and sex, and specify the treatment as passing to someone you hold a bias against. That is, passing to a black person if you hold a pro-white bias and vice-versa. I find no significant treatment effect of passing to someone you are biased against (ATT=0.14, *p=0.625*).

Result 1 *Existence of bias towards receiver does not predict dictator giving*

I.6 Continuous Results

However, we measure IAT score as a continuous variable, and are able to comment not only on the *existence* of implicit bias, but also the *strength* of that bias. As such, one would think that more severe biases would exert more influence on the giving and sorting decisions. To address this dosage question, I standardize the IAT score and outline the following reduced form empirical specification:

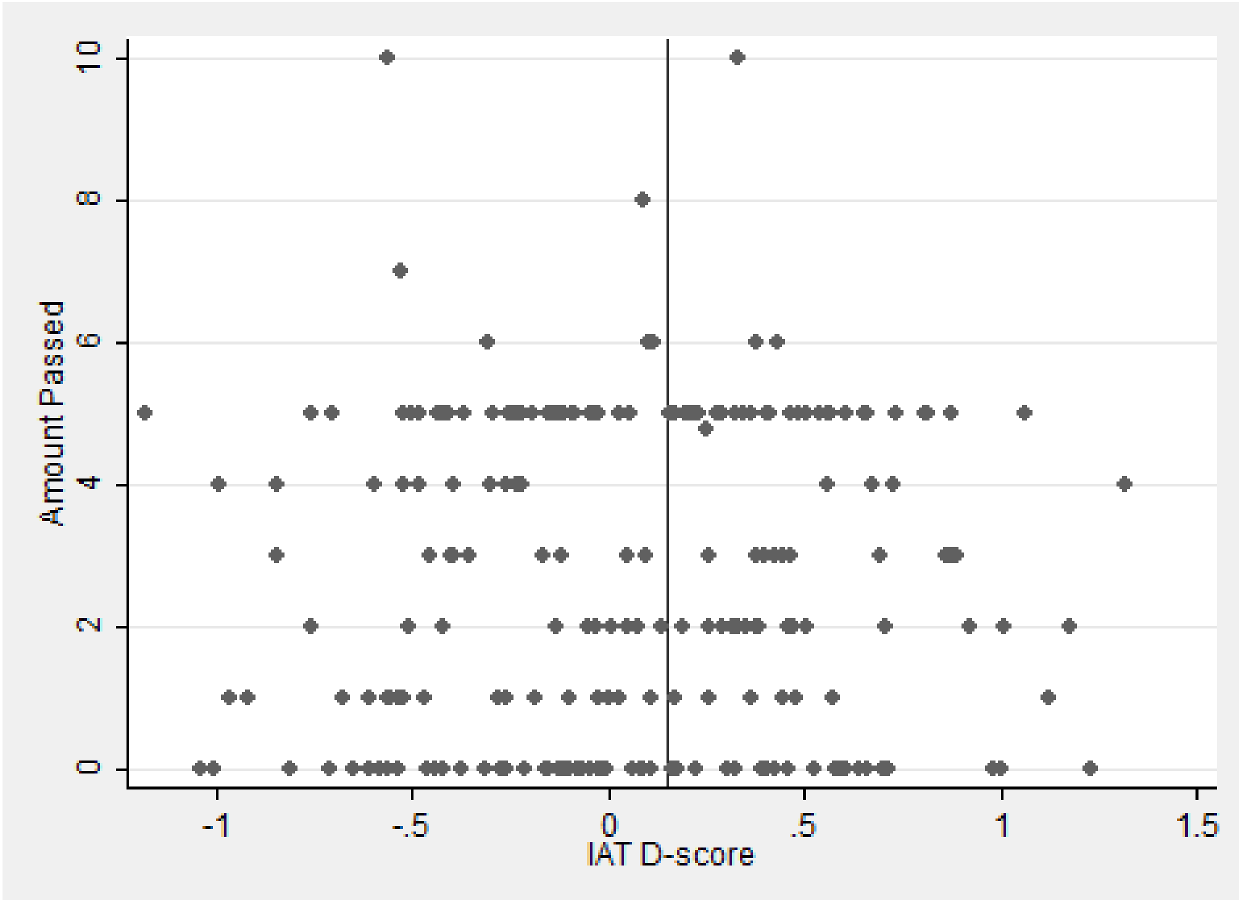
$$Outcome_i = \beta_0 + \beta_1 IAT_i + \beta_2 (IAT_i * Race_j) + \beta' X + \varepsilon_i \tag{I.10}$$

This standardization allows me to interpret coefficients as the effect of a one standard deviation increase in IAT score. In this specification I again regress an outcome variable on two variables of interest: that dictator's IAT score and an interaction term of dictator's IAT score with the race of her recipient, as well as a vector of demographic controls for both dictators and receivers. The interaction term allows us to examine this giving conditional on being paired with the object of one's bias. This interaction is also consistent with the model assumption that the IAT manifests as animus. The controls are necessary because observed differences in the outcome variable may be driven by factors unrelated to a dictator's implicit bias. Different specifications below may highlight different sets of these parameters in my analysis.

I.6.1 Dictator Giving

I continue with a graphical exploration of the IAT's relationship to giving. Figure I.3 shows the amount passed given a dictator's IAT score. Despite the IAT's popularity in academic work, there is no clear linear relationship between IAT score and amount passed ($\rho = -0.01$). Further, in each "column" of IAT score there appears to be a similar bimodal distribution of amount passed. This suggests that levels of implicit bias do not necessarily map into the behaviors of interest.

Figure I.3: Scatter Plot of IAT Score and Amount Passed



Notes: The solid line indicates the IAT-D "bias" threshold of 0.15
Source: Author's illustration

To confirm these findings econometrically, we turn to table I.9 which presents this paper's main estimates. In these models I restrict the sample to only the dictators in sessions with photographs, although the results hold when expanded to the full sample. Additionally, I have used both dictator and receiver dummies for African-American, rather than which race a subject is biased against. While this may be a coarse measure, this modeling technique makes more sense in terms of coefficient interpretation since IAT score is increasing in the level of anti-black bias. Further, these results are consistent with the discrete estimations from section V and robust to the alternate specification of "biased against receiver"[15].

In panel A of table I.9 I start with a simple OLS and regress percent shared on the

[15]See, for instance, tables I.7 and I.13.

parameters of interest. We see that neither implicit bias nor its interaction with a black receiver yields a significant predictor of giving. These results hold true in specifications that control for race and gender of dictator, the receiver, and both. Further, these controls also have no significant effect on giving.

However, the presence of a sorting option consistently and significantly decreases the amount shared by around 10%. This result suggests that in terms of giving behaviors, people aren't acting on their implicit biases, and perhaps are able to control any bias they may hold. Instead, social preferences unrelated to the IAT, especially pressure to give, appear to be strongly influencing these pro-social behaviors (or lack thereof).

Next, to account for the 27% of dictators who either gave nothing or opted out, I replicate the OLS results with a left-censored Tobit model[16]. These results are shown in table I.9, panel B, and are not categorically different than the OLS results. That is, IAT score is positive but insignificant, the interaction term is negative but not significant, controls lack significance, and the presence of a sorting option is strongly and negatively significant.

Following LMW, I assess the determinants of sharing in table I.10. Specifically, I compare the relative importance of implicit bias (in column 1) to the presence of the sorting option, as well as self-reported demographics that could potentially affect sharing (in column two). Again, one's amount of implicit bias does not significantly determine sharing. Magnitudes of these results are similar when I run the full model, including IAT score with demographic controls (column 3). Additionally, I calculate coefficients of partial determination[17]. This measure shows that not only does implicit bias lack statistical significance, but one's IAT score accounts for less than 4% of the unexplained variance and lacks economic significance as well.

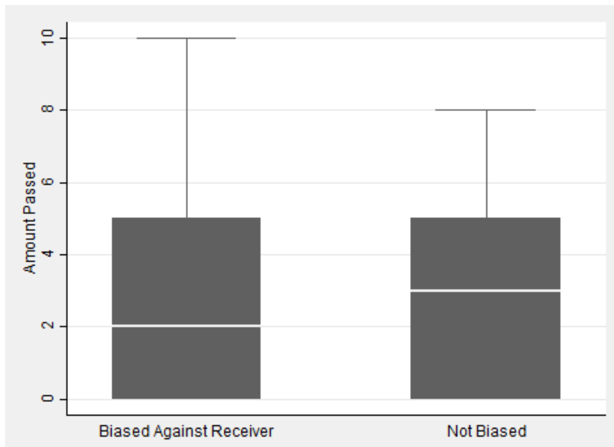
The above exercises hold true when instead of looking at the coarse measure of race of receiver, I look at the finer measure of being biased against one's receiver. In figure I.4, I

[16]Robust standard errors are calculated using jackknife estimation. A double-hurdle model (Cragg, 1971) would be inappropriate here because to account separately for the opt-out process requires restricting the sample to only those in sessions with sorting. Results from this model are presented in Appendix E.

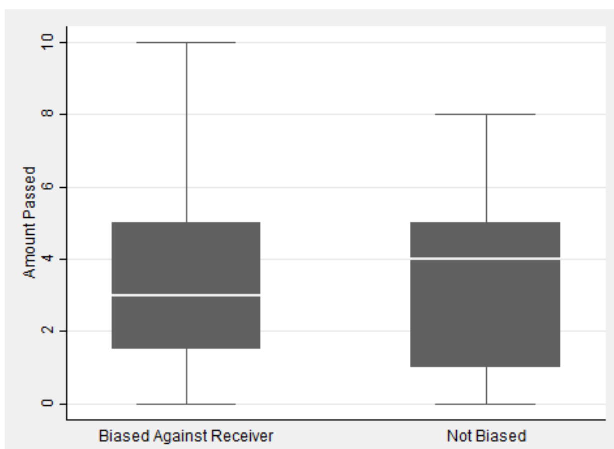
[17]$(R^2 - R_i^2)/(1 - R_i^2)$ where R_i^2 is the R^2 with predictor i removed from the equation.

Figure I.4: Sharing When Dictator is Biased

(a) Whole Sample



(b) Conditional on Staying In



Source: Author's illustration

Table I.9: The IAT's Effect on Percent Shared

Panel A: OLS					
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
IAT D-score	0.0197	0.0169	0.0286	0.0262	0.0389
	(0.0353)	(0.0351)	(0.0450)	(0.0385)	(0.0366)
IATxPassedBlack	-0.0617	-0.0439	-0.0494	-0.0595	-0.0665
	(0.0805)	(0.0795)	(0.0797)	(0.0867)	(0.0843)
Sorting Option		-0.101***	-0.0960***	-0.0944***	-0.0867**
		(0.0355)	(0.0368)	(0.0361)	(0.0377)
Dictator Controls			X		X
Receiver Controls				X	X
Constant	0.269***	0.343***	0.388***	0.339***	0.382***
	(0.0176)	(0.0278)	(0.0509)	(0.0460)	(0.0641)
Panel B: Tobit					
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
IAT D-Score	0.0298	0.0251	0.0361	0.0426	0.0549
	(0.0485)	(0.0481)	(0.0483)	(0.0535)	(0.0516)
IATxPassedBlack	-0.0854	-0.0574	-0.0632	-0.0883	-0.0959
	(0.110)	(0.109)	(0.109)	(0.120)	(0.117)
Sorting Option		-0.145***	-0.138***	-0.132***	-0.122**
		(0.0448)	(0.0462)	(0.0452)	(0.0469)
Dictator Controls			X		X
Receiver Controls				X	X
				(0.0478)	(0.0495)
Constant	0.222***	0.327***	0.366***	0.323***	0.359***
	(0.0254)	(0.0336)	(0.0688)	(0.0576)	(0.0849)
Observations	172	172	172	172	172

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

graph box plots for a further analysis of what happens when a dictator is biased against the race of his or her receiver. For these figures that means both passing to a black receiver when biased against blacks ($Receiver = Black|IAT \geq 0.15$) as well as passing to a white receiver when biased against whites ($Receiver = Black|IAT \leq -0.15$).

Clearly there is no difference in giving when I consider the whole sample in figure I.4a. But, this result also holds when I consider only those dictators who did not take an exit

Table I.10: Determinants of Sharing

Variable	**(1)**	**(2)**	**(3)**	**Partial R^2's**
IAT D-score	-0.00247		0.00850	
	(0.0169)		(0.0195)	0.036
Sorting Option		-0.0847**	-0.0830*	
		(0.0423)	(0.0448)	0.158
Age		0.00642***	0.00614**	
		(0.00245)	(0.00253)	0.148
Male		0.00345	-0.00223	
		(0.0407)	(0.0426)	0.005
Black		-0.0992*	-0.107**	
		(0.0502)	(0.0519)	0.185
Catholic		-0.000280	0.00947	
		(0.0708)	(0.0761)	0.012
Previous Experience		-0.0258	-0.0277	
		(0.0466)	(0.0468)	0.047
Major: Business or Econ		0.00407	0.00614	
		(0.0420)	(0.0429)	0.012
GPA		-0.112***	-0.110***	
		(0.0409)	(0.0411)	0.219
Constant	0.268***	0.657***	0.659***	
	(0.0176)	(0.164)	(0.164)	
Observations	172	146	143	
R-squared	0.000	0.138	0.136	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

option in figure I.4b. We will see a similar result regarding dictators choosing to opt out in the following subsection on Dictator Sorting.

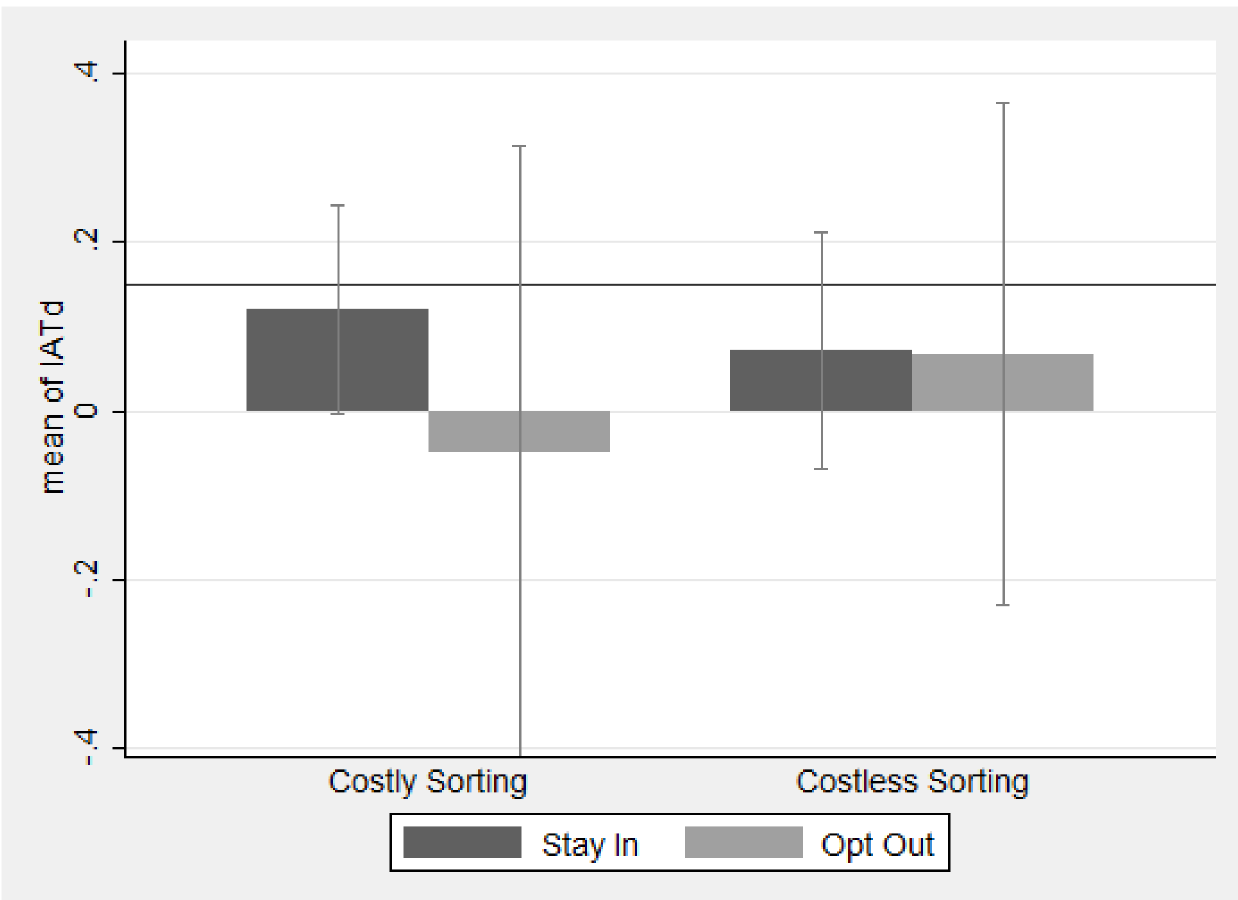
Finally, I compare the picture treatments to the anonymous ones. Using rank-sum tests, amounts given by the dictator do not appear to be different across these two treatment rows ($z = -0.039, p = 0.969$). This holds when we ignore baseline treatments and consider only those with a sorting option ($z = -1.268, p = 0.205$), or restrict the sample to dictators paired with the object of their bias ($z = -0.863, p = 0.388$). Since a dictator cannot see his or her receiver in the anonymous treatments, it is unlikely that implicit racial bias comes into play in this sharing decision. The lack of difference between the two treatment rows here is further indicative of the null results above.

Given the overwhelming evidence above, I now declare the first result, regarding implicit bias and dictator giving:

Result 2 *Amount of Implicit Bias (as indicated by IAT D-score) does not predict dictator giving*

I.6.2 Dictator Sorting

Figure I.5: Bar Graph of IAT Scores and Sorting



Notes: The solid line indicates the IAT-D "bias" threshold of 0.15
Source: Author's illustration

Perhaps the above results are indicative that biased dictators are forward thinking with regard to these biases or otherwise self-aware enough to recognize their biases. If so, they may be simply choosing not to enter sharing environments where they can express this distaste, or similarly choosing to express this distaste through their opt-out. However, we see in figure I.5 the average IAT score for dictators in treatments with an exit option. Under costly and costless sorting schemes, the mean IAT score is descriptively smaller amongst those who stay in (as compared to those who opt out), whereas in costless sorting the mean IAT score is essentially the same. However, in both cases, this difference is not significant (Costly: $t = 1.04$, $p = 0.30$; Costless: $t = 0.03$, $p = 0.98$).

Accordingly, I estimate the probability of opting out in table I.11. This model uses a probit regression and necessarily restricts the sample to only those dictators with an exit option (that is, those in sorting treatments, n=159,). The variable structure is intended to mimic the experimental design, using dummy variables for treatment and a measurement variable to indicate IAT score. In this model, there are no significant coefficients, suggesting that overall, one's IAT score does not seem to influence the decision to sort out, with this result holding even when controlling for both the financial and social costs of sorting.

Nonetheless, this exploration again calls for a deeper analysis. Following equation I.10 I look at the econometric results to confirm. In this case I ignore the anonymous treatments ($n = 127$) and run probit estimations to determine what effect (if any) IAT score has on the probability of opting out. Table I.12 shows the marginal effects of these estimations. Consistent with the results above, the IAT has no significant effect on sorting. This holds when I control for whether the sorting is costly and for race and gender of the dictator, receiver, and both. Similar to the analysis under dictator giving, the signs of these coefficients are also unexpected. We see that more biased dictators opt out less. Specifically, an increase in IAT score by 1 standard deviation leads to roughly an 11% smaller chance of opting out.

As a check, I examine what happens to sorting when a dictator is biased against the race of his or her receiver ($n = 75$). In this case I draw a bar graph in figure I.6. Confirming the results above, there is no evidence that bias has an effect on sorting, even when the dictator

Table I.11: The Probability of Opting Out

Probit Regression **Variable**	**Coefficient**
IAT D-score	-0.095
	(0.114)
Costless Sorting (Pictures)	0.369
	(0.253)
Costly Sorting (Anonymous)	-0.463
	(0.547)
Costless Sorting (Anonymous)	0.116
	(0.370)
Constant	-0.976***
	(0.184)
Observations	159

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

Table I.12: The IAT's Effect on Sorting

Probit Marginal Effects **Variable**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
IAT D-score	-0.114	-0.105	-0.108	-0.111	-0.115
	(0.0727)	(0.0744)	(0.0814)	(0.0871)	(0.0933)
IATxPassedBlack	0.234	0.206	0.214	0.201	0.207
	(0.175)	(0.177)	(0.185)	(0.194)	(0.202)
Costly Sorting		-0.0949	-0.0977	-0.0960	-0.102
		(0.0731)	(0.0730)	(0.0721)	(0.0722)
Dicator Controls			X		X
Receiver Controls				X	X
Observations	126	126	126	126	126

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

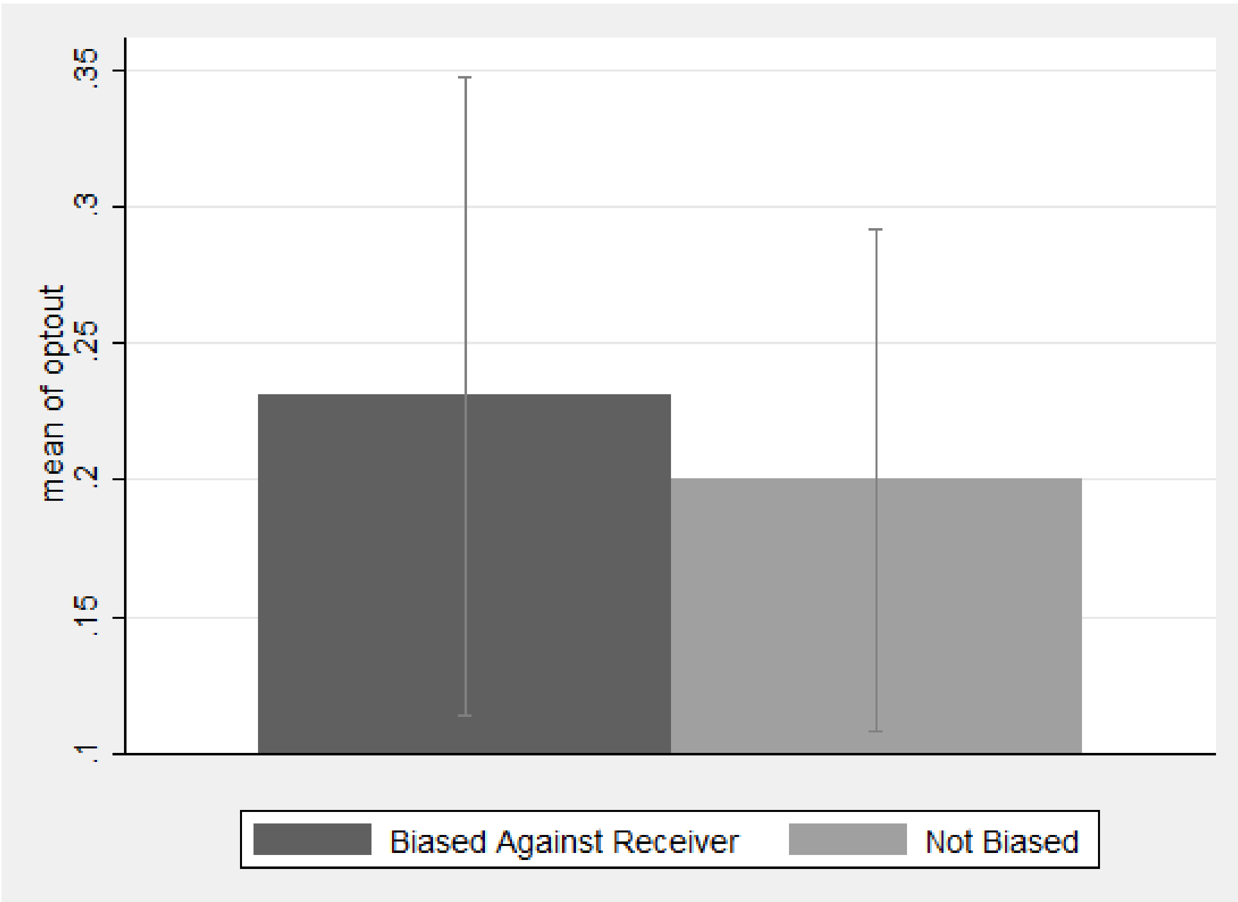
Source: Author's calculation

holds an implicit bias against the receiver's race.

Finally, we extend the cross-treatment exercise from above and compare anonymous sorting to sorting when photo information is present, by way of Pearson's test. Again, there is no statistical difference between opting out in the two treatment rows ($\chi^2 = 0.611, p = 0.434$).

This holds in costly sorting

Figure I.6: Sorting When Dictator is Biased



Source: Author's illustration

($\chi^2 = 0.623, p = 0.430$), and when passing to someone who's race you are biased against ($\chi^2 = 0.707, p = 0.400$), As such it is also unlikely that implicit bias is influencing giving on the extensive margin, inclusive of sorting decisions.

Result 3 *Amount of Implicit Bias (as indicated by IAT D-score) does not predict sorting in or out of the dictator game*

I.7 Small Gifts, a Robustness Check

Thus far, I have suggested that the IAT does not predict giving or sorting behaviors. However, I have also left the door open for dictators to have awareness of their biases, meta-cognitive abilities with respect to it, or both. This may suggest that differences in giving are more subtle than the ones suggested above. For instance, what if biased dictators are giving, but their giving is concentrated in small(er) gifts?

To test for this concentration, I utilize the *Dislike* variable from equation I.9 above, noting that this variable highlights cases of both pro-white and anti-white bias. I also generate dummy variables for various small gift amounts. I then run Pearson's χ^2 tests to see if giving in those small amounts is different for biased and non-biased dictators in each of the pictures treatments. Full results from these tests are depicted in table I.13.

Table I.13: The IAT and Small Gifts

	p-value for Gift Size:			
	0	**$\leq$ 1**	**$\leq$ 2**	**N**
No Sorting	0.738	0.209	0.369	48
Sorting	0.646	0.319	0.968	127
Receiver is Black	0.367	0.256	0.647	130
Sorting & Receiver is Black	0.310	0.181	0.753	94
Whole Sample	0.927	0.893	0.496	175

*** p<0.01, ** p<0.05, * p<0.1

Source: Author's calculation

Small giving is not different between biased and unbiased dictators in every specification. This result suggests that biased giving is not concentrated in small giving, and lends further credence to the above discussion of dictator giving as a whole.

Result 4 *Biased givers are no more likely to give small ($\leq$ \$2) gifts*

I.8 Conclusion

Racial bias is a persistent concern in the social sciences. In the past two decades, a proposed method of detecting it, the Implicit Association Test, has caught fire amongst the academics who study bias. At the time of this writing, the IAT's original paper has over 6600 citations, with researchers claiming it has implications on all sorts of economic outcomes, from workplace discrimination and managerial behavior, to egalitarian ideals and general social welfare. Yet, economists have only recently started to explore these claims in detail.

In this paper I have undertaken an in-depth examination of one of those claims in particular—that implicit (racial) bias is a predictor of pro-social behavior. I focused on these behaviors due to a growing literature suggesting the importance of the relationship between bias and pro-sociality. In doing so, I critique the extant literature stemming from the IAT. I then write a model of giving under conditions of implicit bias and conduct a laboratory experiment to test those model predictions.

Specifically, I test biased giving using a dictator game where acts of giving are both non-strategic and non-spontaneous, and therefore easily controlled (by the subject). Additionally, in some treatments I include a sorting (exit) option to see if biased givers simply choose to avoid the potential giving transactions altogether.

I find that, contrary to model predictions and previous literature, implicit bias fails to predict giving on both the extensive and intensive margins. That is, not only does implicit bias not predict amounts shared in the dictator game, it also does not predict examples of zero sharing, or the choice to exit a giving environments. Furthermore, these results hold not only in fine bins of analysis, but also wider and more powerful ones, such as when I restrict my sample to small gifts, or dictators paired with receivers of the race they hold implicit biases against.

To the best of my knowledge, this is the first paper to explore the implications of a Race IAT in an economics experiment. As such, the analysis in this paper represents a necessary step forward in this line of research that previously consisted of fascinating, but unsubstantiated claims.

The dictator game is a compelling example in that it consists of a very simple economic decision. If the IAT fails to map into this class of *cold-phase* decisions, what are the implications for decisions which may be more complex but also require more deliberation, such as hiring?

However, more research is needed as the dictator game is also a very clear-cut decision, and perhaps the IAT could be better used to predict maps into so-called *fuzzier* or multi-level economic decisions, decisions made in groups or ones where the use of heuristics have been shown to play a prominent role.

In this vein, we might think of implicit bias as mapping into a spectrum of pro-social activities with dictator giving at one end of the spectrum and a potentially different result at the other. If this is true, then future field experiments could prove to be a fruitful area of research.

Finally, as the popularity of the IAT grows in academia, so does its use in the public domain. As such, this paper also speaks to policy in general, and jurisprudence in particular. The typical anti-discrimination statute requires proof that harmful actions were "because of" discrimination. More and more, implicit bias is being recognized as a source of this liability. For instance, in a recent Supreme Court case regarding the Fair Housing Act, Chief Justice Roberts wrote for the majority that:

> Recognition of disparate impact liability under the FHA also plays a role in uncovering discriminatory intent: It permits plaintiffs to counteract *unconscious prejudices and disguised animus* that escape easy classification as disparate treatment. In this way disparate-impact liability may prevent segregated housing patterns that might otherwise result from covert and illicit stereotyping.
>
> (*Texas DoH v. ICP Inc.*, 2015)

Italics are my own. This means that bias is classified under the law as resulting in differential treatment even if one is not aware of the held bias, as in an implicit bias. And hence, we need further explorations of implicit bias and its potential to map into this sort of decision making, else we could be establishing ineffective policies.

Chapter II

Will Girls be Girls?

Risk Taking and Competition in an All-Girls School

II.1 Introduction

At the time of this writing, there are only 21 female CEOs in the S&P 500, and it is well known that this sort of vertical gender segregation exists throughout the labor market[1]. Despite enormous amounts of research on the gender gap for both wages and positions, the underlying reasons for this disparity remain unclear. However, one oft-cited explanation is a basic difference in risk preferences and/or the willingness to compete (Johnson and Powell 1994; Niederle and Vesterlund 2007). This hypothesis suggests that women, either naturally or through institutional environments, such as schools, develop skills not suited for upper management.

In this paper, we follow a promising course of literature in this area, which focuses on these differences and seeks to disentangle innate and immutable causes from those that are the products of our environments—the so-called "nature-or-nurture debate". In short, men and women tend to have different emotional responses to risk and also possess differing levels of confidence. The emotional responses are most likely hardwired ("nature") whereas the confidence levels may be a result of "nurture".

[1]Catalyst, Inc. Knowledge Center (http:// http://www.catalyst.org/knowledge/women-ceos-sp-500).

Given the profound amounts spent on education (over $600 Billion annually, almost 4% of US GDP), we choose to examine this debate in the context of gender-specific peer effects, by running an experiment in two closely matched, yet distinct educational environments—co-ed and single sex schools. Specifically, we study the risk taking and competitive behavior of girls educated in closely matched singer-gender and co-educational environments.

Of course it is extremely difficult to tease apart and isolate these factors, especially absent random assignment. To do so, we structurally estimate risk preference parameters, and directly include these preferences when estimating the decision to compete. We also examine the developmental and formative components of these behaviors by comparing middle school students (grades 7 and 8) with upper school students (Grades 11 and 12), and by including the respective co-educational boys in our analysis. Further, we study the structural differences that exist between families choosing to send their daughters to a single sex school as opposed to a co-educational one that has a similar mission. Consistent with the above hypothesis, we find that though girls educated at a single-gender school are among the most risk averse, they are also the most competitive—comparable in competitive behaviors to their male counterparts.

This chapter proceeds as follows. The next section describes the relevant literature informing our research. Section three outlines our experimental design. Section four describes our data and hypotheses. Section five describes and interprets our results and is followed by a concluding final section.

II.2 Literature Review

This paper follows from a long line of research disentangling the roles of nature and nurture in the gender gap for wages and occupations. Polachek (1981) argues that the gender gap is due to differences in abilities and preferences that result in occupational self-selection. This hypothesis echoes concerns laid out in Heckman (1979). The higher intermittency rate of women in the workforce can be used to explain the gender gap. Further, the apparent lack of women in the upper tiers of corporate management may be indicative that as a group they are less willing to make the risky decisions that are necessitated in the corporate world (Johnson and Powell, 1994). Similarly, Flory et al. (2015) find women shy away from applying to jobs with competitive payment schemes.

These preference-based arguments rely on previously unidentified utility parameters, suggesting a role for both the toolbox of experimental economics and richer data within those experiments. In their review of the experimental literature, Croson and Gneezy (2009) find that, in general, women tend to be more risk averse than men—a view consistent with conventional findings in psychology[2]. Further, they find that men are more likely than women to compete and that male performance levels increase more than female levels when under competition.

Though men are generally more likely to compete, Gneezy et al. (2003) use different incentive schemes and find that the performance of women in single sex tournaments is higher than in the non-competitive treatment. Further studies show women competing as much as men, given large enough rewards (Petrie and Segal, 2014). Thus, there are instances in which women will perform in a competitive environment, however these results focus on tournament entry without controlling for risk attitude. In a similar experiment, Datta Gupta et al. (2013) find that men select the tournament choice more often than women and that a woman's degree of risk aversion influences her decision to select the tournament (as opposed to a non-competitive piece-rate compensation scheme). Their experiment differs from Gneezy et al. (2003) in two key ways. First, they increase the incentives to compete, and second, they allow subjects to select their competitor. They find that participation in the tournament increases with the higher incentive scheme and that the gender gap is reduced when the subjects can choose their competitor[3]. These findings suggest a need for joint estimation.

[2]See Campbell (2002) for a review of these findings.

[3]For a more comprehensive view of the recent literature on the gender differences in competition, we suggest Niederle and Vesterlund (2011).

Given these observed differences, an outstanding question is what drives these differences in competition rates? Niederle and Vesterlund (2007) suggest a major channel is differences in confidence, with men over-competing and women shying away from competition. However, this difference is not inherently universal as Gneezy et al. (2009) observe female competition in a matrilineal society far exceeds that in a patriarchal one. This suggests that decreased competitive behavior may not be intrinsic to female nature, but rather elected by subjects due to familial and societal nurturing. We can also understand nurture to be institutional, such as in Leibbrandt and List (2014) where simple changes in job postings can eliminate the gender gap in wage negotiations.

Given the demonstrations that (1) nurture matters and (2) can affect the later life outcomes, early-life experiments provide an ideal testing ground for these theories. Cárdenas et al. (2012) vary the competitive task and find no gender differences in Colombia, with mixed results in Sweden. While in Austria, Sutter and Glätzle-Rützler (2015) find young boys consistently prefer to compete, and this preference persists throughout adolescence.

Furthermore, educational environments and interventions allow us to study the roles of gender-peer effects. For instance, Fryer and Levitt (2010) note that there is no gender gap for mathematics in kindergarten, yet after six years girls perform worse than their male counterparts. However, a growing body of field and natural experiments suggests that single sex schooling can improve student performance, be it math skills and self-confidence (Eisenkopf et al., 2015) test scores and college attendance (Park et al., 2013), or grades and pass rates (Booth et al., 2014). However, Oosterbeek and Van Ewijk (2014) do not find strong gender effects on performance when the unit of treatment is an economics workgroup.

In the U.K. Booth and Nolen (2012a; 2012b) separately examine risk taking and competitive behaviors of students in single sex schools. Using a one-shot gamble for the risk task and a choice of payment scheme in the competition paper, their results indicate that girls from single gender schools behaved closer to boys from co-ed and single gender schools than girls from co-ed schools. However, only those students who sat for the 11+ exam were eligible for being placed in a single sex school, and perhaps more importantly, their parents needed to be willing to have them sit for the 11+ exam, suggesting the need for further research[4].

In this vein, our paper continues and extends the line of research on the gender peer-effects of risk taking and competitive behaviors. Specifically, we revisit the question of single sex schooling as a mechanism to increase tournament entry using subjects from two closely matched American schools (one single sex and one co-educational)[5]. By structurally estimating a CRRA utility function, we extend previous examinations of risk-preference by also studying the distribution of preference parameters in our data. Further, we include these risk attitudes directly when estimating competitive behavior. Finally, we contribute to the literature by considering the role adolescent development has on these behaviors by including both middle school students and late secondary students from both schools in our analysis.

II.3 Experimental Procedures

To investigate the risk taking and competitive behaviors of young women we utilized two financially incentivized tasks henceforth referred to as the risk task and the competition task, respectively. The subjects were informed that they would be paid based on their performance in the tasks, and that only one task would be selected for payment. In order to study these behaviors in the context of gender peer-effects and education, the experiments were conducted using students in middle school (7th and 8th grade) and late secondary education (11th and 12th graders) at two different, yet closely matched academic institutions.

In the risk task, we follow a Holt and Laury (2002) style multiple price list design to elicit a subject's risk attitude. In it, a subject makes choices over a series of ten gambles.

[4] http://www.elevenplusexams.co.uk/advice/what-is-11-plus.

[5] The schools selected for our analysis serve identical educational markets and have common mission statements. Therefore, these two schools closely compete for applicants and the only major difference is that one is a single sex school and the other is not.

Each gamble consists of a choice over two simple lotteries, one being relatively "riskier" than the other. The 'safe' lottery can produce earnings of either \$8.00 or \$6.40; while the 'risky' lottery can produce earnings of either \$15.40 or \$0.40 (these payments are 4x the baseline levels used in Holt and Laury, 2002). The gambles were presented in order, starting from winning the low prize with certainty, and decreasing that probability 10% for each successive gamble (so in the last gamble, there was a 90 percent chance of the higher-payoff outcome and a 10% chance of the lower-payoff outcome). A subject switching to the 'risky' lottery relatively sooner indicates a lower degree of risk-aversion. To assist with the probabilities, subjects were presented with gambles on successive sheets of paper displaying the gambles numerically (with probabilities expressed in terms of the throw of a 10-sided die) and also graphically (as pie charts). Examples of these risk decision sheets are in appendix F. For the purposes of this experiment we assume a CRRA utility function and structurally estimate the risk parameter, ρ, in a fashion similar to Andersen et al. (2008).

We choose to measure competitiveness as a self-selection into a competitive setting. Specifically, our competition task elicits a subject's willingness to compete following Gneezy et al. (2009). In it, each subject is randomly assigned a member of his or her session as a competitor and given ten chances to lob a tennis ball underhand into a bucket from three yards away and paid based on his or her performance in this task. Subjects are told that the assignments were determined before the session, and that they would not be told who they are paired with in the experiment. However, they can observe the other subjects in their group. Since these assignments are pre-determined the person they are competing against could have chosen either payment method, and assignments are not commutative. That is, the person you are competing against is not necessarily competing against you, let alone competing at all. Before beginning to lob, each individual subject must select a preferred payment scheme—which is either "piece-rate" or "competitive".

The performance of one's competitor is immaterial for the piece-rate scheme. Subjects are simply paid \$2 for each successful lob (meaning the ball was thrown inside of the bucket).

However in the competitive payment scheme, subjects are paid \$0 if they fail to lob more than their competitor; if they lob more than their competitor, they are paid \$8 for each successful ball that exceeds their competitor's number of successful lobs. For example, if a subject elects to compete and successfully lobs 4 balls, while her competitor successfully lobs 2 balls, she earns \$16. On the other hand, if she had only lobbed 1 ball in the bucket, she earns nothing. Ties are paid \$0 and negative earnings are not permitted.

Paper instructions are given to subjects, and read out loud to all participants in the session. After completing the competition task, we randomly selected which task is chosen for payment by way of a coin flip. All payments were made to subjects in the form of gift cards to either Starbucks or Chick-Fil-A fast-food restaurant (subjects chose which they preferred to receive). After all tasks were completed subjects filled out a short demographic questionnaire. Complete subject instructions as well as the questionnaire can be found in appendix F.

In many of the previous experimental studies of gender and education, there have been obvious sample selection issues. Ours is not unique in this respect. However, gender differences aside, the schools were chosen based on similar mission statements and the respective educational markets they serve. Furthermore, we are the first to undertake an experimental study of risk and competition within a United States single sex institution. While our confidentiality agreement with the schools prohibits us from naming the schools, the similarities in the two institutions lead us to believe, *a priori*, that observed differences in choices will be due to nurture rather than nature. That is, they arise from the differences in the schooling environments, parenting, and peer influences. Therefore, our study is on the structural differences that exist between families that elect to send their daughters to private single sex school versus an alternate private co-ed school with a similar mission. While we are not able to precisely isolate the mechanism that generates the differences observed, our results are informative to the debate surrounding these preference parameters and gender. In the next section we discuss this selection argument further.

II.4 Data

Table II.1: Number of Subjects

	Females (SS)	Females (CE)	Males (CE)	Total
Middle School (Grades 7-8)	80	25	40	145
Upper School (Grades 11-12)	42	56	34	132
Total	122	81	74	277

Our experimental subjects were housed in the middle and upper grades of their schools. Table II.1 describes the breakdown of experimental subjects by school and grade level. Overall, 277 students participated in the experiment, with a roughly equal split between upper schools and middle schools. 44% of the subjects attended the SS (Single sex) school, and boys comprised 47.7% of the CE (Co-ed) subjects, and 26% of the sample in all.

As part of the experiment, we aimed to demonstrate the demographic closeness of these schools. Publicly available data from the National Center for Education Statistics are shown in Figure II.1. While the SS school serves a higher percentage of students of color, they are otherwise strikingly similar, particularly with respect to the universe of private education as a whole.

Table II.2: Summary Statistics—Sports Participation

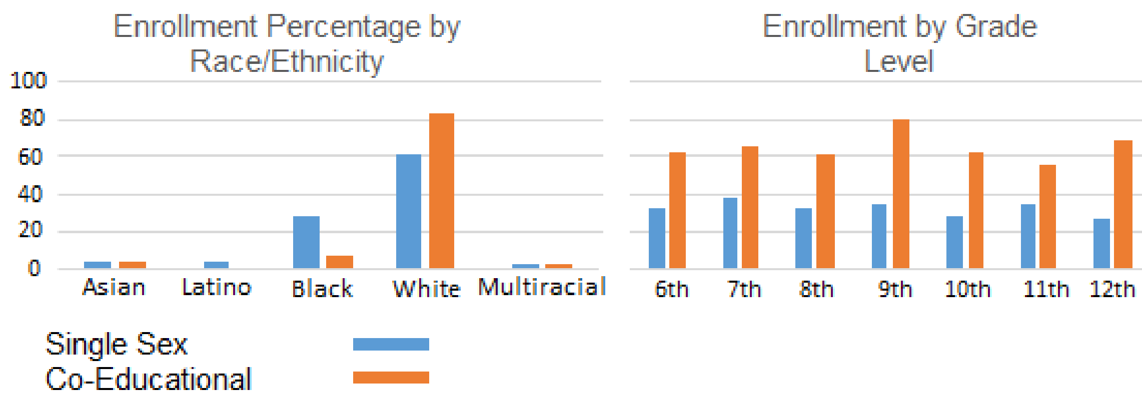
Panel A: Grades 7-8			
Plays in a:	Females (SS)	Females (CE)	Males (CE)
Team Sport	0.75 (0.43)	0.72 (0.46)	0.5 (0.42)
Hand-Eye Sport	0.71 (0.46)	0.72 (0.46)	0.8 (0.41)
Lettering Sport	0.86 (0.35)	0.88 (0.34)	0.88 (0.34)
Panel B: Grades 11-12			
Plays in a:	Females (SS)	Females (CE)	Males (CE)
Team Sport	0.53 (0.50)	0.34 (0.48)	0.38 (0.49)
Hand-Eye Sport	0.56 (0.50)	0.43 (0.5)	0.35 (0.49)
Lettering Sport	0.73 (0.45)	0.57 (0.5)	0.59 (0.5)

Notes: Presented as difference in mean with standard deviation in parentheses

Additionally, we gathered data on sports participation, educational attainment, and fam-

Figure II.1: Demographic Comparison of SS and CE Schools

Characteristics	**Single Sex**	**Co-Educational**
Locale	City: Large(11)	City: Large(11)
Type	Regular (elementary or secondary)	Regular (elementary or secondary)
Affiliation	Nonsectarian	Nonsectarian
Student Body	All-female	Coed
Days in Year	178	178
Hours in Day	7	8
Library	yes	yes
Total Teachers (FTE)	24.3	91.8
Total Students	223 (grades 6-12)	725 (grades K-12)
Student/Teacher Ratio	9.2	7.9



Source: National Center for Education Statistics

ily environment within these schools. These data not only serve to help us test hypotheses relating to risk and competition (described below), but also further illustrate the demographic similarities of the schools. Table II.2 provides descriptive statistics on the rates of sports participation among subjects. As shown in the table, participation rates appear equal within age groups regardless of gender of school. Participation hovers around 75% in middle school, but begins to diverge as the children get older, presumably because older children have access to a greater array of extracurricular activities.

In fact, we illustrate this closeness by failing to reject the hypothesis that team sport participation differs between schools in either the middle school or upper school test groups

(Kruskal-Wallis test, $p = 0.93$, $p = 0.22$, respectively). However, we can conclude that participation in team sports decreases across the two age groups (Pearson's χ^2, $p = 0.00$)[6].

Table II.3: Summary Statistics—Educational Attainment

Panel A: Grades 7-8			
Variable:	Females (SS)	Females (CE)	Males (CE)
HW hours/night	2.02 (1.24)	1.69 (0.79)	1.43 (0.98)
Math test hours	2.23 (7.25)	1.70 (1.22)	1.22 (1.2)
Highest math class	2.58 (.522)	2.56 (0.51)	2.83 (0.75)
# of AP classes	0.05 (0.28)	0.04 (0.2)	0.03 (0.16)
# of AP Math/Sci	0.00 (0)	0.04 (0.2)	0.03 (0.16)
Panel B: Grades 11-12			
Variable:	Females (SS)	Females (CE)	Males (CE)
HW hours/night	2.53 (1.09)	2.23 (1.3)	2.24 (1.21)
Math test hours	2.02 (4.52)	1.56 (1.27)	1.76 (1.9)
Highest math class	6.09 (1.14)	6.00 (0.74)	6.08 (0.67)
# of AP classes	2.2 (2.08)	3.08 (2.25)	2.61 (2.41)
# of AP Math/Sci	0.84 (1.1)	1.23 (1.43)	1.05 (1.74)

Notes: Presented as difference in mean with standard deviation in parentheses

We see a similar pattern when looking at educational attainment per student. Table 3 examines activities such as coursework and time spent studying. Again, group means are fairly similar by age but differ between the middle and upper schools. Perhaps the two most illuminating measures of educational attainment we have are the coursework variables: number of AP classes taken and highest math class taken since these measures are housed entirely within the schools. With few exceptions, middle schoolers do not sit for AP classes, and within the upper schools, students take roughly the same number of AP classes (Kruskal-Wallis test, $p = 0.13$). Furthermore, when we restrict the scope of educational attainment to just mathematics, the students remain similar. Every subject in the sample has progressed beyond 6th grade math. In middle school, the highest math class a subject has taken is usually Pre-Algebra or Algebra 1, (KW, $p = 0.30$), and Pre-Calc for high schoolers (KW,

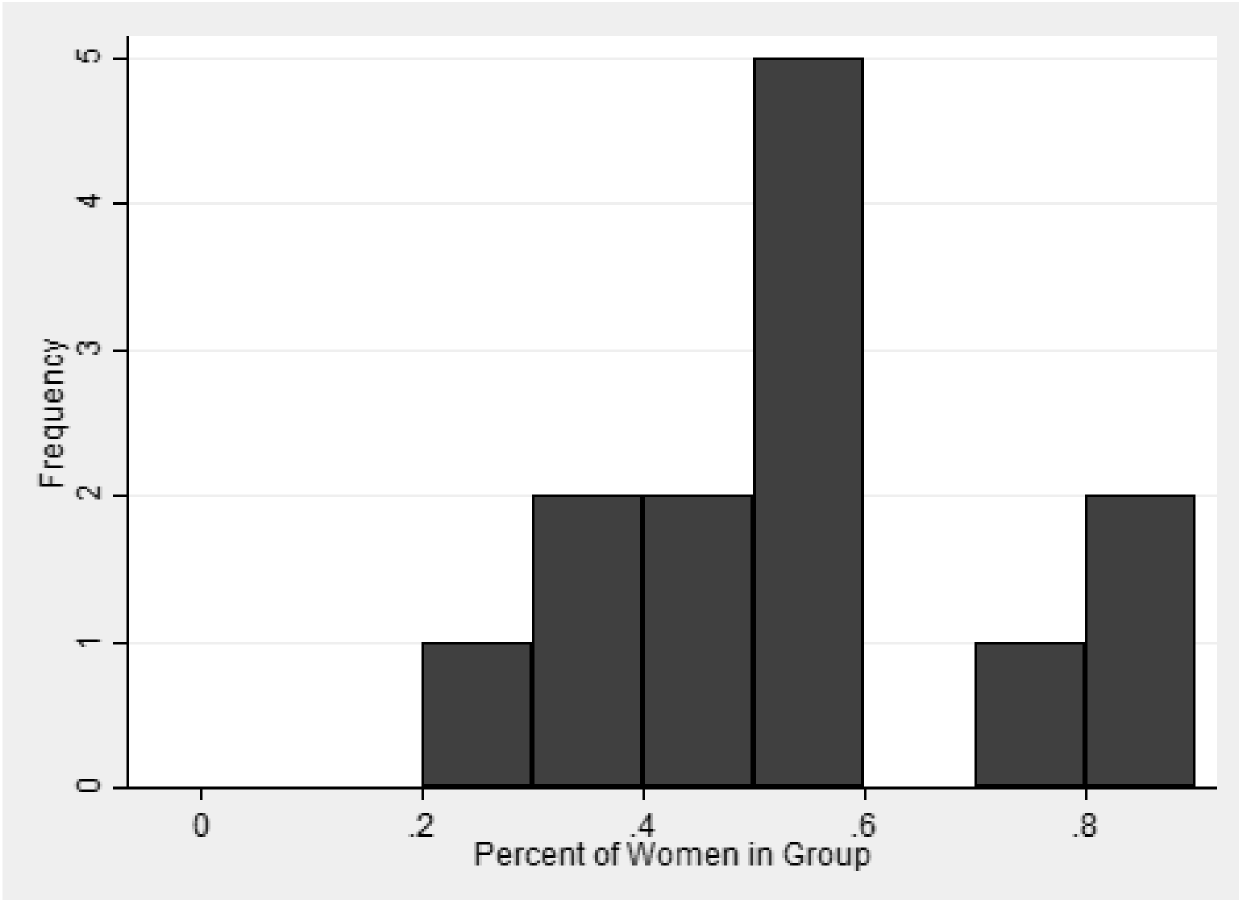
[6]To clarify, by team sports we mean baseball, basketball, cricket, football, hockey, kickball, lacrosse, soccer, softball, ultimate frisbee and volleyball.

$p = 0.36$).

Unfortunately, we cannot randomly assign subjects into a single sex school. As such, a concern with selection into the schools is that there is an unobserved component of the education production function that affects both the risk taking and competitive behavior of girls as well as their school choice. However, the prior closeness between the two schools as well as these demographic similarities and our survey data on athletics and family environment suggest that any family considering enrollment in one likely considers the other as well. One potential difference is that the co-ed school serves more grades. If increased exposure to single sex schooling increases its impact on risk taking and competitive behaviors, we would expect to see larger difference in the older girls. Regardless, we are comfortable positing that our estimates reflect the treatment effect of the all girls school, conditional on selecting into it.

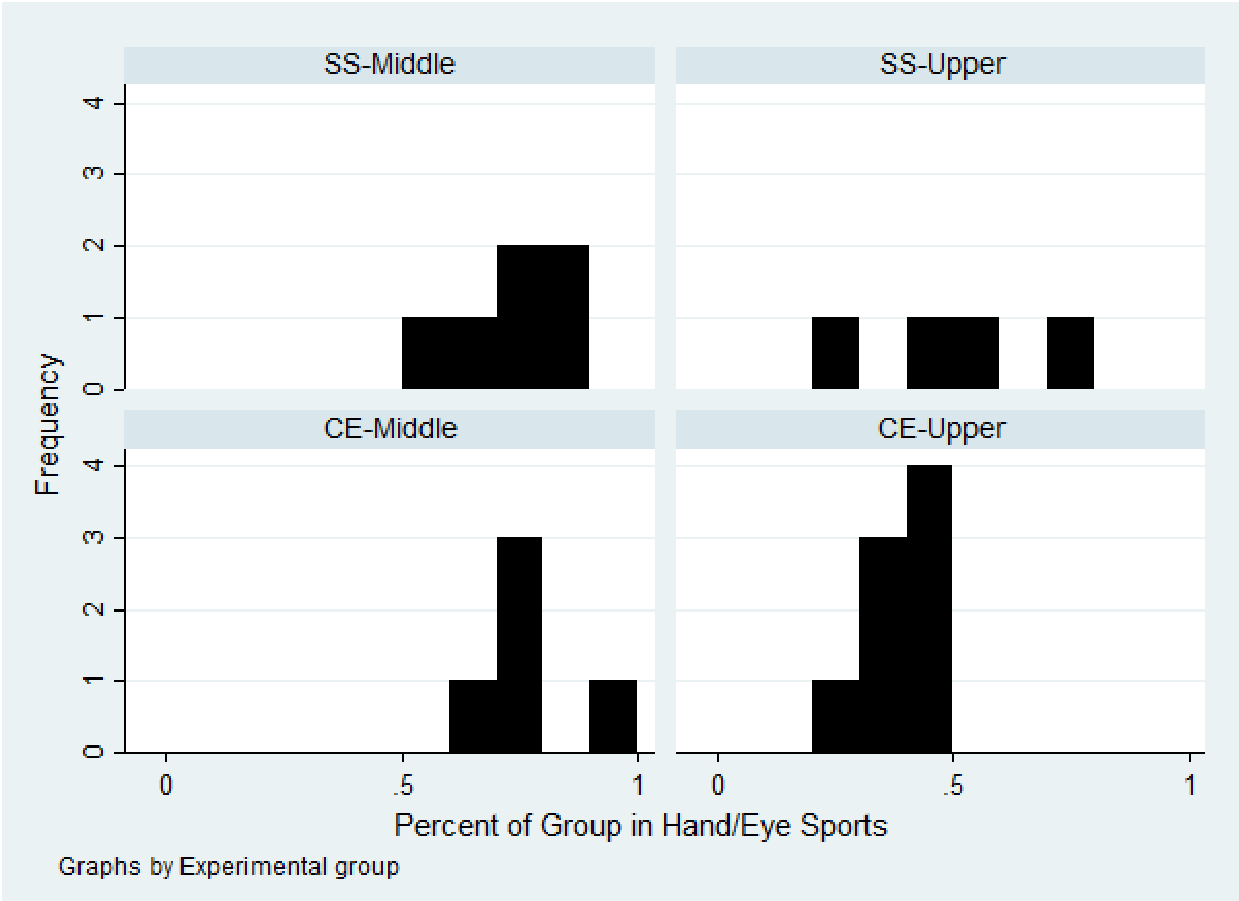
Our interest in demographics also reflects our prior belief that a subject's willingness to compete is not only affected by personal preferences, but also the composition of the experimental group. This belief echoes the concerns made by Gneezy and Rustichini (2004), among others. However, here group composition goes beyond number of athletes in one's group, and also reflects the gender composition of the group in the co-ed school, since experimental sessions were conducted within (but not across) schools and subjects were not separated into all-male or all-female groups at the co-ed school. We look more closely at gender composition below; however we present an overview of the group composition in Figure II.2, which shows the distribution of CE groups by their percent women. Obviously, SS groups are entirely women, so adding them to the figure would merely skew the distribution. CE groups are 52% female on average. It is noteworthy that while we do not have any entirely female groups within the CE sessions, we do have groups that are mostly women (75% or greater). Furthermore, 31% of groups are more than 50% female with another 27% being 50% exactly. This variation in gender mix allows us to control for the SS-CE comparison econometrically.

Figure II.2: Gender composition of CE groups



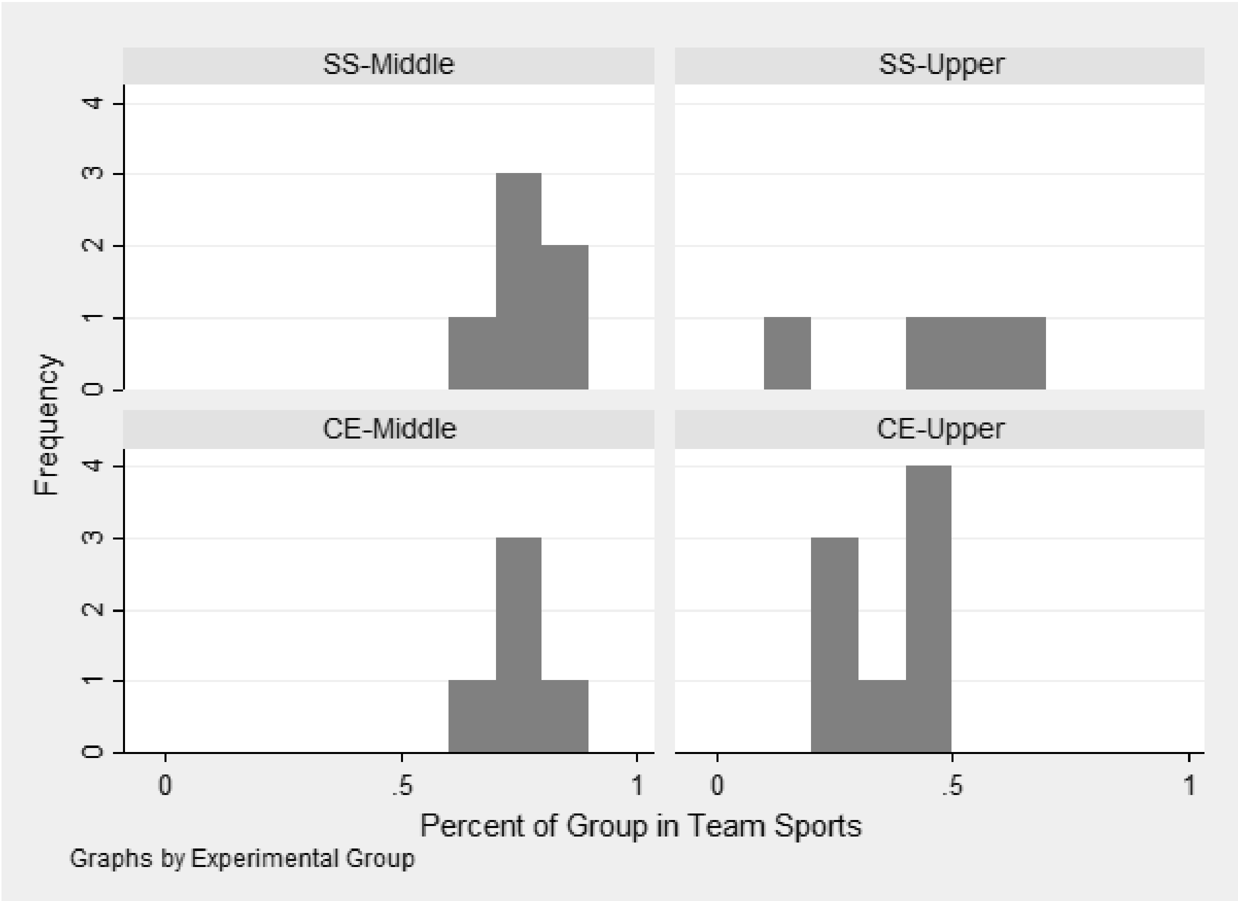
Figures II.3 and II.4 examine the athletic composition of the participant groups by percent of players on team sports and players on sports that require substantial hand-eye coordination, respectively. Across figures II.3 and II.4 these distributions are similar since many team sports involve significant hand-eye coordination (and vice-versa). Within each figure, we see the distributions starting out very tightly packed in middle school and becoming diffuse as we move to older students, confirming the earlier conclusion that sport participation declines as we transition from middle to upper school. Specifically, the SS upper schoolers have the widest distributions of both team and hand-eye sports. As discussed earlier it is important to note that here, the CE sessions were, by construction, co-educational. Thus, figures 3 and 4 only consider the 2x2 design of experimental sessions (CE and SS; middle school and upper school), as opposed to looking at all 6 test groups separately.

Figure II.3: Hand-Eye sports participation by experimental group



Recall, the purpose of these experiments is to revisit competing theories of nature versus nurture. In particular, we examine the nurturing effects that selecting into an all-girl schooling environment has on the risk taking and competitive behavior of women. Given this goal and the resulting data, it is now appropriate to discuss relevant hypotheses. First we consider the 4 central hypotheses, which are informed by conventional wisdom and the recent literature:

Figure II.4: Team sports participation by experimental group



Hypothesis 1 *Girls educated in a single sex school are less risk averse than girls educated in a co-education school.*

Hypothesis 2 *The risk profile of girls at a single sex school is similar to that of boys at a co-educational school.*

Hypothesis 3 *Girls in a single sex school are more likely to elect to compete than girls in a co-education school.*

Hypothesis 4 *The competition profile of girls at a single sex school is similar to that of boys at a co-educational school.*

We exploit the closeness of our sample and the richness of our dataset in our examination of these hypotheses. To this end, we also consider several tangential hypotheses, which follow from H1-H4, and are unique to our experiment:

Hypothesis 5 *5.a: Competitive sports positively influence competitive behavior.*
5.b: Competitive sports positively influence risk taking behavior.

This influence may come about from multiple channels. For example, people who participate in sports may be more willing to compete in the experiment, non-athletes may be less willing to compete when in an athletic group, or some combination thereof.

These experiments also investigate the influence adolescent development has on this behavior. Accordingly, we specify a hypothesis addressing differences across age groups.

Hypothesis 6 *The behavior of SS young girls is approximately equal to the behavior of CE young girls.*

Given our supposition that "nurture exceeds nature", we expect younger girls to be similar across schools. However, we want to know if the differences between girls are causing selection into individually optimal educational environments, or are different schools shaping different behaviors in female students? In addressing this question, we implicitly view boys as a reference group, and assume that the difference between girls and boys is inversely related to the year in school (which proxies the length of exposure to educational environment). We now specify two final hypotheses:

Hypothesis 7 *Due to exposure, differences in competition rates between SS and CE girls are greater in the older cohorts.*

Hypothesis 8 *Competition rates increase with the number of girls in a group.*

Differential rates in competition imply that the presence of boys stabilizes or normalizes female behavior. Thus there may be group effects with regards to the gender makeup of the group. Coming full circle, these suggestions are closely related to hypothesis 1—the risk posture of girls across schooling. However, hypothesis 8 does not identify the channel for choice to compete. That is, we remain agnostic whether we have observed SS girls being more competitive or a preference for competing (not competing) against girls (boys).

II.5 Results

Table II.4: CRRA Risk Attitudes

Panel A: Summary Statistics					
SS Middle	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
0.46	0.50	0.40	0.38	0.36	0.38
(0.05)	(0.01)	(0.05)	(0.08)	(0.04)	(0.05)

Notes: Presented as as means with standard deviation parentheses

Panel B: Rank-Sum Test Results	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
SS Middle	0.00 (35396)	0.00 (16510)	0.00 (48132)	0.00 (30261)	0.00 (24416)
SS Upper		0.00 (6628)	0.00 (21388)	0.00 (12858)	0.00 (10165)
CE F-Middle			0.53 (9563)	0.00 (5488)	0.16 (4243)
CE F-Upper				0.01 (18093)	0.56 (14431)
CE M-Middle					0.07 (8458)

Notes: Presented as p-value with adjusted variance in parentheses

Hypotheses 1 and 2 concerned risk attitudes, so first we consider the risk profiles of our subjects. Using the data elicited from our risk task, we specify a CRRA utility function and conduct a structural estimation of risk attitudes. Table II.4, panel A, lists summary statistics on risk preference. The reader will note that preferences range from risk neutral ($\rho = 0.028$) to solidly risk averse ($\rho = 0.53$), where SS subjects are the most risk averse, regardless of age group. There is no statistical difference in risk aversion between age groups ($p = 0.376$). However, boys are considerably less risk averse than girls ($p = 0.00$), consistent with findings in the previous literature. Both in general and throughout the sample, SS girls tend to be the most risk-averse, with CE girls being less so, and CE male subjects are the least risk averse.

Figure II.5: Box plot of risk Preferences across test groups

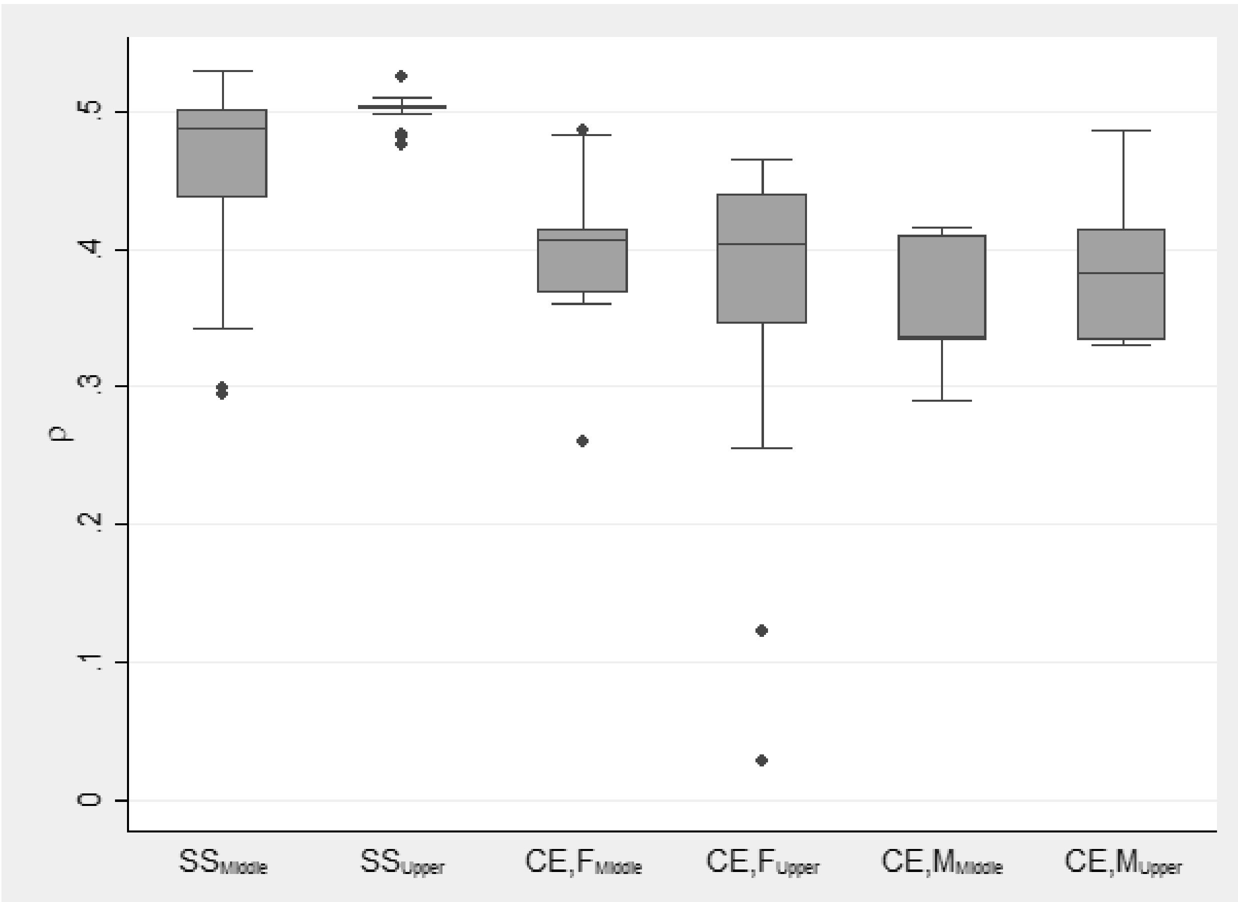


Figure II.5 shows information on risk preference using box plots. It not only illustrates that the central tendency of the SS girls is to be more risk averse, as stated above, but also how uniform SS girls are in their risk preference. Boys in our sample were the least risk averse. Hypothesis testing confirms this result—that SS girls are in fact more risk averse than both CE boys and girls. This result is remarkably robust regardless of specification. It holds in the middle and upper schools as well as both parametrically and non-parametrically.

As a formal test of hypothesis 1 we conduct Mann-Whitney (MW) tests (Table II.4, panel B) against the null hypothesis that risk attitudes between groups have the same central tendency. For both middle and upper school girls, we strongly reject this null in favor of the alternative that SS are more risk averse ($p = 0.00$, $p = 0.00$). Furthermore, the results hold when we ignore grade level and just compare school type (SS vs. CE, $p = 0.00$). Formal testing of hypothesis 2 requires the same methodology, with the specification modified to compare SS girls to (CE) boys. Table II.5, panel B, shows that in both the middle school and the upper school, there is a significant difference between SS girls and (CE) boys ($p = 0.00, 0.00$); thus, we strongly reject the nulls from both hypotheses 1 and 2.

Result 1 *Girls from the single sex school are not significantly less risk averse than coed girls*

Result 2 *Girls from the single sex school are significantly more risk averse than coed boys*

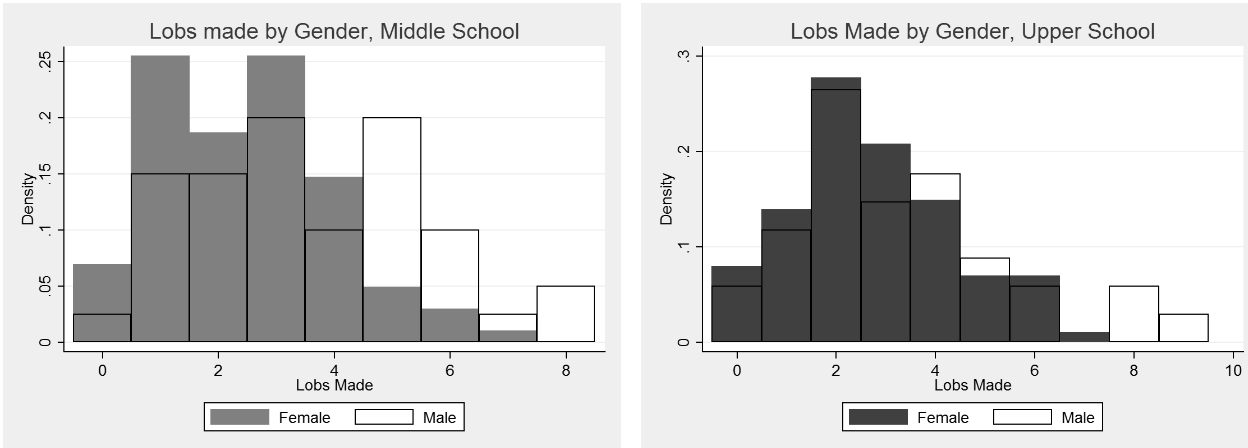
However, despite this apparent risk aversion, it is remarkable that SS girls opt to compete more than their CE counterparts. This result is discussed in more depth below, following a discussion of the gender neutrality of the competition task.

Table II.5: Lobs made by test group

Group:	**Mean**	**SD**	**Min**	**Max**
SS-Middle	2.363636	1.485979	0	7
SS-Upper	2.955556	1.91828	0	7
CE, F-Middle	2.8	1.632993	0	6
CE, F-Upper	2.571429	1.38639	0	6
CE, M-Middle	3.625	2.034163	0	8
CE, M-Upper	3.323529	2.211822	0	9
Total	2.841155	1.782576	0	9

We examine this appropriateness three ways: loosely termed performance, earnings, and rationality. Performance is simply how well a subject did in the task. That is, how many balls were successfully lobbed in the bucket, described in detail in table II.5. We see that at least one person in each test group missed every ball, and no subject successfully lobbed all ten. However, there is some difference between boys and girls. This difference is perhaps better illustrated in figure II.6, the histograms of performance by gender, given grade level, which shows this difference is mostly coming through the tails. More girls are successfully

Figure II.6: Histogram of performance by gender, by age

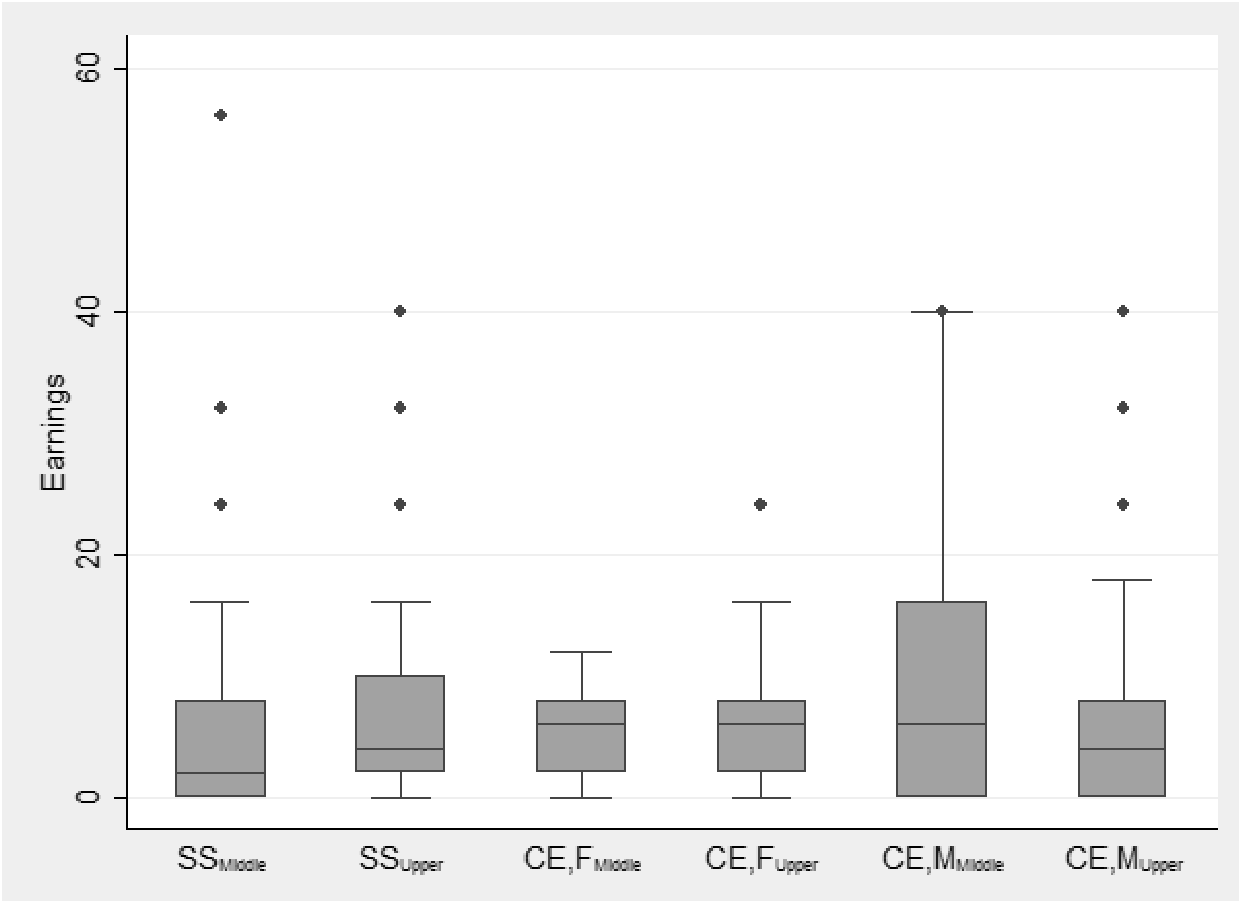


lobbing fewer than three and no girl lobbed more than seven. While there is some significance to this performance gap (Pearsons $\chi^2, p = 0.01$), there is also a great deal of overlap between the distributions. This overlap holds for both age groups, and the difference is only coming through the right tail. As such, we are comfortable drawing statistical inferences from this task.

Furthermore, subjects tend to earn the same amount across genders and age groups. Figure II.7 depicts box plots of earnings; note that each of our test groups had similar median earnings. Most of the sample earned less than \$20, and girls are well represented among the highest earners. In fact, of the 15 possible MW pairwise tests in table II.6, the only test groups who experienced significantly different earnings were the SS middle schoolers compared to the SS high schoolers ($p = 0.03$).

Finally, we ask the question: how much would subjects have earned had they switched their competitive choice (If one who chose to compete opted for piece-rate payment and vice versa). For exposition, we term this maximizing argument "rationality". We measure rationality in absolute (differenced-earnings) terms. That is if a subject would have earned more in the other payment scheme (differenced-earnings are less than zero), he or she made an "irrational" choice. In these terms, Mann-Whitney tests in table II.7 show no group was pairwise more rational than another. Specifically, the smallest p-values were 0.08 and 0.09,

Figure II.7: Earnings by Test Group



for SS middle schoolers when compared to their two female upper school counterparts. All other p-values were greater than 0.31[7].

Given these similar earnings patterns across the sample, we formally examine the decision to compete in hypotheses 3 and 4. Table II.8 outlines descriptive statistics and the results of our tests of these hypotheses, conducted using Pearson's χ^2 tests. Middle school boys appear to be the most competitive, opting to compete 60% of the time. Remarkably, even though SS girls are more risk averse than CE girls, they also choose to compete more often. Addressing hypothesis 3, SS middle schoolers compete significantly more than both CE middle school girls ($p = 0.00$) and upper school girls ($p = 0.01$). SS upper schoolers compete more than CE middle girls ($p = 0.04$) but their choices are not significantly different than the

[7]These results are robust to when we consider only the binary state (that is rational or irrational), where no p-value is less than 0.10.

Table II.6: Experimental Earnings

Panel A: Summary Statistics					
SS Middle	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
5.17	8.13	5.04	5.89	9.25	7.35
(8.14)	(9.74)	(3.66)	(5.25)	(10.88)	(9.58)

Notes: Presented as as means with standard deviation parentheses

Panel B: Rank-Sum Test Results	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
SS Middle	-2.10	-1.20	-1.65	-1.86	-1.04
	(0.036)	(0.23)	(0.00)	(0.06)	(0.30)
SS Upper		0.731	0.561	-0.192	0.68
		(0.464)	(0.575)	(0.85)	(0.494)
CE F-Middle			-0.166	-0.683	0.078
			(0.868)	(0.494)	(0.938)
CE F-Upper				-0.66	0.233
				(0.51)	(0.816)
CE M-Middle					0.646
					(0.518)

Notes: Presented as z-score with p-value in parentheses

CE upper school girls ($p = 0.53$). There is no statistical difference in the choice to compete between SS middle schoolers and either middle school or upper school boys (Pearson's $\chi^2, p = 0.67, 0.57$; respectively). In the SS upper school we see a similar result. Though middle school boys compete marginally more, this result is not statistically significant. Thus, our findings regarding competition are predominately consistent with hypotheses 3&4—that SS girls compete similar to boys and more than CE girls.

Result 3 *Middle-school girls in a single sex school are significantly more likely to compete than girls in a co-educational school.*

Result 4 *The competition profile of girls at a single sex school is not significantly different from that of boys at a co-educational school.*

While non-parametric tests provide a good overview of the competition choice problem, a structural model allows us to recover deep preference parameters and jointly estimate

Table II.7: Subject Rationality

Panel A: Summary Statistics					
SS Middle	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
5.17	8.13	5.04	5.89	9.25	7.35
(8.14)	(9.74)	(3.66)	(5.25)	(10.88)	(9.58)

Notes: Presented as as means with standard deviation parentheses

Panel B: Rank-Sum Test Results	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
SS Middle	0.0855	0.3658	0.0837	0.0687	0.6571
	(35168)	(16309)	(47649)	(30033)	(24159)
SS Upper		0.4378	0.7444	0.7193	0.3980
		(6591)	(21224)	(12817)	(10112)
CE F-Middle			0.6071	0.3091	0.8350
			(9453)	(5436)	(4200)
CE F-Upper				0.4418	0.4748
				(17933)	(14311)
CE M-Middle					0.3170
					(8451)

Notes: Presented as z-score with p-value in parentheses

those parameters while providing a more complete interpretation of the decision to compete. Of course, there are additional methods for examining this decision to compete, which we consider for robustness and sensitivity. Assuming the choice to compete is a latent process, a great benefit of our design is our use of multiple tasks to identify these latent parameters. Furthermore, there is reason to believe risk attitudes affect the decision to compete. For instance, Datta Gupta et al. (2013) find that a woman's degree of risk aversion influences her decision to select each compensation scheme. Using our experimental data, we estimate a joint structural model of risk attitude and the decision to compete. However, we show that these attitudes do not significantly impact the decision to compete.

Marginal effects from the relevant estimation are shown in table II.9, with upper school boys serving as the reference group[8]. These effects show significant differences in risk attitude between upper school boys and both male and SS middle schoolers. Consistent with our

[8] Consistent with the existing literature, we conduct our estimation using frequency weighting.

Table II.8: Decision to Compete

Panel A: Summary Statistics					
SS Middle	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
0.558	0.4	0.16	0.339	0.6	0.5
(0.499)	(0.495)	(0.374)	(0.477)	(0.496)	(0.507)

Notes: Presented as as means with standard deviation parentheses

Panel B: Pearson's χ^2 Test Results

	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
SS Middle	0.091	0.001	0.012	0.666	0.569
	(2.8519)	(12.0587)	(6.2573)	(0.1858)	(0.3244)
SS Upper		0.038	0.529	0.066	0.376
		(4.2955)	(0.3962)	(3.3887)	(0.7849)
CE F-Middle			0.098	0.001	0.007
			(2.7324)	(12.1467)	(7.2648)
CE F-Upper				0.011	0.131
				(6.4136)	(2.2768)
CE M-Middle					0.388
					(0.7438)

Notes: Presented as p-value with χ^2 score in parentheses

earlier results, boys are less risk averse and girls are more risk averse. Specifically, as we shift focus to middle school boys the average coefficient of relative risk aversion decrease by 0.06.

Turning to SS middle schoolers, CRRA coefficients increase by 0.127. However, the lack of significance in the competition column suggests these differences in risk attitude are not driving decisions to compete. This result is consistent with Wozniak et al. (2014) as well as Kuhn and Villeval (2015) who find controlling for risk attitude has little impact on female willingness to compete and cooperate, respectively.

Another measure of competitiveness, "residual competitiveness," has recently been used in the economics literature (Buser et al. 2014, further explored by Reuben et al. 2015). This measure is obtained by saving the residual from a (robust) linear probability model regression of the decision to compete on performance and CRRA risk attitude, two variables

Table II.9: Joint Estimation (Marginal Effects)

	ρ	Compete
SS Middle	0.127*	
	(0.0138)	
SS Upper	0.0105	
	(0.0149)	
CE, F-Middle	0.0174	
	(0.0171)	
CE, M-Middle	-0.0619*	
	(0.0152)	
CE, F-Upper	-0.00156	
	(0.0142)	
ρ		-0.133
		(0.493)
Constant	0.395*	-0.139
	(0.0113)	(0.212)
N	30,470	30,470

Notes: Standard errors in parentheses, * $p < 0.01$

considered to be implicit in the choice to compete[9]. Thus, higher (i.e. more positive) residuals are indicative of more competitive behavior. We report these means and compare them formally in table II.10. Clearly, SS middle schoolers are highly competitive with CE girls (all ages) being the least competitive. Specifically, students in the SS middle school are more competitive than upper school boys ($p = 0.01$) and upper school SS girls are as competitive as (that is not significantly less competitive than) middle school boys ($p = 0.47$) or upper school boys ($p = 0.32$).

Earlier, we considered extracurricular activities such as participation in sports. Given the nature of these activities, it is reasonable to think that they influence competitive behavior. This influence can come across as athletes competing differentially and/or subjects having preferences over competing against athletes. For these reasons, we specified hypotheses 5.a and 5.b—that competitive sports positively influence both risk taking and competitive behaviors. To test these, we first ran a probit regression of team sports participation and the

[9]Probit regressions have roughly the same deviance residual, but are considerably less intuitive to interpret.

Table II.10: Residual Competitiveness

Panel A: Summary Statistics					
SS Middle	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
0.117	-0.047	-0.292	-0.112	0.132	0.039
(0.5)	(0.493)	(0.38)	(0.478)	(0.491)	(0.508)

Notes: Presented as as means with standard deviation parentheses

Panel B:Rank Sum Test Results	SS Upper	CE F-Middle	CE F-Upper	CE M-Middle	CE M-Upper
SS Middle	1.77 (0.077)	3.4 (0.0001)	3.68 (0.0002)	2.24 (0.025)	2.39 (0.017)
SS Upper		2.32 (0.02)	1.47 (0.142)	0.079 (0.937)	0.446 (0.656)
CE F-Middle			-1.6 (0.11)	-2.43 (0.015)	-1.85 (0.063)
CE F-Upper				-1.08 (0.28)	-0.85 (0.396)
CE M-Middle					0.179 (0.858)

Notes: Presented as z-score with p-value in parentheses

proportion of subjects in the group who compete in a team sport on the subject's decision to compete. Table II.11, column (1) illustrates the marginal effects. We see that those who participate in a team sport compete significantly more. Specifically, if a subject plays a team sport, he or she is almost 17% more likely to compete. Looking over to column (2), which includes controls for test group (with upper school boys as a reference group) we see that those who participate in a team sport still compete about 17% more than others, and having more athletes in your group has a significant negative impact on the choice to compete. The implication here is athletes are more apt to compete, and subjects dislike competing against athletes[10]. However it does not appear that participation in sports influences risk attitudes. Thus we find support for Hypothesis 5.a, but not for 5.b.

Result 5 *Competitive sports favorably influences competitive behavior, but has no effect on*

[10]This interpretation is conditional on subjects knowing the sports their group members participate in, a safe assumption given our subject pool.

risk attitude.

Table II.11: Probit of Sports on Decision to Compete (Marginal Effects)

	(1)	(2)
Plays in a Team Sport	0.167**	0.176***
	(0.0660)	(0.0674)
Percent of Group in a Team Sport	-0.0637	-0.584**
	(0.157)	(0.286)
SS-Middle		0.214
		(0.143)
SS-Upper		-0.0514
		(0.117)
CE, F-Middle		-0.228
		(0.144)
CE, F-Upper		-0.181*
		(0.101)
CE, M-Middle		0.245*
		(0.146)

Note: N=277, Standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$

Earlier we noted that middle schoolers tend to be less risk-averse than the upper schoolers in general. This observation suggests that exposure to the school environment may influence risk attitude. As such, we examine the effect that school choice has on the behavior of young girls in hypothesis 6, positing that girls are similar across school types. However, we have shown that in middle school, SS girls are significantly more risk averse (Table II.4, $p = 0.00$), yet compete significantly more than their CE counterparts (Table II.8, $p = 0.00$), thus we reject hypothesis 6.

Result 6 *In young girls, both risk taking and competitive behavior is significantly different across school types.*

Table II.12: "Difference-in-Differences" Estimations

	(CRRA)	(Competed?)
Upper School	-0.0242*	0.179
	(0.0138)	(0.115)
Single Gender	0.0616***	0.398***
	(0.0132)	(0.110)
Upper*SS	0.0625***	-0.338**
	(0.0175)	(0.146)
Constant	0.403***	0.160*
	(0.0115)	(0.0958)
R-squared	0.416	0.072

Note: N=203, Standard errors in parentheses, $***p < 0.01, **p < 0.05, *p < 0.1$

The difference between both risk attitudes and competition decisions suggests we can consider the treatment effect of exposure to single sex schooling on girls. Hypothesis 7 asks whether these differences increase with said exposure. As a formal test of it, we restrict our attention to female subjects and calculate a "pseudo difference-in-differences" estimation:

$$y = \beta_1 + \delta_1(\text{Upper School} + \beta_2(SS) + \delta_2(\text{Upper*SS}) + \mu$$

where y can be specified as the decision to compete or the CRRA coefficient and δ_2 is the treatment effect of exposure to single sex education, conditional on selecting in. Table II.12 presents the results from this estimation. In both specifications, the coefficient on exposure, δ_2, is significant, yet the sign of the effect switches depending on whether we are measuring its effect on risk attitude or the decision to compete. Specifically, we measure that the coefficient of relative risk aversion increases by 0.06, while the decision to compete decreases by 34%.

However, taking seriously the simulations of Bertrand et al. (2004), we note this "treatment" is at the school level, and clustering is therefore infeasible. We instead refer to the above evidence as supporting the following observation:

Observation 1 *Conditional on selecting into it, the treatment effect of increased exposure to single sex schooling is that girls become more risk averse, and compete less.*

At first glance, this observation may seem bizarre, but it is actually quite profound, and underscores the fact that the various differences between girls are coming through at the middle school grade levels[11].

Table II.13: Female Competition in Groups of Majority (> 50%) Girls

Panel A: Summary Statistics		
	Decision to Compete	Residual Competitiveness
	0.428	-0.021
	(0.502)	(0.505)

Note: All groups were CE, Upper, Presented as as means with standard deviation parentheses

Panel B: Rank Sum Test Results		
	Decision to Compete	Residual Competitiveness
vs. SG, Middle	1.269	2.225
	(0.204)	(0.026)
vs. SG, Upper	-0.256	0.432
	(0.798)	(0.666)

Notes: Presented as z-score with p-value in parentheses

Bearing the above in mind, there is a shortcoming that needs to be addressed. In sessions conducted at the co-ed school, we did not have any all-female groups. However, there is significant variability in the CE group composition (see Figure II.2). This variability led us to specify hypothesis 8, which relates competitive behavior to the gender composition of the subject group. This suggests that the SS preference to compete may actually be a preference to compete against girls (or reluctance to compete against boys)[12]. In our sample, 35 CE (upper school) girls had groups of at least 50% girls. We leverage this variability in group size and composition, and conduct Mann-Whitney U tests on this sub-sample to test hypothesis 8. An argument in favor of a female preference to compete against girls arises when we restrict the analysis to this sub-sample and present these findings in Table II.13. It shows that CE girls in these heavily female groups now opt to compete 43% of the time—over 50% more than CE girls on average. Furthermore, they are no longer significantly different from

[11]One seminar participant went so far as to suggest that SG schools may have a homogenizing effect on young women.

[12]There are other possibilities, for instance opting to (not) compete could suggest a preference (aversion) to feedback.

SS girls in either middle school ($p = 0.204$) or upper school ($p = 0.798$). Thus we fail to reject the null of hypothesis 8, which may suggest that group gender composition is driving observation 1. That is, with more girls in a CE group, the more similar it looks to an SS group, or in other words:

Observation 2 *Tournament entry rates significantly increase with the number of girls in a group.*

II.6 Concluding Remarks

In a November 2014 article of the New York Times, a third grade math teacher commented that she turns math lessons into games because her male students enjoy competition (Rich, 2014). Research has demonstrated that this sort of preference has implications for both educational and later-life outcomes. Recently, policy makers have proposed increases in single sex schooling as a way to increase equity in education and decrease the gender gap that may result from differences in these preferences. However, these proposals lack a theoretical understanding of why these differences exist and may lead to unintended consequences.

In this paper, we run an experiment to disentangle competing theories in the context of single sex education. While much of the evidence on the subject suffers from the question of who is selecting into these schools, we attempt to mitigate these sample selection issues. We conduct our experiment in two closely matched schools and focus our results on the structural differences that exist conditional on selecting into a single gendered environment.

Our experimental results suggest that indeed, girls educated in coed environments do shy away from competition. However, girls schooled in single sex environments compete just as much as their male, coed counterparts. Additionally, they elect to compete despite an apparent risk aversion across financial gains.

These results speak to the nature-versus-nurture debate and lend credence to the conclusion that nurture matters when it comes to making competitive choices. Furthermore, single sex schooling appears to be nurturing in such a way that the competition is rational—that is, it is not being guided by a preference for risk, but rather an appropriate calculation of that risk.

These findings are relevant to recent policy proposals and controversies regarding gender segregated education. They demonstrate several effects that single sex schooling can have on economically relevant decisions. Our paper has expanded this line of research in both experimental design and estimation techniques. However, we conclude by stressing the need for further research in this area.

Chapter III

The Richness of Giving:

Charity Selection and Charitable Gifts in a Large Field Experiment

III.1 Introduction

In 1772, Alexander Hamilton famously arrived in New York City from St. Croix as the result of charitable donations from the local populace. Over 200 years later, benevolent giving to charity now hovers at slightly above 2% of US GDP, and approximately 76% of Americans contribute to at least one charitable organization each month (World Giving Index, 2013). However, we still know very little about the *decision* to donate to charity, the motivations behind it, and in particular, the mechanism behind one's charitable choice.

In this paper we examine this mechanism by means of an economic experiment. Charitable giving in many lab and field experiments (artefactual or otherwise), has previously been modeled as a 2-dimensional dictator game. That is, a game played between oneself and a charity. However, to the best of our knowledge, no one has yet examined the mechanism behind one's choice of said charity or operationalized it in these experimental environments. In these giving experiments the charities to which a subject can donate tend to be either anonymous, presented without a choice, or only selected through what we call the restricted choice method (such as from an *ad-hoc* list of ten charities). Accordingly, in this paper we run an experiment where the unsure components are not only how much the subjects chose to give, but also to which charities they chose to give. In this way, our experiment examines this choice mechanism, and uses it to disentangle motives behind giving.

These charitable and philanthropic motives were first discussed by Becker (1974). In this essay he outlines the following three rationales: improvement of the general well being, receiving social acclaim, and avoidance of scorn. In the years hence, we have come to think of these motives as efficiency concerns, warm glow, and social pressures, respectively.

Whereas the formalization of social pressures remains relatively novel (DellaVigna et al. 2012; Lazear et al. 2012), the concepts of efficiency and warm glow, sometimes thought of a pure and impure altruism, have long been discussed. In an effort to improve model predictions and comment on the nature of individual donations, Andreoni (1989) introduces impure altruism, and then (1990) formalizes a theory of warm-glow giving, where a person exhibiting warm glow can have both altruistic and egoistic motives for giving. Later papers have suggested alternate reasons for this warm-glow, such as signaling benefits of giving to others (Bénabou and Tirole 2006, Ellingsen and Johannesson 2008).

Taken in the context of our model, we can think of pure altruism (as opposed to warm-glow) as dependent on the relative price of giving. One notable implication of here, is that pure altruism is only concerned with the amount of public good ultimately provided. As such, giving can be fully crowded out by what other individuals, firms, or governments provide. In a broad overview of the subject, Vesterlund (2006) further discusses these motives for giving, citing observational and experimental evidence in the context of a classical demand setting.

Distinguishing between these motives is important to both further our economic understanding of giving and, subsequently, inform optimal policies. Given these *pro forma* descriptions, economic experiments have previously tried to disentangle motives for giving by focusing on design and design choice. In particular, experiments have tended to use either linear public goods games (e.g. Andreoni 1993; Goeree et al. 2002) or dictator games, similar to the modified one we use in this paper[1].

In the first paper to introduce real charities as a dictator recipient, Eckel and Grossman (1996) vary receipts between anonymous subjects and the American Red Cross. They find that subjects are more willing to donate to charity and conclude (some form of) altruism is a motivating factor. To disentangle altruistic motives, Crumpler and Grossman (2008) extend this design, with the stipulation that the experimenter will fully crowd out any amount donated to charity. As such, any subject still opting to donate is taken as an incidence of warm glow.

As a response to Crumpler and Grossman, Tonin and Vlassopoulos (2014) use multiple to decisions to identify warm glow, noting that otherwise the observed effect could in fact be purely altruistic feelings towards the experimenter. In doing so, they are able to place bounds on the magnitudes of various motives for giving. However, these quantities are not parameter estimates, and the authors note that their results may not generalize to other settings.

As such, we design our experiment to create a richness of data and subject pool. This allows us to impose a standard utility (CES) framework and estimate a structural model to fully quantify (rather than rule out) motives for giving.

Though this emphasis on structure is relatively novel in the literature its importance has been clearly stated. Recent papers (e.g. Huck et al. 2015) have used this methodology to focus on the fundraising side of philanthropy. However, DellaVigna et al. (2012) stress the importance of these approaches to estimate *individual* motives for giving. We extend this line of research to examine these motives inframarginally.

Further, Vesterlund (2006) cautions against interpreting aggregate elasticities when not all contributors experience the same changes in marginal tax rate. Rather, to better inform tax policy, individual elasticity estimates should be used. Chay et al. (2005) voice a similar concern that individual elasticities are necessary for unbiased welfare estimates in the market

[1]For this reason I will focus on dictator games in this brief review.

for clean air.

Taking these concerns into account, we set up our experiment such that the data act as a lens to examine charitable choices, i.e., the choices one undertakes when deciding where to give, whether to give, and how much to give to charity. Previous dictator designs have tended to not vary the price of giving or subject income in the experiment. The novelty of our design, as well as the richness of our data and that fact that we observe it at the individual-level, allows us to impose a structure (discussed further below) that we use to recover the deep parameters of these preferences for giving. In this note we examine rationality of these choices and decompose our parameter estimates across demographics and the types of selected charities.

We also take seriously the advice of DellaVigna et al. (2013) who note the inherent difficulties of predicting which causes individuals will give to. We ask experimental subjects to choose their most preferred charity from a much richer set of organizations. In doing so we are also able to gather data that inform us of charity preference. Though non-causal, we correlate these data with subject demographics and estimated parameters. We ask if there are interesting co-movements between distributional preferences and various these chosen charities. If so, do these preference parameters for charitable giving differ from what we have observed in the past with person-to-person gifts? Finally, we ask if we can exploit this information and the in-subject heterogeneity to increase charitable giving in the naturally-occurring world.

As such, our paper speaks to the literature in both public and behavioral economics. Our contribution to these literatures is our development and estimation of the structural model in the context of a richer set of both givers and recipient organizations. In so doing, we estimate and comment on full parameter distributions, rather than simple sample averages.

This paper proceeds as follows: Section 2 presents our contextual framework, experimental design, and data. Section 3 discusses the consistency of choices and our structural estimates. Section 4 looks at the mechanism of charitable choice and selected charity. A final section concludes.

III.2 Experimental Design

III.2.1 The ALP

For the purposes of structurally measuring distributional preferences in our experiment, we required a large and diverse sample of subjects that differs from the standard undergraduate population. In this sense, we classify our work as an *artefactual field experiment* to use the terminology of Harrison and List (2004). Accordingly, we chose to embed our software in the RAND Corporation's American Life Panel (ALP henceforth) and conduct and incentivized experiment. The ALP is a 6,000 member, U.S.-based, Internet panel. Its unique interface allows researchers to conduct sophisticated experiments matched with individually rich demographic data. Panelists are recruited through a number of different ways, including randomized recruitment, providing us with a representative sample, as well as HRS-style demographic data[2].

The composition of our subject pool is described in table III.1, which has demographic information on those who completed the experiment, those who started but did not finish, as well as comparative US population data from the American Community Survey[3].

Subjects in our experiment hail from every state except Alaska. They range in age from 22 to 92. Our sample is 55% female and slightly less than 80% white. 45% of our subjects hold a college degree and the employment figures are roughly similar to the US population as a whole.

[2]For more information on the ALP, please visit: https://mmicdata.rand.org/alp/.

[3]For the purposes of GARP and CES analysis, we consider a "complete" experiment as completing 45 or more dictator allocations.

Table III.1: Subject Demographic Averages

Variable	Completed	Started	US Adults
Female	0.55	0.566	0.58
Age (Median)	57	57	37.4
18-44	0.229	0.217	0.599
65+	0.269	0.29	0.137
Caucasian	0.799	0.783	0.763
African American	0.1	0.11	0.137
Native American	0.012	0.016	0.017
Asian or Pacific Islander	0.022	0.022	0.063
Hispanic or Latino	0.169	0.165	0.169
HS Diploma	0.955	0.955	0.862
College	0.448	0.446	0.267
Employed	0.553	0.54	0.577
Unemployed	0.059	0.051	0.058
Not in the Labor Force	0.389	0.409	0.361

Note: US Population Data come from the American Community Survey
http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

III.2.2 Charity Navigator

While there is vast anecdotal evidence that charitable givers prefer to give locally, empirical evidence on the subject remains mixed. Furthermore, local (as well as national) causes can be highly idiosyncratic in nature. As such, we required a large and diverse set of charities in addition to the diversity of our subject pool.

To develop this set we turned to the website Charity Navigator, an agency used to rate charitable organizations[4]. Founded in 2001, Charity Navigator rates over 7000 charities and is the largest and most used rating agency in the United States. Using information from IRS form 990 returns, Charity Navigator assigns each charity a star rating on the bases of efficiency (financial health), and accountability (transparency)[5]. Empirical evidence suggests a positive impact of these star ratings on charitable behavior (Gordon et al. 2009; Brown et al. 2014).

The Charity Navigator website is organized as follows. Each rated charity is placed into

[4]http://www.charitynavigator.com.

[5]For a more detailed yet concise description of this star rating methodology, please refer to Table I in Gordon et al. (2009).

one of nine categories, and each category has as associated list of (between 2 and 6) unique causes. For instance, "Patient and Family Support" is a cause in the "Health" category. At the time of this writing, the top rated charity in that cause is Camp John Marc.

For the purposes of this experiment, we scraped the Charity Navigator website for the top 10 ranked charities within each cause, providing us with a list of 340 charities in total. The complete list (with associated categories and causes) can be found in appendix G.

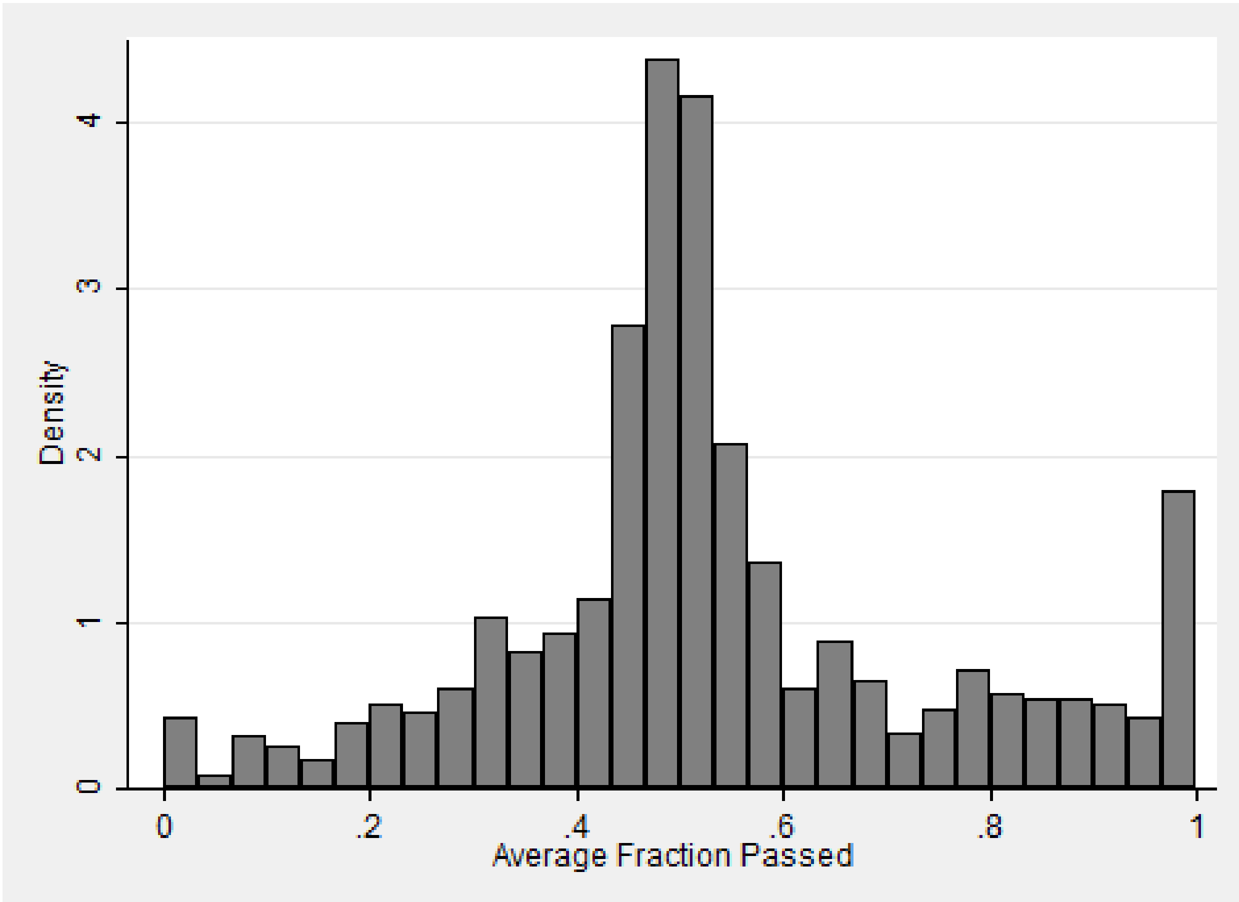
III.2.3 Experimental Procedures

As such, subjects in this experiment were first obliged to select their preferred charity from this list (that is, the charity to which they wish to donate). Immediately after, each subject makes a series of allocation decisions reflective of a charitable contribution. These decisions each consist of a modified dictator game played between the subject and his or her chosen charity. In the standard dictator game, an active player (the dictator) is given an endowment of wealth, w, and divides it between herself and a passive player such that total payoffs are given by $\pi_d + \pi_o = w$. This stipulation necessarily restricts the slope of the budget line to -1.

We differ from this standard dictator framework in two crucial ways. Firstly, the passive player is a charity of the dictator's own choosing. Secondly, we follow the generalization of the dictator game developed by Andreoni and Miller (2002). Budget sets are still *linear* in this game form, but the *price* of giving (donating) varies such that total payoffs are now represented as $x_i + p_g g = w$ where x_i is the payoff to oneself and p_g is the *relative price of donating to charity.* The purpose of this variation is twofold. First, we answer the seminal question: can this giving be consistent with individually rational behavior? That is, are patterns of giving are consistent with the axioms of revealed preference? Secondly, the price and income variation over several decisions allows use to trace out each subject's preferences for giving.

When subjects first log onto the experimental interface, they are given a instructions

Figure III.1: Average Fraction Donated



and a broad overview of the experiment. We explain to them that they will be selecting a charity of their choice and then given opportunities to allocate between themselves and that charity. To select a charity, subjects are taken through a set of expandable and collapsible tables adapted from the Charity Navigator website. The tables consist of populated lists. A subject is first instructed to pick a category. Each choice of category feeds into its associated list of causes. The selection of a given cause expands the list to the top 10 highest rated charities (as determined by the Charity Navigator website) for that particular cause. From there the subject is instructed to pick his or her preferred charity. If the subject so desires, or if no charity is to his or her liking, a charity write-in option is provided.

After selecting a charity, each session consists of 50 independent dictator problems. The number number of choices allows us to generate the rich data needed for individual level statistical inference[6]. In each round, the subject is asked to allocate tokens between two accounts: the personal account of the participant (henceforth *self*), and the account of the chosen charity (henceforth *charity*). Each decision problem starts by having the computer select a budget line randomly from the set of lines that intersect at least one axis at or above the 50 token level and intersect both axes at or below the 100 token levels. The budget lines selected for each subject in his decision problems are independent of each other and of the budget lines selected for other subjects in their decision problems.

Next, the subject chooses an allocation along the budget line. To choose an allocation, subjects use the mouse or the arrows on the keyboard to move the pointer on the computer screen to the desired allocation. This *point-and-click* design is adapted from Fisman et al. (2007). The benefits to using this software are manifold. Among them, the ability to represent consumer choice problems graphically is simple and easy for subjects to understand, and the choice environment allows for the generalization of individual preferences. Additionally, the design can be easily adapted to many other kinds of individual choice problems. A subject's view of the computer program dialog window is shown in the attached experimental instructions located in appendix H.

The payoff for each decision round is determined by the number of tokens each account (*self* and *charity*). At the end of the experiment, the computer randomly selects one decision round for payment for each participant. The subject is then paid the amount earned in that round using to the conversion rate 2 tokens = 1 dollar.

The average subject passed 54% of her tokens. Figure III.1 shows this distribution of percent of tokens donated to charity at the subject level.

[6]This number reflects the Bronars calibration in Choi et al. 2007 who show that the distribution of consistency values is skewed to the left as the number of budget sets increases.

III.3 Results

Broadly speaking, motives for giving tend to reflect either the efficiency or "warm-glow" concerns discussed above. We can think of the parameters for these concerns as boiling down to whether or not the act of giving itself is independent of the price of giving. Accordingly, in this section we examine the parameters within this analytical framework.

III.3.1 Rationality

However, before we can examine these equity/efficiency tradeoffs, we must first confirm that the data generated by each experimental subject are governed by the principles of utility maximization. Given that each subject has preferences over herself, x_i and the charity, g, we want to know whether observed choices can be expressed via a utility function $U_i = u_i(x_i, g)$, and subsequently, maximize said function. That is, are the data consistent with the *Generalized Axiom of Revealed Preference* or GARP. The practice of seeing if the data satisfy GARP follows Varian (1982)[7].

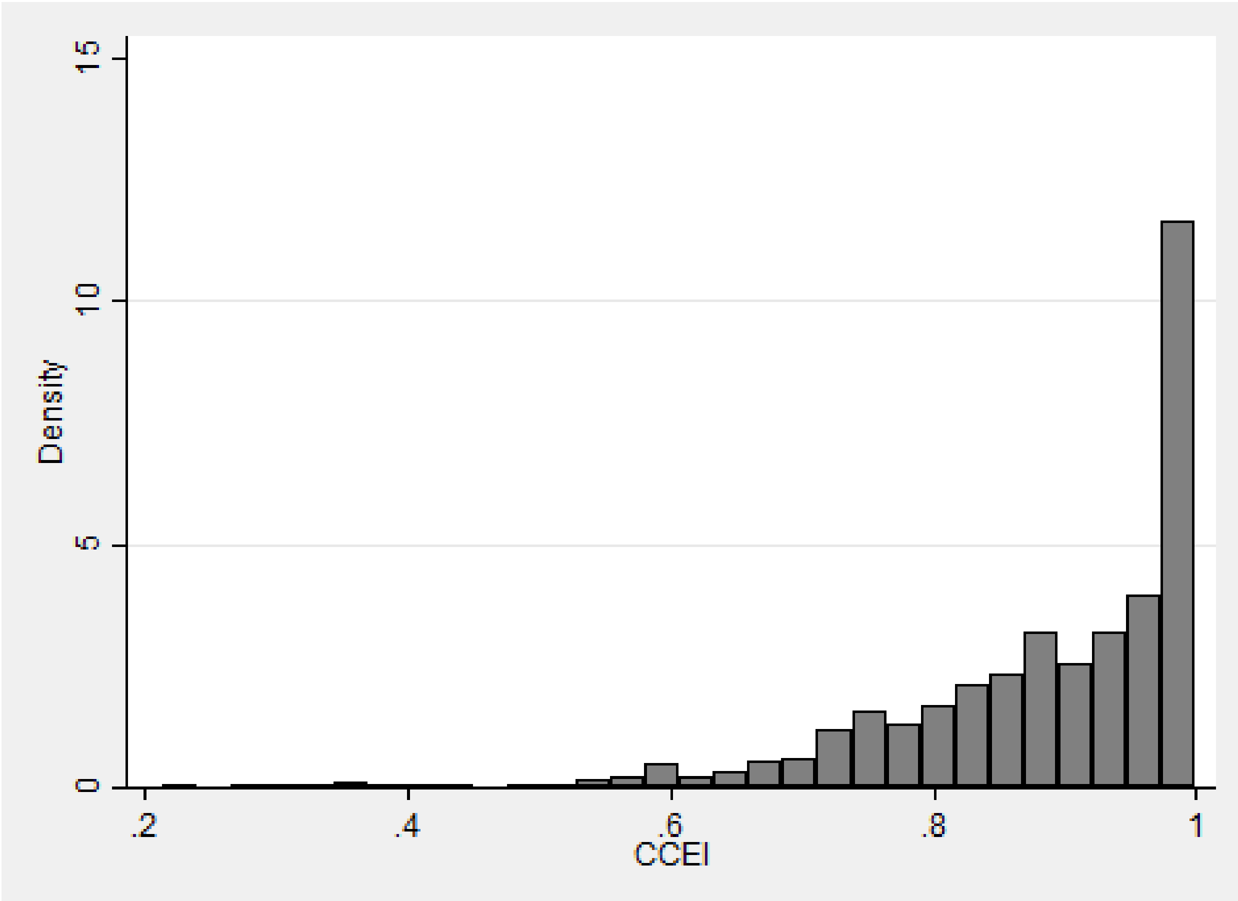
In order to see if, how, and to what extent the data comply with GARP we use Afriat's (1972) *Critical Cost Efficiency Index* (CCEI) as follows:

> Define a generalization of the revealed preference relation $R^D(e^t)$ such that $x^t R^D(e^t) x$ iff $e^t p^t x^t \geq p^t x$, that is, x would not be affordable at a fraction e^t of the income available when the person chose x^t. Define $R(e^t)$ as the transitive closure of $R^D(e^t)$. Then define GARP(e^t) as "if $x^t R(e^t) x^s$, then $e^t p^s x^s \leq p^s x^t$." Then the CCEI is the highest value of e^t such that there are no violations of GARP(e^t). (Andreoni and Miller 2002)

In short, the CCEI measures the extent to which budget constraints need to be relaxed in order for violations to no longer violate. By construct, it is bounded between 0 and 1,

[7]Note that we restrict our analysis to *Walrasian* budget sets and therefore implicitly treat preferences as *well-behaved*.

Figure III.2: Distribution of Aftriat's CCEI



with scores closer to 1 being closer to satisfying GARP. That is, scores closer to 1 are "more rational". Among those who completed the experiment, the distribution of CCEI scores are illustrated as a histogram in figure III.2.

In our data subjects exhibit significant rationality. Over 75% of the data have a CCEI score greater than 0.8. The mean CCEI score is 0.883, and the modal score in the histogram's highest bin[8].

Given this consistent behavior on the part of our panel it is appropriate to think of charitable giving as a standard utility maximizing activity. Further, we have generated enough data to estimate these standard parameters at the individual level. As such, we now test the structural properties of each subject's individual utility function.

[8]i.e., the bin closest to (and inclusive of) 1.

III.3.2 Structural Model

Similar past experiments (e.g. Fisman et al., 2015) have shown that subjects exhibit remarkable heterogeneity. Given these demonstrations and our observed pattern of CCEI scores sufficiently close to 1, we hereby treat the data as generated by a well-behaved utility function, u_d.

Further, to maintain consistency with the previous literature, we assume u_d is both separable and homothetic. These two assumptions taken in concert with the restriction imposed by our design that choices must be budget balanced imply that u_d is of the family of CES utility functions (equation III.1). In fact, we estimate these CES utility functions using non-linear tobit MLE (discussed in turn below), and find that this subject heterogeneity is more pronounced *within* demographics and charity types, rather than *across* them.

$$u_d = [\alpha \pi_d^{\rho} + (1-\alpha)\pi_o^{\rho}]^{\frac{1}{\rho}} \tag{III.1}$$

This utility function only has 2 components. Since people are behaving consistently (maximizing) we choose to keep the model parsimonious and do not include additional parameters. However, given that some subjects are inconsistent, future research would perhaps benefit from imposition of additional parameters or more flexible functional forms.

Within the context of a CES framework, these two parameters of interest are ρ and α. Previously, these parameters have been interpreted in terms of person-to-person giving (see e.g., Fisman et al., 2007). However, they are also immensely important in terms of charitable giving, and can easily generalize to other settings as well.

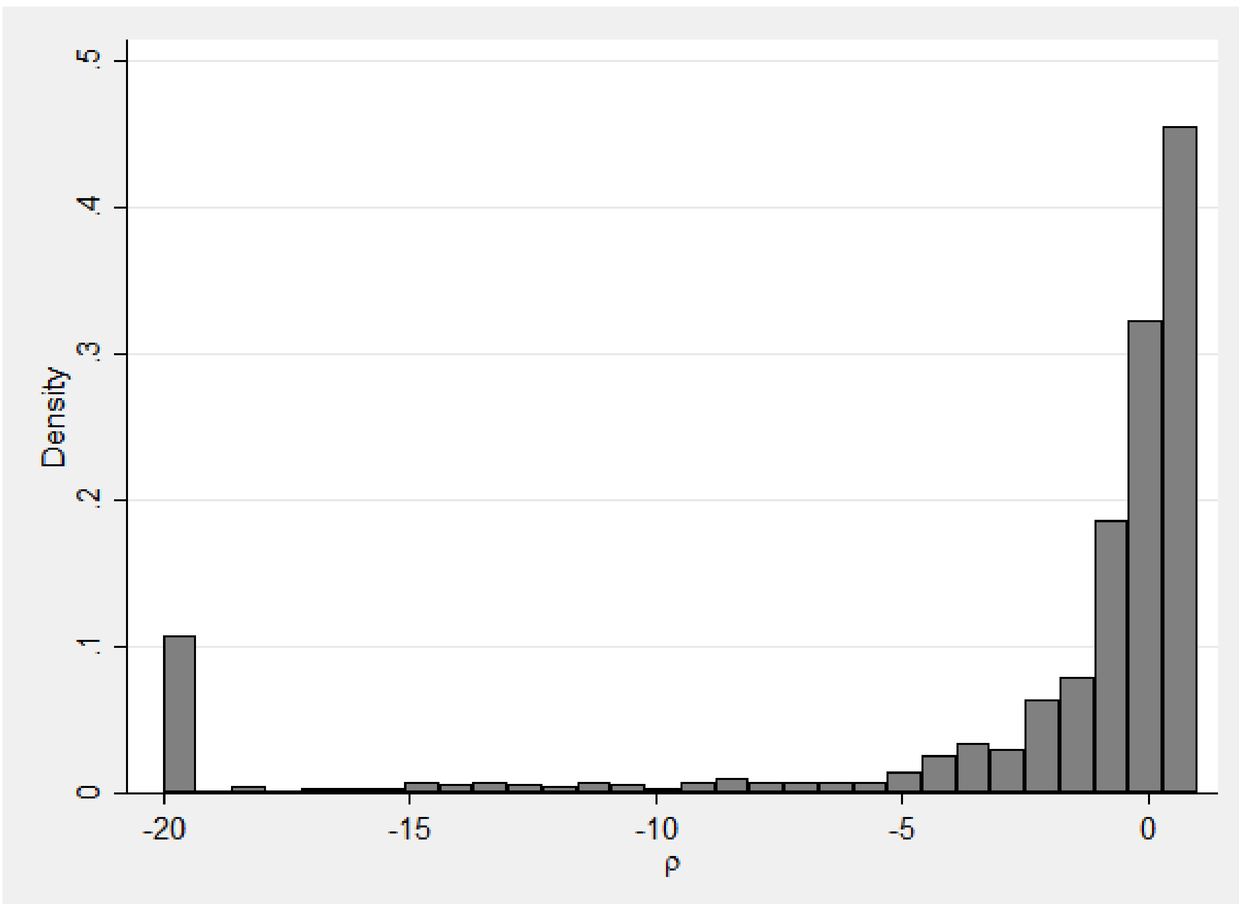
The parameter ρ represents the curvature of each individual's indifference curves, and thereby her sensitivity to price[9]. Further, α represents the relative weight one puts on *self* as opposed *charity*. As such, we interpret α as the warm-glow parameter.

For any ρ, when $\alpha = \frac{1}{2}$ the subject equally weights the payoff to herself and her selected

[9]It follows that $\frac{1}{\rho-1} = \sigma$ is the (constant) elasticity of substitution.

charity. Meanwhile, when $\rho > 0$, charitable preferences are weighted towards efficiency in the sense of increasing total payoffs or the "size of the pie". When $\rho < 0$, preferences are weighted towards equity–reducing differences in payoffs or the "slices of the pie"[10]. For these reasons, we extend previous interpretations of ρ and tie it in closely with the concepts of "pure" altruism in previous giving literatures and "efficiency" in the public goods literature.

Figure III.3: Distribution of Estimated Rho (ρ)

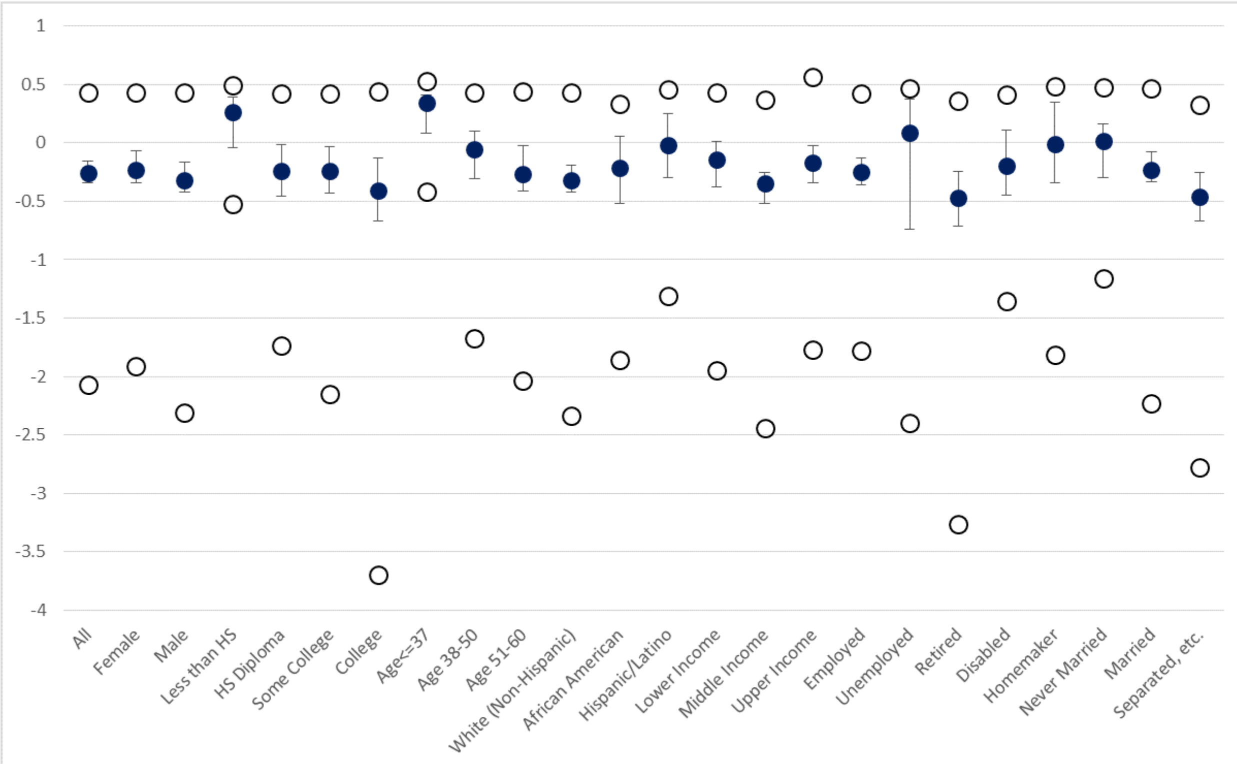


Rho (ρ)

The mean estimated ρ is -2.70 and the median is -0.259. However, the distribution of estimated rho parameters (depicted in figure III.3) is perhaps more informative. The distribution is noteworthy in that is highly skewed left with a large spike at $\rho \approx -20$. Most

[10]i.e. the elasticity of substitution is < -1.

Figure III.4: Estimated Median Rho (ρ) Values by Subgroup



Dots indicate median values, circles indicate 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles, bars indicate 95% confidence intervals for medians

subjects (58%) have a $\rho < 0$, while 441 (42%) have $\rho > 0$. In figure III.4 we decompose the ρ distributions with socio-demographic data collected by the ALP. While the medians differ by subgroup, overall, the distributions appear similar.

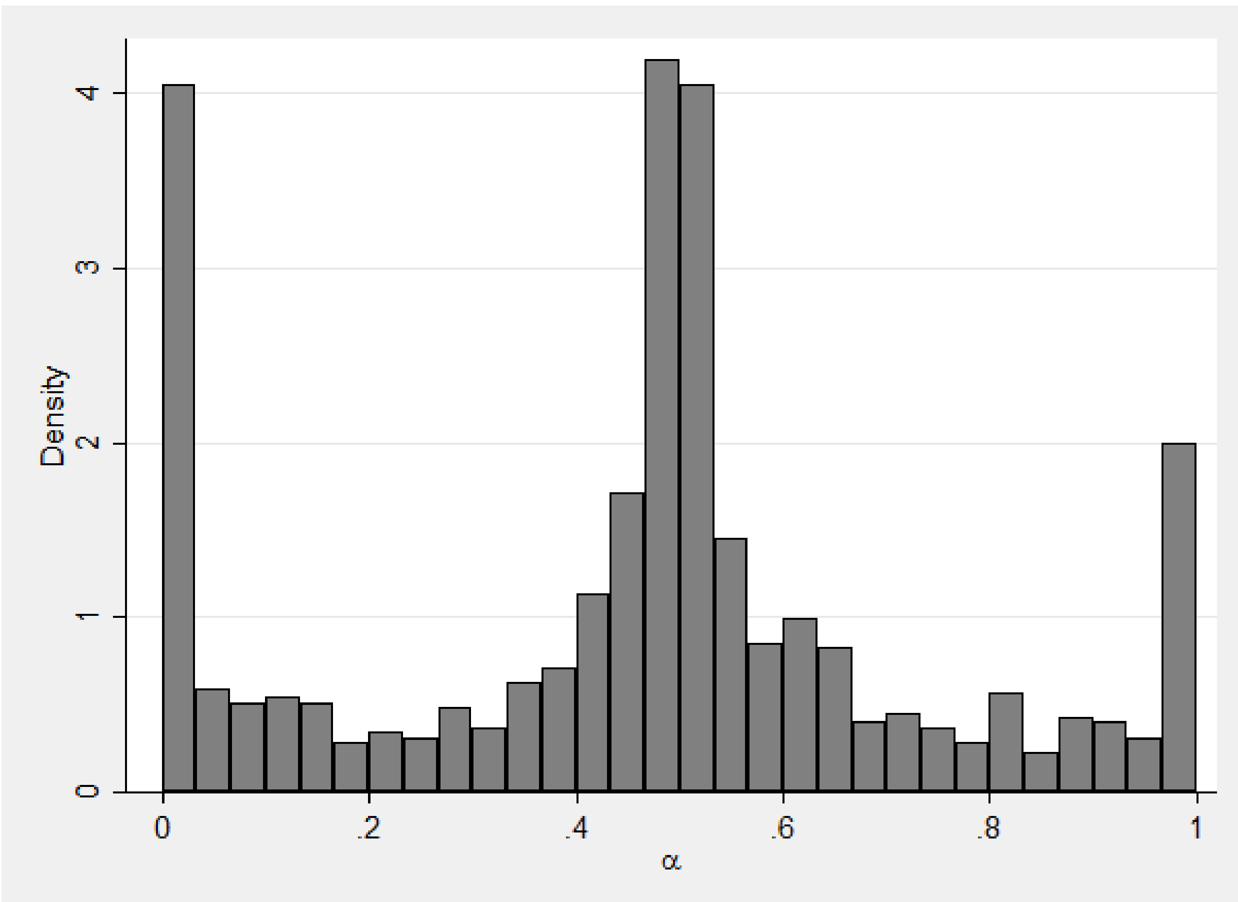
We examine this further with pairwise distributional (Kolmogorov-Smirnov) tests in table III.2. In these tests, the comparison group is all subjects not in the specified subgroup. For instance, the comparison group for men is women, for a college degree it is everyone with less than a college degree. P-values for the ρ distributions are shown in the 3$^{\text{rd}}$ column of the table. While some subgroups (7 out of 22) exhibit statistical difference in their distributions, this is likely an artifact of multiple hypothesis testing. Further, there are *fewer* differences in distribution, then there are differences in means.

Alpha (α)

Similar to ρ we interpret α in a slightly different fashion than previous work. Whereas similar person-to-person giving functions have interpreted α as fair-mindedness, in our context (charitable giving and contribution to a public good) we find "warm-glow" to be a more appropriate interpretation.

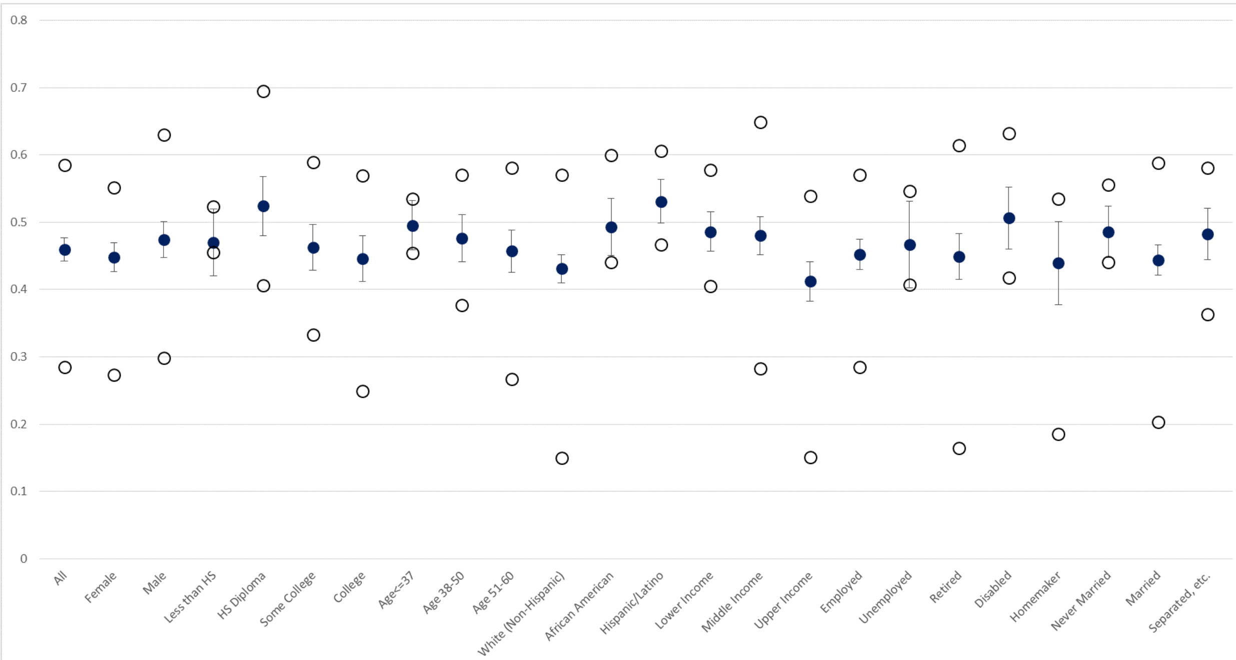
The distribution of estimated alpha parameters is depicted in figure III.5.

Figure III.5: Distribution of Estimated Alpha (α)



Again, the full distribution appears to be more informative than the summary stats. The mean estimated α is 0.46, and the median is 0.49, both of which are in line with Fisman et al.'s (2007) interpretative definition of "fair-mindedness". Interestingly however, the distribution appears to be tri-modal with large spikes at 0, 0.5, and 1. Overall, 366 subjects (34.7%) have an α between 0.45 and 0.55, while just 75 (7%) have $\alpha > 0.95$. Again, we decompose the α distributions with socio-demographic data in figure III.6. The results are striking in that the means are very close to 0.5 with tight confidence bounds.

Figure III.6: Estimated Mean Alpha (α) by Subgroup



Dots indicate mean values, circles indicate 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles, bars indicate 95% confidence intervals for means

P-values for the distributional tests on α are shown in table III.2, column 5. The tests were calculated the same way as above. As above, there are *fewer* differences in distribution then there are differences in means. Further, 10 out of the 22 subgroups exhibit statistical significance, although again, no corrections were made for multiple testing.

III.4 Charity Selection

Since Eckel and Grossman (1996), standard experimental practice has been to give to actual charities. However, within these experiments there is often not enough data to comment on charity preference. Several studies offer no choice at all (e.g. Davis et al. 2005). Furthermore, in studies that offer no choice, several conceal the charitable organization. Harrison and Johnson did not identify their charity (the ACLU of South Carolina) in their

Table III.2: Pairwise Kolmogorov-Smirnov Tests by Subgroup

	Rho (ρ)		**Alpha (α)**	
Variable	**D**	**P-Value**	**D**	**P-Value**
Male	0.0399	0.778	0.986	0.013**
Less than HS	0.2307	0.014**	0.1833	0.086
HS Diploma	0.0669	0.604	0.1515	0.006***
Some College	0.0566	0.538	0.0549	0.577
College	0.0849	0.112	0.0607	0.448
Age<=37	0.2504	0.00***	-0.1753	0.001***
Age 38-50	0.0945	0.092	0.0854	0.161
Age 51-60	0.048	0.701	0.0343	0.957
White (Non-Hispanic)	0.0823	0.091	0.2195	0.00***
African American	0.0923	0.366	0.1308	0.069
Hispanic/Latino	0.1144	0.037**	0.2161	0.00***
Lower Income	0.0699	0.239	0.1073	0.014**
Middle Income	0.0953	0.019**	0.0963	0.017**
Upper Income	0.1439	0.00***	0.1344	0.00***
Employed	0.0646	0.228	0.0691	0.167
Unemployed	0.1317	0.241	0.1042	0.516
Retired	0.1039	0.018**	0.0749	0.182
Disabled	0.0915	0.322	0.1565	0.01***
Homemaker	0.1245	0.183	0.0864	0.606
Never Married	0.1082	0.067	0.1367	0.009***
Married	0.0836	0.056	0.0783	0.086
Separated, etc.	0.1148	0.009***	0.0606	0.428

Note: Comparison group is all subjects not in the specified subgroup, *** p<0.01, ** p<0.05

experiment, but afforded subjects the opportunity to examine the check written. Similarly, Buchheit and Parsons (2006) disguise the name of the charitable organization. Furthermore, in papers with choices, the choice at hand is necessarily reflective of the actual choice one makes in charitable giving. For instance, donors inMcDowell et al. (2013) are only given a choice of two charities. In what has become something of a standard, Eckel and Grossman (2003) provide a list of ten charities to choose from. They note that:

> The charities were selected to *reflect as broad a range of services and client groups as possible*. The sample included *international* charities (African Christian Relief, Doctors Without Borders USA, and Feed The Children); *national* charities (I Have A Dream Foundation); and *local* organizations (Women's Haven of Tar-

rant County and American Red Cross, Tarrant County Chapter). The charities covered *health* (AIDS Outreach Center and Cancer Care Services); *environmental* (Earth Share Texas); and *social service* charities (YMCA of Arlington). Charities were selected from the Texas State Employee Charitable Campaign booklet for 1997, which was provided to state employees during the workplace charity campaign. All charities included in the booklet meet state tax eligibility standards. A brief description of each charity was given to the subjects, taken verbatim from the Texas State Employee Charitable Campaign booklet.

(Eckel and Grossman 2003)

Italics are my own. Other charitable giving papers following this procedure include Grossman et al. (2012), Tonin and Vlassopoulos (2013), and Harrison and Phillips (2013).

We account for these previous restrictions of choice by allowing subjects to not only select from a markedly broader list of charities, but also by allowing them the option to switch their charitable choice once they know the rules of the game.

Table III.3: Selected Categories and Causes

Category	**N**	**Top Cause**	**N**
Animals	328	Rights,Welfare, & Services	272
Arts, Culture, & Humanities	42	Performing Arts	19
Education	131	Other Programs & Services	101
Environment	65	Protection & Conservation	59
Health	298	Medical Research	109
Human Services	307	Food Banks & Distribution	97
International	31	Development & Relief Services	13
Public Benefit	26	Advocacy & Civil Rights	12
Religion	93	Media & Broadcasting	29

Given our diverse subject pool, it follows that subjects have a wide range of charitable concerns. Table III.3 gives information on the distribution of selected types and the most popular cause within each type. The three most popular charitable categories are Animals, Human Services, and Health, respectively. Three divisions which by conventional wisdom, rely on very different donor types[11].

Figure III.7: Word Cloud of Selected Charities

Excluded Words: *Association, for, foundation, in, institute, of, the, to*

Of course, the universe of selected charities within these categories presents even finer and more diverse bins. As a means of description, figure III.7 displays a word cloud comprised of the names of every selected charity in our dataset. Clearly the most prominent words are related to our top categories (e.g. *Rescue, Research and Cancer*), but an interesting observation is also the high number of geographic indicators in the cloud (e.g. *Southeastern, Greenville, Downtown*).

[11] We have more data on charitable selection than structural estimates as most subjects classified as not completing the experiment still selected a charity (*N=1321*).

III.4.1 Rationality Across Type

Above, we noted that subjects exhibit rational, utility maximizing behavior. We build on this analysis further by asking the follow-up question: does this rationality differ across charity type? By way of example, we ask whether a subject who prefers to donate to charities supporting education is more consistent in those charitable choices (as indicated by CCEI score) than one who donates to religious causes? In this way, we follow the framework of Choi et al. (2014).

Figure III.8: Distributions of CCEI Scores

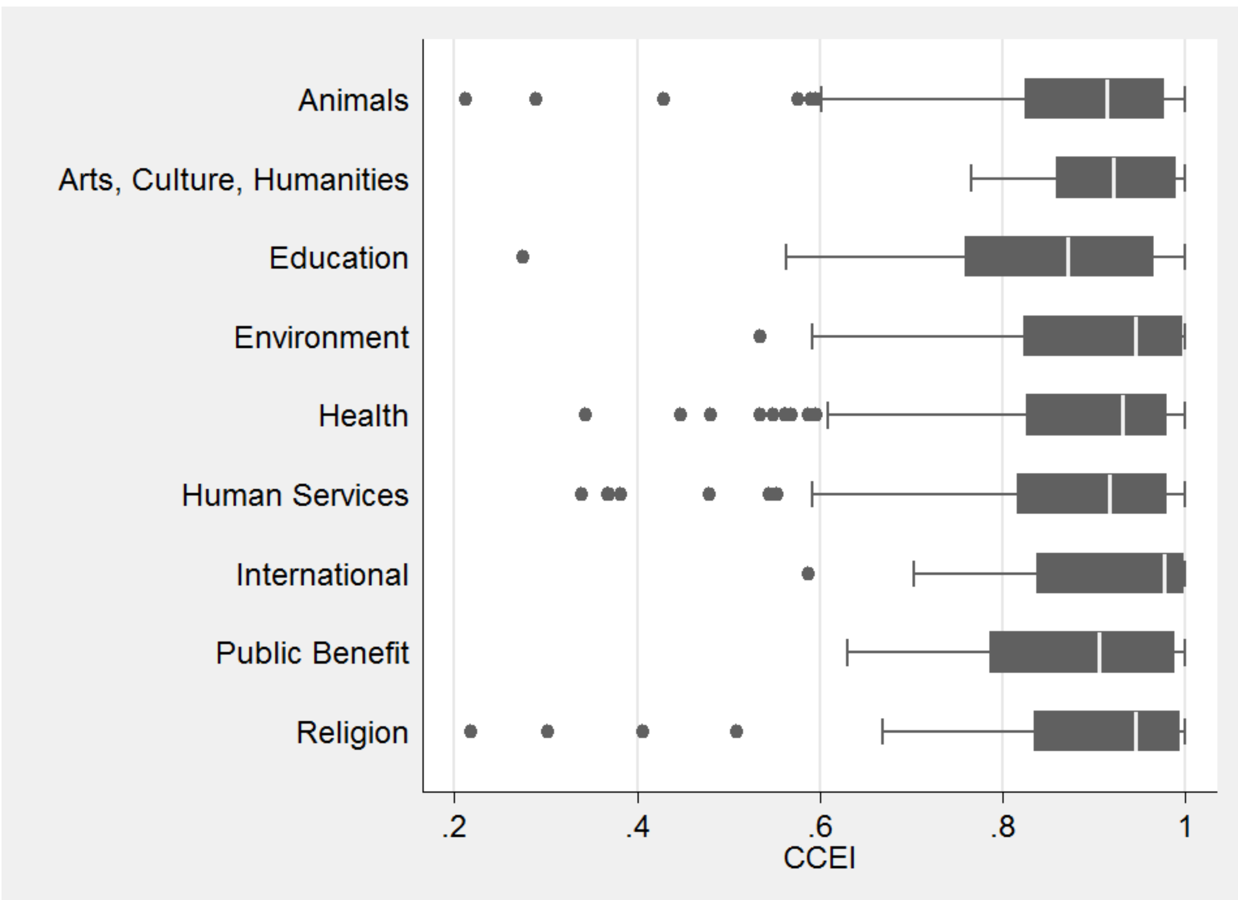


Figure III.8 illustrates the distributions of these scores across all charitable categories. First we note that every type of charity has a mean CCEI greater than 0.84. Further, while that leaves the door open for some slight variation in means, the distributions of rationality are all very similar.

Table III.4: Correlation between CCEI and Selected Category (OLS)

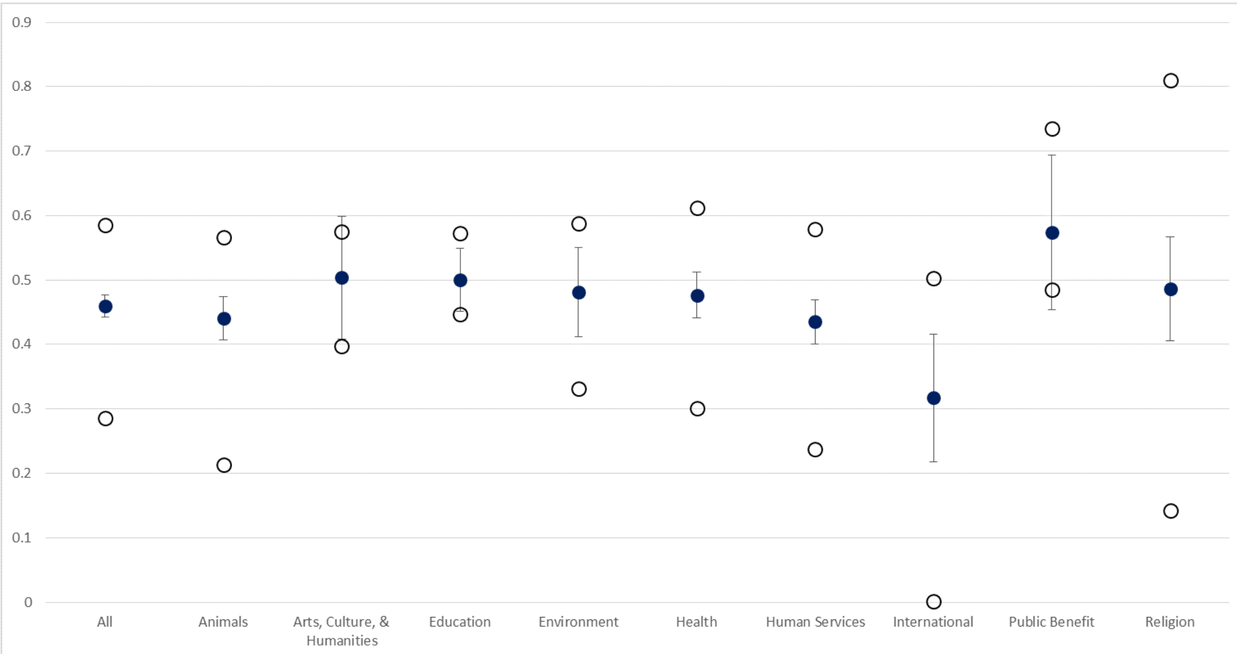
Category	**Coefficient**
Arts, Culture & Humanities	0.0414***
	(0.0140)
Education	-0.0291*
	(0.0150)
Environment	0.00620
	(0.0179)
Health	0.00225
	(0.0108)
Human Services	-0.00445
	(0.0109)
International	0.0264
	(0.0260)
Public Benefit	0.00210
	(0.0297)
Religion	-0.00179
	(0.0207)
Constant	0.884***
	(0.00730)
Observations	1051
R^2	0.01

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

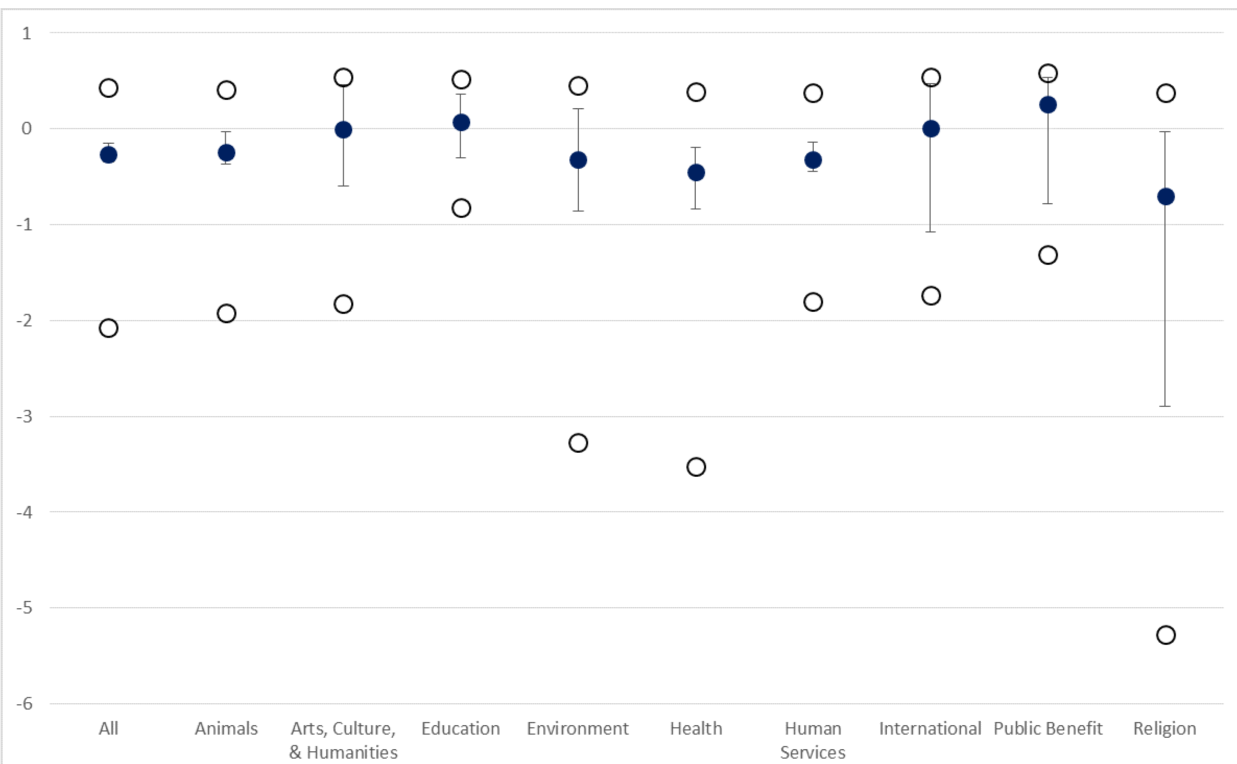
We continue to explore this observation using OLS in table III.4, where only one category (Arts, Culture & Humanities) exhibits significantly more rationality than the reference group (Animals). Kolmogorov-Smirnov testing confirms this difference ($p = 0.046$).

Figure III.9: CES Parameter Distributions Across Charity Types

(a) Estimated Mean $\hat{\alpha}$



(b) Estimated Median $\hat{\rho}$

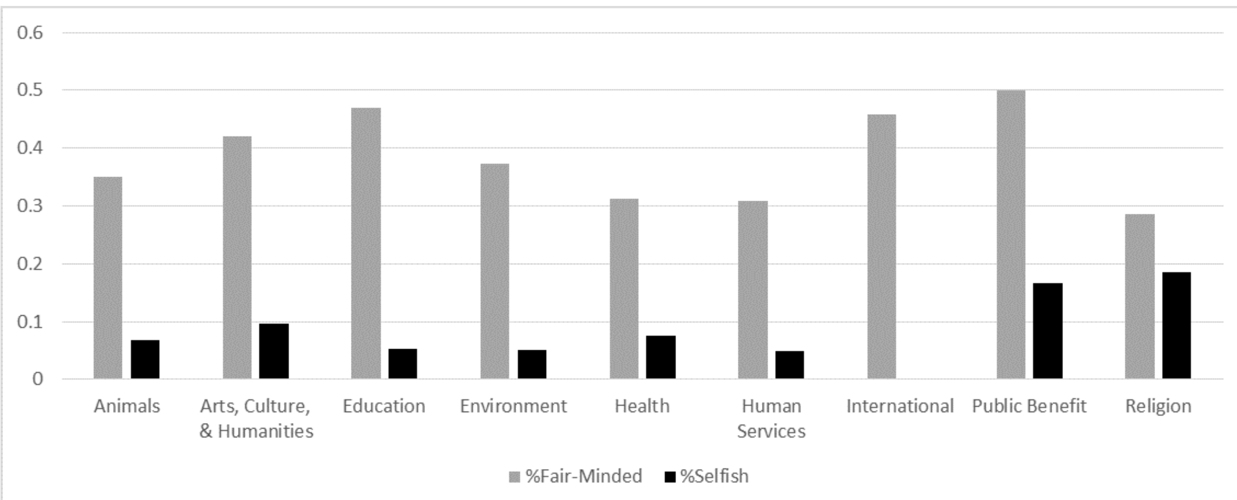


III.4.2 Structural Estimates Across Type

Given the rationality exhibited both throughout the dataset and within specific charitable types, we follow the logic from the previous section and examine estimated parameters across charity type.

Again, we compare the distributions of the key parameters (namely ρ and α) across charitable categories. Figure III.9 shows the distributions of our imputed parameters. An observation of note is that there appears to be more within category variation than across category variation.

Figure III.10: Classifying Subjects' Preference over Own Income

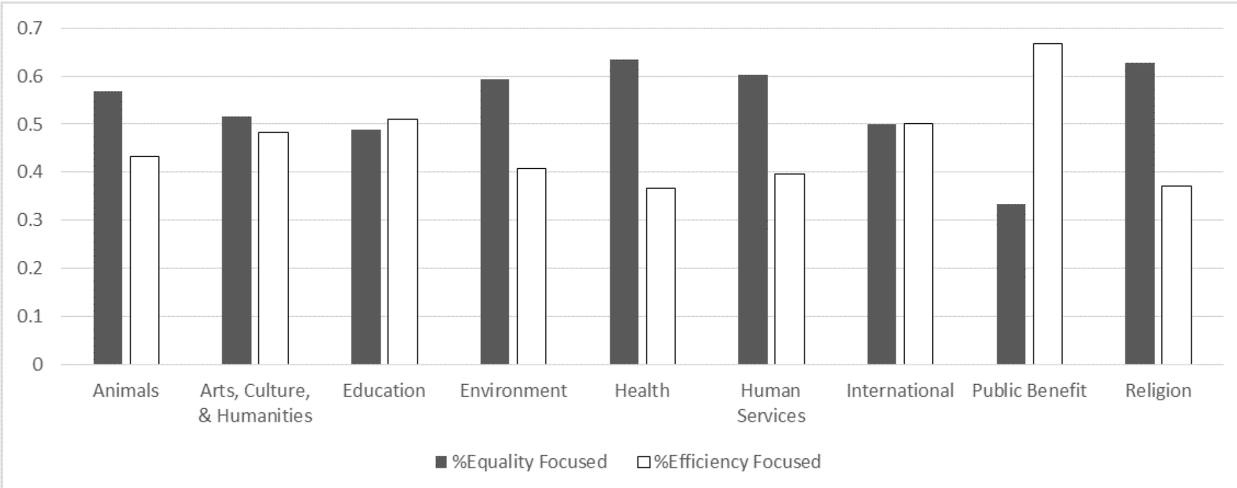


The bars indicate the percentage of subjects in each cell.
We classify a subject as fair-minded if $0.45 < \hat{\alpha}_n < 0.55$; a subject is classified as selfish if $\hat{\alpha}_n > 0.95$

Using Fisman et al.'s 2015 definitions of fair-mindedness and selfishness, we further compare $\hat{\alpha}$ distributions in figure III.10. Interestingly enough, the givers to religious charitable causes have the highest number of selfish participants.

We further examine a subject's equity-efficiency trade offs by means of $\hat{\rho}$ distributions in figure III.11. Again, Religious causes provide interesting insight into motives for giving, where givers' preferences for equality vastly outnumber those who are efficiency-focused. Compare this to givers of Public Benefit causes, where we see almost the exact opposite. Public Benefit Charities (e.g. The United Way) are the only category which efficiency-focused

Figure III.11: Classifying Subjects' Preference over Total Income



The bars indicate the percentage of subjects in each cell. We classify a subject as equality focused if $\hat{\rho}_n < 0$; a subject is classified as efficiency focused if $\hat{\rho}_n > 0$

subjects outnumber those who are equality focused.

III.5 Concluding Remarks

This paper examines motives for giving, and further decomposes those motives based on *where* subjects actually prefer to donate.

Our study concerns an experimental design that is strongly informed by theory. In doing so, we create a rich data set that allows us to impose the structure of a standard utility function (CES) to better understand these motives.

The benefits of this design are manifold. In particular, we are able to estimate parameters at the individual level, and thereby comment on full parameter distributions, rather than simple sample averages.

Further, decomposing by charity is a necessary step to build on previous literature. Rather than rely on conventional wisdom, we find that differences within outnumber differences across demographic and charity types. In doing so, we reveal empirical evidence for how charities can best target potential donors.

As such, our paper speaks to multiple literatures including public and behavioral economics, and structural econometrics. Of course, more work in this area is needed, and this paper is merely the first in a series of rich questions to be asked. In forthcoming projects, we aim to further disentangle motives for giving and better inform policy by adding extra dimensions to the charitable choice such as highlighting individual receivers within a charitable organization.

Chapter IV

Summary and Personal Reflection

In a recent conversation...

...with a colleague, he lamented the tendency of economics experiments to attempt to answer several questions at once—a practice at odds with other experimental sciences such as chemistry and physics. I agree with this standpoint and note that this tendency is particularly severe when we forsake simplicity in pursuit of a *magnum opus*.

And yet, when my advisor asked me to reflect on what I've learned over the course of my graduate work, I immediately lost sight of my own advice. I began to draft (what I hoped was) an irreverent and holistic economic model of dissertation writing. In it, I quoted luminaries such as Ackerloff, Friedman, and Varian; and used words like stochastic and lexicographic.

Those words are not in this draft. In the former, I neglected to *reflect*. Reflect on what areas of research interest me, why they do, and what I've learned from them. Most importantly, I neglected to reflect on who I am as a scholar.

This reflection is surprisingly difficult. One of the great joys of an economics dissertation is it can be structured to highlight a broad set of research skills. However, this can make it hard to draw unifying insight from the work as a whole.

What follows is the story of my experience writing the dissertation. While the results of my research are presented in the preceding chapters, here I aim to focus on the lessons gleaned from *conducting* the research, replete with its associated laurels and hurdles.

IV.1 Implicit Bias

I had been vaguely aware of the concept of implicit bias as it grew in the cultural consciousness. I had read Blink (Gladwell, 2005), and I had seen the episodes of Oprah. In spite of this, I didn't become interested in it as an area of research until the fall of 2012, when my father happened to be in town for a conference on employment law.

One day, I met him for lunch and he began telling me about his experience taking an IAT a few sessions earlier. "I could *feel* myself slowing down" he was saying. Fresh off my first graduate field courses in experimental, I was struck by what I viewed as several flaws with this comment.

While my goals in writing the paper have become more nuanced since then, many of those insights are still in chapter 1: the differences between cheap talk and incentivized action, the role of marketplace decisions, the vectors of skills that various tasks could be highlighting.

Outside of the paper's own results, a great lesson here is not only that economic inspiration can come from anywhere, but also the dividends that payoff in wheting and honing a thesis through the process of workshopping and presenting.

Having said that, the paper was not without its difficulties. From the stance of implementation, my subject pool was not as balanced as I had envisioned in designing the experiment. However, there is no perfect lab, and this realization was actually quite liberating for me. On the analysis side, this chapter has null results which come with their own provisos and limitations. After relentlessly scrutinizing and torturing the data I have come to believe that these are true null results, and learned many things in the process about how to better ask and answer questions. I think my committee would want me to stress the importance of asking questions that aren't mere statistical exercises, but look for *economic* significance, and in doing so learning to ask questions where even null results are interesting results.

IV.2 Girls' School

In the fall of 2013, I was called into a meeting with my advisor and the (then) director of graduate studies. Given all my neuroses, I assumed something terrible was about to happen. That "something terrible" was the girls' school paper, what is now chapter 2 of this dissertation, and my first coauthorship.

While any first is bound to have a learning curve[1], perhaps my most valuable lessons came from this paper's "seconds," by which I mean the follow-up experiments. Though not part of chapter 2, *per-se*, this experience offers a microcosm of dissertation lessons as a whole.

These follow-up experiments had elements that were both extremely fortunate and extremely lacking. For instance, our IRB application was accepted without revisions, but our subject recruitment was underwhelming. It is not my aim to disparage any of the parties involved, or even comment on the merits of lab versus field environments. Rather I wish to pay deference to that fact that, whether providential or ill-fated, any experimental endeavor is replete with chance. Experimentation is necessarily a process where the outcome is unknown, and as such, makes the researcher susceptible to Murphy's Law. In course, I've found that there are great benefits to additional groundwork and preparedness.

IV.3 Charitable Giving

Experimental economics is a top-heavy process. Rather than cleaning and coding data, we spend a great deal of our time tweaking and adjusting designs, such that when the data are ready the analysis tends to be fairly straightforward. I found this to be the case even in my third chapter, which was by far the most structural of the three.

[1] Including the importance of debugging code, where special thanks are due to Glenn Harrison

In many ways, this chapter was the synthesis of what I had already been assimilating while writing the dissertation. Along with my coauthors I was able to refine the question to fill an interesting hole in the literature, I was perpetually writing, even during some of (what felt like) the more stagnant experimental design phases, and I learned to better lean on the advice and assistance of others.

In the chapter's opening footnote, I remark that the paper is part of a larger, ongoing project. While this is particularly true with regards to "The Richness of Giving," the same could be said for the research process as a whole.

IV.4 Concluding Remarks

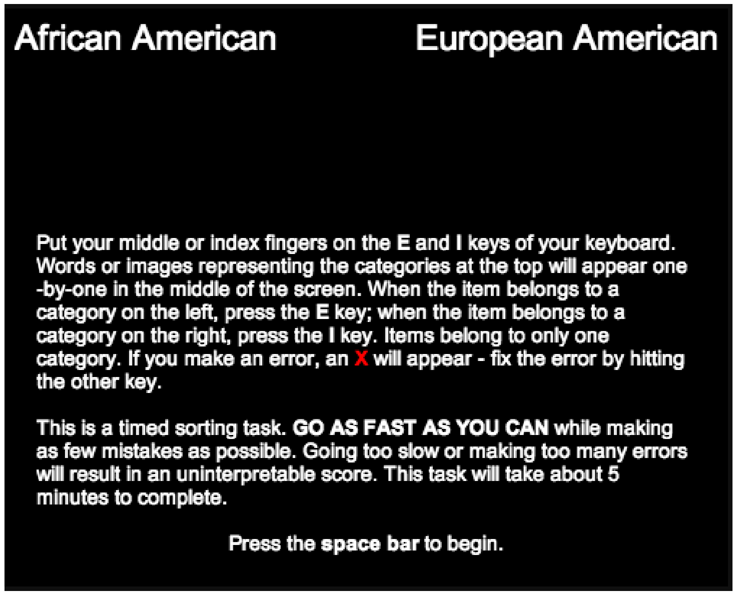
Though we often forget from the solitude of our laptops, this reflection serves to remind me how collaborative academe is. Not only the act of coauthorship itself, but the entire research process: brainstorming ideas, workshopping papers, debugging, writing, editing. This collaboration is compounded in economics experiments where we need assistance to run the damn things.

A common thread from my lessons in collaboration is the maintenance of perspective. Whether it is the context of a broader literature in your problem, how to best communicate that problem to others, or just proper respect paid to those you collaborated with in the first place.

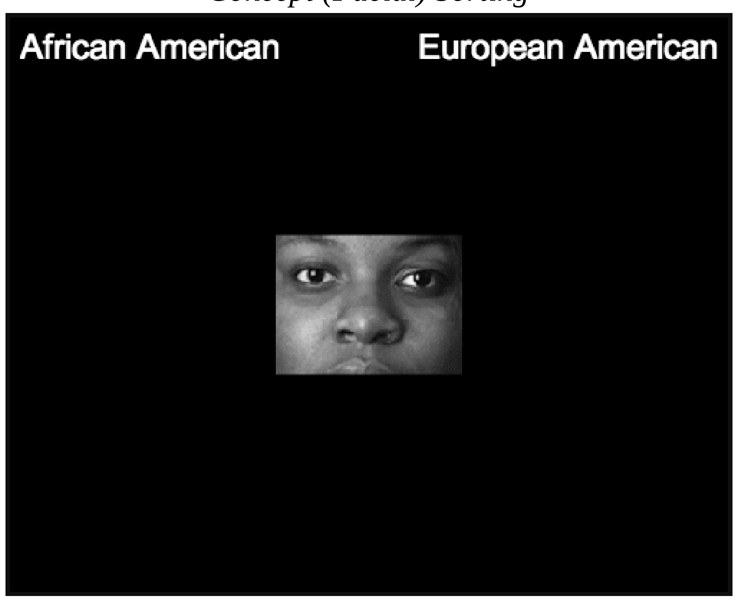
In this vein, as I stand on the precipice of my doctorate, I am very grateful for the support of my Chair, Dr. Susan Laury. My committee members Drs. Charles Courtemanche, John A. List, and Michael K. Price. Non-committee coauthors Drs. Shachar Kariv and Kurt E. Schnier, as well as numerous peers, assistants, funding agencies and anyone else whose invaluable assistance made "Three Essays on Social Issues in Experimental Economics" a possibility. Many kind thanks are owed to all.

Appendix A) IAT Screenshots

On Screen Instructions

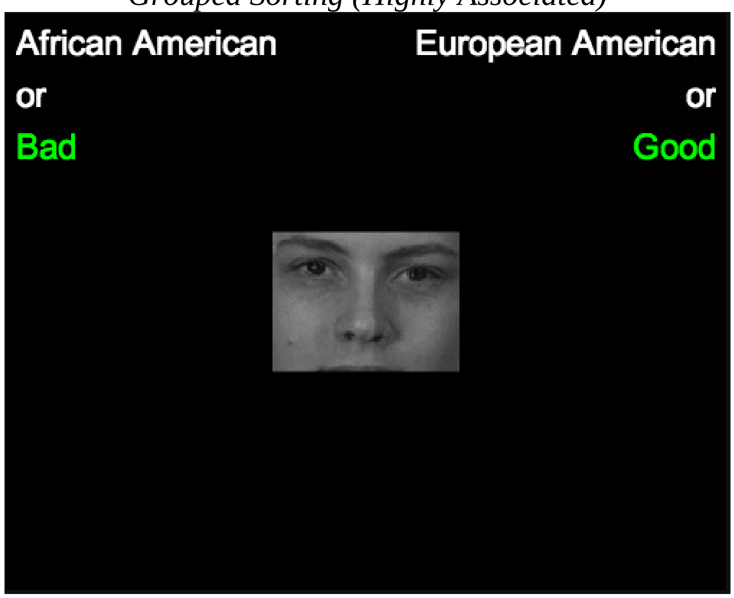


Concept (Facial) Sorting

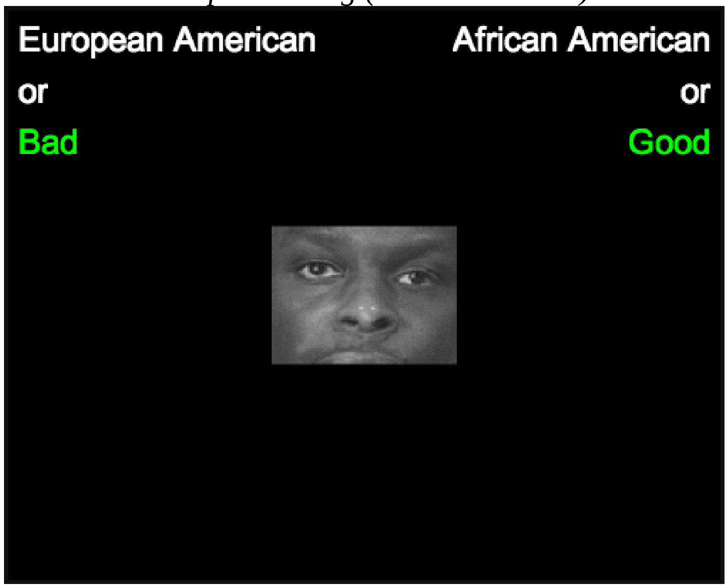


Attribute (Word) Sorting

Grouped Sorting (Highly Associated)



Grouped Sorting (Less Associated)



Appendix B

In this appendix, I adapt the language of Lazear et al. (2012) and include formal definitions of Reluctant, Willing, and Non-sharers

Definition 1. *A **Willing Sharer** (i) shares a positive amount in a sharing environment and (ii) prefers to be in such an environment when $w = w'$.*
(i) $\arg\max_{x \in [0,w]} U(1, x, w - x) < w$
(ii) $\max_{x \in [0,w]} U(1, x, w - x) > U(0, w, 0)$

Definition 2. *A **Reluctant Sharer** (i) shares a positive amount when in a sharing environment but (ii) prefers to not have the option when there is no financial reward to sharing.*
(i) $\arg\max_{x \in [0,w]} U(1, x, w - x) < w$
(ii) $\max_{x \in [0,w]} U(1, x, w - x) < U(0, w, 0)$

Definition 3. *A **Non-Sharer** (i) does not share, even if the environment allows for it.*
(i) $\arg\max_{x \in [0,w]} U(1, x, w - x) = w$

Appendix C) Subject Instructions

Treatment: Pictures, Costly Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

In the first part of this experiment, you have been given the choice of whether or not to participate in the following activity. That is, participating in this activity is optional.

You have been randomly paired with the participant displayed on your screen. This person is completing a different task that may include different payments, and does not know that he or she is participating with you. If you choose to participate, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. If you choose to participate, I will explain the activity to the other person. That is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not see your picture. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Moreover, you may decide to not participate in the above activity. If you choose this option, you will receive a fixed amount of $9 (plus the $5 for participation). The other person will receive $5 for participation. He or she will not receive any information about this activity. Please indicate your choice on the sheet below.

Decision Sheet

I wish to (circle one) Participate/Not Participate

If you are participating, please indicate

Amount of money to give to the other person: ________________

Amount of money to keep for yourself : ________________

(these two quantities must sum up to $10.00)

Treatment: Pictures, Costless Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

In the first part of this experiment, you have been given the choice of whether or not to participate in the following activity. That is, participating in this activity is optional.

You have been randomly paired with the participant displayed on your screen. This person is completing a different task that may include different payments, and does not know that he or she is participating with you. If you choose to participate, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. If you choose to participate, I will explain the activity to the other person. That is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not see your picture. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Moreover, you may decide to not participate in the above activity. If you choose this option, you will receive a fixed amount of $10 (plus the $5 for participation). The other person will receive $5 for participation. He or she will not receive any information about this activity. Please indicate your choice on the sheet below.

Decision Sheet

I wish to (circle one) Participate/Not Participate

If you are participating, please indicate

Amount of money to give to the other person: _________________

Amount of money to keep for yourself : _________________

(these two quantities must sum up to $10.00)

Treatment Pictures, No Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

You have been randomly paired with the participant displayed on your screen. In this part of the experiment, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. After you make your decision, I will explain the activity to the other person, that is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not see your picture. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Decision Sheet

Amount of money to give to the other person: __________________

Amount of money to keep for yourself : __________________

(these two quantities must sum up to $10.00)

Treatment: No Information, Costly Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

In the first part of this experiment, you have been given the choice of whether or not to participate in the following activity. That is, participating in this activity is optional.

You have been randomly paired with a participant in this room. This person is completing a different task that may include different payments, and does not know that he or she is participating with you. If you choose to participate, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. If you choose to participate, I will explain the activity to the other person. That is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not learn who you are. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Moreover, you may decide to not participate in the above activity. If you choose this option, you will receive a fixed amount of $9 (plus the $5 for participation). The other person will receive $5 for participation. He or she will not receive any information about this activity. Please indicate your choice on the sheet below.

Decision Sheet

I wish to (circle one) Participate/Not Participate
If you are participating, please indicate
Amount of money to give to the other person: _________________
Amount of money to keep for yourself : _________________
(these two quantities must sum up to $10.00)

Treatment: No Information, Costless Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

In the first part of this experiment, you have been given the choice of whether or not to participate in the following activity. That is, participating in this activity is optional.

You have been randomly paired with a participant in this room. This person is completing a different task that may include different payments, and does not know that he or she is participating with you. If you choose to participate, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. If you choose to participate, I will explain the activity to the other person. That is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not learn who you are. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Moreover, you may decide to not participate in the above activity. If you choose this option, you will receive a fixed amount of $10 (plus the $5 for participation). The other person will receive $5 for participation. He or she will not receive any information about this activity. Please indicate your choice on the sheet below.

Decision Sheet

I wish to (circle one) Participate/Not Participate

If you are participating, please indicate

Amount of money to give to the other person: _________________

Amount of money to keep for yourself : _________________

(these two quantities must sum up to $10.00)

Treatment: No Information, No Sorting

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group A Instructions

You have been randomly paired with a participant in this room. In this part of the experiment, you will be given $10. It is your task to decide how much to distribute between yourself and the person with whom you are paired. In other words, you must decide how much money, between $0 and $10 to give to the other person and how much to keep for yourself. You may select any amount between $0 and $10. For example, you may decide to give $9 to the other person and keep $1 for yourself, or you may instead decide to give $1 to the other person and keep $9 for yourself. After you make your decision, I will explain the activity to the other person, that is, the other person will learn the rules of the allocation task and the assigned amounts you assigned. He or she will not learn who you are. The assigned amounts will then be paid to both you, in addition to your show-up fees.

Decision Sheet

Amount of money to give to the other person: __________________

Amount of money to keep for yourself : __________________

(these two quantities must sum up to $10.00)

Receiver Instructions (Constant Across Treatments)

Instructions

Thank you for agreeing to participate. This is an experiment in two parts. We are interested in how people make decisions in social situations. Please read the instructions carefully, as your task may not be the same as those around you. During the session please do not talk or communicate with the other participants. If you have a question, please raise your hand and a research assistant will come answer it.

Everyone has already earned $5 for showing up. Additionally, you may have an opportunity to earn more. We will pay you privately in cash at the end of the session. None of the other participants will know the amount you have earned.

Group B Instructions

In the first part of this experiment, you are asked to complete the attached questionnaire. You will earn money based on how you answer these questions. After finishing, you will be asked to participate in an additional activity. The additional activity will not affect your payment in this part of the experiment.

In this questionnaire, you will be presented with a table that contains information on 10 different decisions that you must make. For each of the 10 decisions you must select either option 1 or option 2. The outcome of each option depends on the role of a 10-sided die. You will be paid based on your decisions in this questionnaire and partly on chance. Below is an example of the first three decisions you will make:

Decision	Option 1				Option 2			
1	Roll	1 for $2	or	2-10 for $1.60	Roll	1 for $3.85	or	2-10 for $0.10
2	Roll	1,2 for $2	or	3-10 for $1.60	Roll	1,2 for $3.85	or	3-10 for $0.10
3	Roll	1-3 for $2	or	4-10 for $1.60	Roll	1-3 for $3.85	or	4-10 for $0.10

Here is how I will pay you for this activity: I will first roll the 10-sided die to determine which decisions will receive payment and then re-roll the 10-sided die to determine your final earnings based on whether or not you selected option 1 or 2. All die rolls will be conducted after you have completed the experiment.

In the example above, suppose that I roll the 10-sided die and it lands on 1. Then the first row will be selected for payment. Now supposed I reroll the die and it lands on 6. If this is true you will receive $1.60 if you had selected Option 1 and $0.10 if you had selected Option 2. However, if the 10-sided die lands on 1 you will receive $2.00 if you had selected Option 1 and $3.85 if you had selected Option 2.

Please indicate your decision for each of the 10 rows on the opposite side of this sheet:

Please indicate your choice by circling either option 1 or 2 in the far right column. Only choose one option for each decision:

Decision	Option 1				Option 2				My Choice	
1	Roll	1 for $2	or	2-10 for $1.60	Roll	1 for $3.85	or	2-10 for $0.10	Option 1	Option 2
2	Roll	1,2 for $2	or	3-10 for $1.60	Roll	1,2 for $3.85	or	3-10 for $0.10	Option 1	Option 2
3	Roll	1-3 for $2	or	4-10 for $1.60	Roll	1-3 for $3.85	or	4-10 for $0.10	Option 1	Option 2
4	Roll	1-4 for $2	or	5-10 for $1.60	Roll	1-4 for $3.85	or	5-10 for $0.10	Option 1	Option 2
5	Roll	1-5 for $2	or	6-10 for $1.60	Roll	1-5 for $3.85	or	6-10 for $0.10	Option 1	Option 2
6	Roll	1-6 for $2	or	7-10 for $1.60	Roll	1-6 for $3.85	or	7-10 for $0.10	Option 1	Option 2
7	Roll	1-7 for $2	or	8-10for $1.60	Roll	1-7 for $3.85	or	8-10for $0.10	Option 1	Option 2
8	Roll	1-8 for $2	or	9,10 for $1.60	Roll	1-8 for $3.85	or	9,10 for $0.10	Option 1	Option 2
9	Roll	1-9 for $2	or	10 for $1.60	Roll	1-9 for $3.85	or	10 for $0.10	Option 1	Option 2
10	Roll	1-10 for $2	or	- for $1.60	Roll	1-10 for $3.85	or	- for $0.10	Option 1	Option 2

Appendix D) Demographic Survey

Demographic Survey

Below are several questions relating to your background. Your answers here will help us in conducting statistical analysis. Your name will not be matched with your responses and all information will be kept confidential. Please indicate if you prefer not to answer a particular question or if you would like to leave the study at any time. Please answer the questions honestly and to the best of your ability.

1) What is your age? __________

2) What gender do you identify with:
- ☐ Male
- ☐ Female
- ☐ Prefer Not to Answer

3) Which of these groups best describes you?
- ☐ White
- ☐ Black or African-American
- ☐ Hispanic
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ Other
- ☐ Prefer Not to Answer

4) What religion do you currently identify with?
- ☐ Catholic
- ☐ Protestant
- ☐ Muslim
- ☐ Jewish
- ☐ Agnostic
- ☐ No Religion
- ☐ Don’t Know
- ☐ Prefer Not to Answer
- ☐ Other

5) Have you participated in an economics experiment previously?
- ☐ Yes
- ☐ No
- ☐ Don’t Know
- ☐ Prefer Not to Answer

6) What is your current year in school?
- ☐ Freshman
- ☐ Sophomore
- ☐ Junior
- ☐ Senior
- ☐ Graduate Student
- ☐ I am not currently enrolled in school
- ☐ Prefer Not to Answer

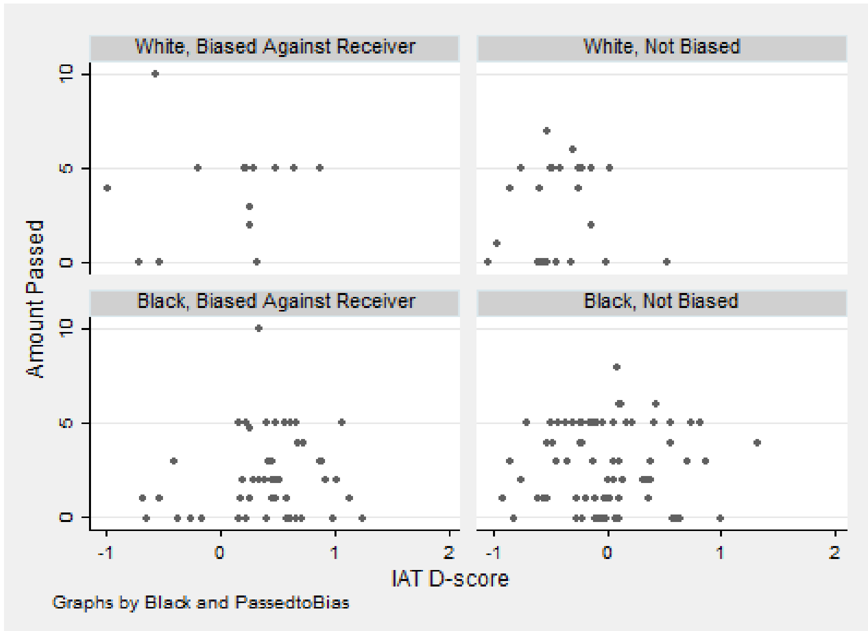
7) What is your GPA?
- ☐ _____________
- ☐ Prefer Not to Answer

8)What is your Major?

____

Appendix E

Scatter plots of IAT score and Amount Passed by Race and Bias of Dictator



In this appendix we start by looking at giving in finer bins in the photo treatments. Specifically the bias of the dictator. Here our definition of bias is IAT scores beyond ±0.15. The greatest difference in means exists between passing to the same and other for those holding an Anti-Black bias. This difference is not significant.

Table 1: Average Amounts Passed By Bias and Equivalence of Race

	IAT Threshlod		
	Anti-White	None	Anti-Black
Same Race	2.964 (1.971, n=28)	2.179 (2.342, n=28)	2.516 (2.206, n=48)
Other Race	2.742 (2.756, n=31)	2.600 (2.591, n=10)	3.233 (1.960, n=30)
Avg. Pass	2.847	2.289	2.792
Total Obs	59	38	78

Std. Deviations & Observations in Parentheses

IAT cutoffs at bias thresholds of ≤ -0.15 and ≥ 0.15

Next we look at giving in the context of a double-hurdle model (Cragg, 1971). This more flexible model, displayed in table 2 allows two separate processes for choosing to stay in, and how much one participates conditional on staying in. However, in order to specify these processes, it is necessary that I restrict the sample to those sessions with a sorting option (n=126). Even with this increased flexibility, IAT is still neither a significant predictor of dictator giving nor sorting.

Table 2: The IAT's Effect on Percent Shared Hurdle Model

Panel A: Hurdle				
Variable	**(1)**	**(2)**	**(3)**	**(4)**
IAT D-score	0.0584	0.107	0.108	0.154
	(0.116)	(0.118)	(0.119)	(0.226)
Receiver is Black		0.340	0.338	0.350
		(0.271)	(0.271)	(0.276)
Receiver is Same Gender		-0.559**	-0.558**	-0.561**
		(0.253)	(0.253)	(0.253)
Costly Sorting			-0.0440	-0.0535
			(0.236)	(0.239)
IATxPassedBlack				-0.129
				(0.533)
Constant	0.428***	0.525**	0.550*	0.552*
	(0.116)	(0.257)	(0.290)	(0.290)
Panel B: Above				
Variable	**(1)**	**(2)**	**(3)**	**(4)**
IAT D-score	-0.363	-0.383	-0.381	-0.976
	(0.311)	(0.303)	(0.300)	(0.611)
Receiver is Black		-0.487	-0.554	-0.735
		(0.696)	(0.692)	(0.702)
Receiver is Same Gender		0.878	0.922	1.029*
		(0.606)	(0.600)	(0.604)
Costly Sorting			-0.589	-0.609
			(0.570)	(0.564)
IATxPassedBlack				1.600
				(1.412)
Constant	3.251***	3.167***	3.510***	3.555***
	(0.345)	(0.636)	(0.696)	(0.688)
Sigma	2.306***	2.265***	2.246***	2.223***
	(0.264)	(0.256)	(0.252)	(0.247)
Observations	126	126	126	126

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

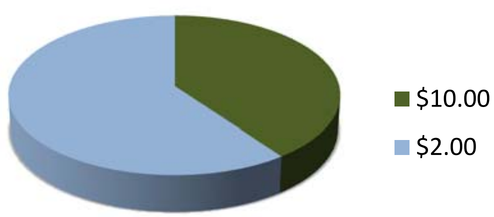
Appendix F

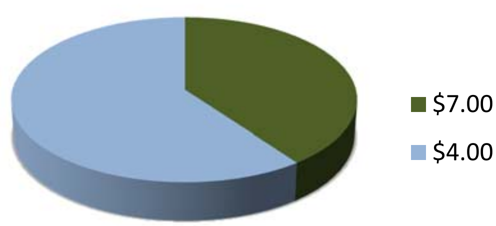
STUDENT INSTRUCTIONS

There are two separate games that you are going to play today. You will first play Game 1 and then proceed to Game 2. Listed below is a description of the two games and the tasks/decisions you will make. After you have completed both games we will flip a coin to determine which game you will receive payment. You will only receive payment for one of the two games.

GAME 1

In game one you will be presented with a folder that contains information on 10 different decisions that you must make. For each of the 10 decisions you must select either option A or option B. The outcome of each option depends on the role of a 10-sided die. The decision task is illustrated two different ways. At the top of each page you will see a pie chart that graphically illustrates the probable outcomes of your decision. Listed below the pie chart is a verbal description of the game. To determine the outcome of the game the study leader will roll a 10-sided die with your payment depending on the die number. After you have made all 10 of your decisions, the study leader will first roll the 10-sided die to determine which of the 10 decisions you will receive payment and then re-roll the 10-sided die to determine your final earnings based on whether or not you selected option A or B. All die rolls will be conducted after you have completed both Game 1 and 2. Below is an example of the decision task you will make:





Option A
$10 if throw of die is 1 – 4
$2 if throw of die is 5 – 10

Option B
$7 if throw of die is 1 – 4
$4 if throw if die is 5 – 10

Suppose that the study leader rolls the 10-sided die and it lands on 6. If this is true you will receive $2.00 of you had selected Option A and $4.00 if you had selected Option B. However, if the 10-sided die lands on 3 you will receive $10.00 if you had selected Option A and $7.00 if you had selected Option B. The color-coded pie chart illustrates how probable each outcome is, the larger the area the more probable the outcome. For instance, under Option A there is a 40% chance that you will earn $10.00 and a 60% percent chance you will earn $2.00. Under Option B there is a 40% chance you will earn $7.00 and a 60% chance you will earn $4.00.

GAME 2

In Game 2 you will be asked to under hand lob (you must under hand lob) a tennis ball ten times into a basket that is 10 feet away. You must remain behind the line when tossing each of the tennis balls. Before playing the game you must decide which payment method you wish to receive. Under option A you will receive $2 for every tennis ball that you successfully under hand lob into the basket. Under option B you will be randomly paired with one of your fellow students and if you lob more tennis balls in the basket then they do you will receive $8 for each tennis ball that exceeds the number lobbed into the basket by your competitor. At no time during the game will you know which student in your class is the student you have been paired with; you will be informed of your earnings upon completion of both Game 1 and 2. For instance, if you lob 6 tennis balls into the basket and your paired competitor lobs 4 tennis balls into the basket you will earn (6-4)*$8=2*$8=$16 in the game. However, if they had lobbed 6 tennis balls in the basket as well you would earn (6-6)*$8=0*$8=$0. In the case that your competitor lobs more tennis balls in the trash can than you, you will earn $0 (it is not possible to lose money).

Student Survey
Student Number____________

1. How old are you? _________________________
2. What grade are you in at school? ________________________
3. Please circle your gender.
    a. Girl
    b. Boy
4. What sports teams have you played on in the past year (it is ok to write “none” if you don’t play any sports)?

5. Please circle all of the math courses that you have taken, including your current math course.
    a. 6th grade math
    b. pre-algebra
    c. algebra I
    d. geometry
    e. algebra II
    f. pre-calculus
    g. calculus
    h. statistics
    i. AP calculus
    j. economics
6. How much time do you spend doing homework every day? _____________________
7. If you had a math test coming up, how many hours would you spend studying for it? _____________________
8. How many AP courses have you taken? _______________________
9. How many AP courses have you taken in math or science? ______________________
10. On average how many hours do you spend with a parent every day? __________
11. How many times a week do you eat dinner as a family? __________

Appendix G

Complete List of Charitable Organizations

Category: Animals
Cause: Animal Rights, Welfare, and Services
Greenville Humane Society
American Veterinary Medical Foundation
Southeastern Guide Dogs
Banfield Charitable Trust
Puppies Behind Bars
Michigan Anti-Cruelty Society
San Francisco SPCA
Days End Farm Horse Rescue
Peggy Adams Animal Rescue League
Main Line Animal Rescue
Cause: Wildlife Conservation
Big Cat Rescue
Wildlife Conservation Network
WildAid
International Primate Protection League
Orangutan Foundation International
American Bird Conservancy
Pheasants Forever
Trout Unlimited
American Eagle Foundation
Pollinator Partnership
Cause: Zoos and Aquariums
Clearwater Marine Aquarium
North Carolina Aquarium Society
Columbus Zoo and Aquarium
Houston Zoo
Texas State Aquarium
Birmingham Zoo
Detroit Zoological Society
National Aquarium, Baltimore
Cincinnati Zoo & Botanical Garden
The Philadelphia Zoo
Category: Arts, Culture, and Humanities
Cause: Libraries, Historical Societies and Landmark Preservation
Archaeological Conservancy
The Seattle Public Library Foundation
Minnesota Historical Society
Louisville Free Public Library Foundation
Civil War Trust
Western Reserve Historical Society
George Washington's Mount Vernon
Center for Jewish History
The New York Public Library
The Friends of the Saint Paul Public Library
Cause: Museums
Chrysler Museum of Art
Walking Mountains Science Center
Santa Barbara Museum of Art
Los Angeles County Museum of Art
Children's Museum of Richmond
Honolulu Museum of Art
Missouri History Museu
Children's Museum of Houston
American Museum of Natural History
Metropolitan Museum of Art
Cause: Performing Arts
Houston Ballet
La Jolla Playhouse
Fractured Atlas
Colorado Springs Fine Arts Center
Paul Taylor Dance Foundation
Brooklyn Academy of Music
On the Boards
San Francisco Ballet
Boston Lyric Opera
The Raymond F. Kravis Center for the Performing Arts
Cause: Public Broadcasting and Media
Vermont Public Radio
NPR
Center for Investigative Reporting
StoryCorps
Maine Public Broadcasting Network
Twin Cities Public Television
New Hampshire Public Radio
Graywolf Press
Texas Public Radio
KUOW Puget Sound Public Radio
Category: Education

Cause: Universities, Graduate Schools, and Technological Institutes
University of Delaware
Medical College of Wisconsin
Emory University
Northeastern University
Drexel University
New York University
Dartmouth College
Baylor University
Carnegie Mellon University
Cornell University
Cause: Private Elementary and Secondary Schools
Epiphany School
The Kinkaid School
Pace Academy
KIPP DC
Grand Rapids Christian Schools
Kimball Union Academy
Mercersburg Academy
Wesleyan School
Abraham Joshua Heschel School
Harlem Academy
Cause: Private Liberal Arts Colleges
Davidson College
St. Olaf College
Claremont McKenna College
University of Puget Sound
Furman University
Wheaton College
Lafayette College
Gustavus Adolphus College
Wellesley College
Spelman College
Cause: Other Education Programs and Services
Step Up For Students
DonorsChoose.org
Communities In Schools National Office
The Parent-Child Home Program
National Medical Fellowships
GreatSchools
I Know I Can
Small Steps Nurturing Center
The BISON Children's Scholarship Fund
BELL
Category: Environment
Cause: Environmental Protection and Conservation
North Cascades Institute
Rare
Living Lands and Waters
Alliance for the Great Lakes
World Resources Institute
The Sierra Club Foundation
Trees Atlanta
Texas Parks and Wildlife Foundation
Conservation Law Foundation
Teton Science Schools
Cause: Botanical Gardens, Parks, and Nature Centers
Western Pennsylvania Conservancy
New England Wild Flower Society
San Francisco Parks Alliance
Grand Teton National Park Foundation
Golden Gate National Parks Conservancy
Naples Botanical Garden
Thomas Irvine Dodge Nature Center
World Forestry Center
Cincinnati Parks Foundation
The Battery Conservancy
Category: Health
Cause: Diseases, Disorders, and Disciplines
United Cerebral Palsy of Greater Chicago
National Kidney Foundation of Michigan
Glaucoma Research Foundation
National Alopecia Areata Foundation
Cure Alzheimer's Fund
Immune Deficiency Foundation
HelpHOPELive
Breast Cancer Connections
Children's Organ Transplant Association
FSH Society
Cause: Patient and Family Support
Camp John Marc
Sharsheret
Make-a-Wish Foundation of the Texas Gulf Coast and Louisiana
Camp Sunshine, Maine
Mercy Medical Airlift/Mercy Medial Angels

Operation Access
Make-a-Wish Foundation of Masachusetts and Rhode Island
Dream Foundation
Tom Coughlin Jay Fund Foundation
Boston Ronald McDonald House
Community Volunteers in Medicine

Cause: Treatment and Prevention Services

International Planned Parenthood Foundation/Western Hemisphere Region
St. Petersburg Free Clinic
Resource Center
Cenikor Foundation
Venice Family Clinic
Arlington Free Clinic
Wisconsin Women's Health Foundation
Planned Parenthood of Maryland
Fan Free Clinic

Cause: Medical Research

Masonic Medical Research Laboratory
The Lustgarten Foundation for Pancreatic Cancer Research
Breast Cancer Research Foundation
The Multiple Myeloma Research Foundation
Sabin Vaccine Institute
Alliance for Aging Research
Cancer Research Institute
Sansum Diabetes Research Institute
Damon Runyon Cancer Research Foundation
Oklahoma Medical Research Foundation

Category: Human Services

Cause: Children's and Family Services

Forever Young Foundation
Harlem Children's Zone
Cradles to Crayons
Mary's Center
Dave Thomas Foundation for Adoption
Martha's Table
Jewish Family Service of San Diego
Families Forward
Emergency Family Assistance Association
Court Appointed Special Advocates of Collin County

Cause: Youth Development, Shelter, and Crisis Services

Do Something
Boys & Girls Clubs of Central Florida
Place of Hope
Girls Inc. of Omaha
Royal Family Kids
Big Brothers Big Sisters of Eastern Missouri
Harlem RBI
St. Anne's
United Friends of the Children
Boys & Girls Clubs of Metropolitan Phoenix

Cause: Food Banks, Food Pantries, and Food Distribution

Midwest Food Bank
The Billings Food Bank
Weld Food Bank
San Antonio Food Bank
Second Harvest Food Bank of North Central Ohio
Central Illinois Foodbank
Second Harvest Food Bank of Northwest Pennsylvania
Northern Illinois Food Bank
The Food Bank of Lower Fairfield
Ozarks Food Harvest

Cause: Multipurpose Human Service Organizations

New York Cares
Good Sports
National Fallen Firefighters Foundation
Higher Ground Sun Valley
Special Olympics Arizona
United States Soccer Foundation
100 Club of Arizona
Armed Services YMCA
Adaptive Sports Association
All Hands Volunteers

Cause: Homeless Services

Preble Street
Homeless Emergency Project
Durham Rescue Mission
The Lord's Place
SOME
Abode Services
Safe Haven Family Shelter

Downtown Women's Center
Primavera Foundation
The INN
Cause: Social Services
Emergency Outreach Colorado
Air Warrior Courage Foundation
Fisher House Foundation
Boca Helping Hands
Eva's Village
Navy SEAL Foundation
Homes for Our Troops
Special Operations Warrior Foundation
Surrey Services for Seniors
REDF
Category: International
Cause: Development and Relief Services
Wings of Hope
Life in Abundance International
Aga Khan Foundation, USA
Child Aid
ECHO
Fistula Foundation
Kids Alive International
International Institute of Rural Reconstruction (IIRR)
Kiva
GlobalGiving
Cause: International Peace, Security, and Affairs
V-Day
Polaris Project
United Nations Foundation
Women's Learning Partnership
StandWithUs
Shared Hope International
International Center for Journalists
Institute of International Education
Birthright Israel Foundation
Human Rights Watch
Cause: Humanitarian Relief Supplies
Books for Africa
Project C.U.R.E.
Heart to Heart International
Direct Relief
Matthew 25: Ministries
Outreach
Feed My Starving Children
Brother's Brother Foundation
Project HOPE
MedShare International
Cause: Foreign Charity Support Organizations
Palestine Children's Relief Fund
CommonHope
Hadassah, The Women's Zionist Organization of America
The Citizens Foundation, USA
American-Israeli Cultural Foundation
Sankara Eye Foundation, USA
American Society for Yad Vashem
Fonkoze USA
BRAC USA
Solid Rock International
Category: Public Benefit
Cause: Advocacy and Civil Rights
Equal Justice Initiative
Chicago Foundation for Women
Physicians for Reproductive Health
Acton Institute for the Study of Religion and Liberty
Compassion & Choices
Injured Marine Semper Fi Fund
NumbersUSA
Institute for Justice
National Immigration Law Center
Freedom From Religion Foundation
Cause: Fundraising Organizations
Charities Aid Foundation America
Robin Hood Foundation
Jewish Community Federation of San Francisco, the Peninsula, Marin and Sonoma Counties
Greater Kalamazoo United Way
Arthritis National Research Foundation
United Way of Summit County
United Way of Cass-Clay
The Rose Foundation for Communities and the Environment
Elton John AIDS Foundation
AIDS United
Cause: Research and Public Policy Institutions

RESULTS Educational Fund
Carnegie Institution for Science
Public Interest Projects
Center for Food Safety
North Carolina Agricultural Foundation
The Brookings Institution
Mount Desert Island Biological Laboratory
The Federalist Society for Law and Public Policy Studies
Kentucky Youth Advocates
Woods Hole Oceanographic Institution

Cause: Community Foundations

The Community Foundation for Northeast Florida
The Community Foundation of Louisville
Community Foundation of North Texas
Orange County Community Foundation
The Columbus Foundation
Parasol Tahoe Community Foundation
Community Foundation for Southeast Michigan
Princeton Area Community Foundation
Community Foundation of Middle Tennessee
Community Foundation of New Jersey

Cause: Community and Housing Development

Habitat for Humanity of Greater Los Angeles
Habitat for Humanity of East Bay
San Gabriel Valley Habitat for Humanity
Rebuilding Together
Habitat for Humanity of Washington, D.C.
Enterprise Community Partners, Inc.
Cleveland Housing Network
Habitat for Humanity of Monroe County, IN
Habitat for Humanity of Omaha
Houston Habitat for Humanity

Category: Religion

Cause: Religious Activities

Mission Waco Mission World
Young Life
International Messengers
Urban Youth Impact
Commission To Every Nation
Hebrew Free Burial Association
Maoz Israel Ministries
Forward Edge International
Asian Access
Mission Arlington/Mission Metroplex

Cause: Religious Media and Broadcasting

Andrew Wommack Ministries
Ramesh Richard Evangelism and Church Health
Educational Media Foundation
Lutheran Bible Translators
WAY Media, Inc.
Pioneer Bible Translators
Lamb & Lion Ministries
Billy Graham Evangelistic Association
SAT-7
Blue Ridge Broadcasting Corporation

Appendix H

Experimental Screens and Instructions

Instructions (TEXT ONLY)

Welcome to the survey
Login code:______________

Please remember that participation in this survey is voluntary and you may skip over any questions that you would prefer not to answer. You will not be identified in any reports on this study.

This is an experiment in decision-making. Your payoffs will depend partly on your decisions and partly on chance. Please pay careful attention to the instructions as a considerable amount of money is at stake.

During the experiment we will speak in terms of experimental tokens instead of dollars. Your payoffs will be calculated in terms of tokens and then translated into dollars at the end of the experiment at the following rate:

2 Tokens = 1 Dollar

You are free to stop at any time. If you do not complete the experiment now, you may return to complete the experimental session at any time between now and 04-01-2016. If you do not complete the experiment before then, you will not receive any payment. Details of how you will make decisions and receive payments will be provided below.

This is an experiment in two stages. For stage one, you will be presented with information on several charitable organizations taken from the website www.CharityNavigator.com; afterwards you will be asked to select a preferred organization.

In stage two you will participate repeatedly in 50 independent decision problems that share a common form. We next describe in detail the process that will be repeated in all decision problems and the computer program that you will use to make your decisions.

In each decision problem you will be asked to allocate tokens between yourself and the charitable organization you selected in the previous stage. We will refer to the tokens that you allocate to yourself as tokens that you **Hold**, and tokes that you allocate to the chosen charity as **Pass**.

Charity navigator is a website that evaluates organizations which rely on public support and actively solicit donations from the public. It rates organizations which file IRS Form 990 along several dimensions and has been acclaimed by numerous publications as among the best or most useful websites.

They have identified 9 charitable categories and several causes within each category. The table on the next screen is adapted from the charity navigator website and contains information on the top ten charities within each cause. Please review the information in this table carefully and select your most preferred charity. If you like, you can also write in a different charity of your choice.

The Charity I select is___________

Each choice will involve choosing a point on a line representing possible token allocations to you (**Hold**) and to your charity (**Pass**). In each choice, you may choose any Hold / Pass pair that is on the line. Examples of lines that you might face appear in the diagrams below. In each graph. **Hold** corresponds to the vertical axis and **Pass** corresponds to the horizontal axis. The points on the diagonal lines in the graphs represent possible token allocations to **Hold** (tokens to you) and **Pass** (tokens to the charity) that you might choose.

By picking a point on the diagonal line, you choose how many tokens to hold for yourself and how many to pass to the charity. You may select any allocation to **Hold** or **Pass** on that line.

If, for example, the diagonal line runs from 50 tokens on the **Hold** axis to 50 tokens on the **Pass** axis (See Diagram 4), you could choose to hold all 50 tokens for yourself or pass all 50 tokens to the

To further illustrate, in the example below, choice A represents an allocation in which you hold **y** tokens and pass **x** tokens. Thus if you chose this allocation you will keep **y** tokens for yourself and pass **x** tokens to the charity. Another possible allocation is B, in which you hold **w** tokens and pass **z** to the charity.

Each of the 50 decision problems will start by having the computer select a diagonal line at random. All of the lines that the computer will select will intersect with at least one of the axes at 50 or more tokens, but will not intersect either axis at more than 100 tokens. The lines selected for you in different decision problems are independent of each other and depend solely upon chance.

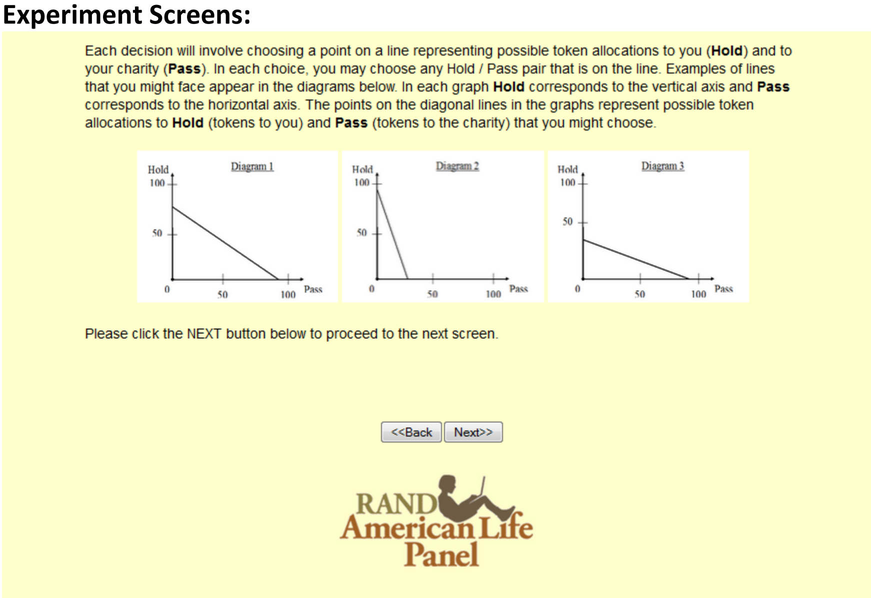
The computer program dialog window is shown here. In each round, you will choose an allocation by using the mouse to move the pointer on the computer screen to the allocation that you wish to choose (note that the pointer does not need to be precisely on the diagonal line to shift the allocation). When you are ready to make your decision, left-click to enter your chosen allocation. After that, confirm your decision by clicking on the OK button. Note that you can choose only Hold and Pass combinations that are on the diagonal line. Once you have clicked the OK button, your decision cannot be revised. After you submit each choice, you will be asked to make another allocation in a different decision problem involving a different diagonal line representing possible allocations. Again, all decision problems are independent of each other. This process will be repeated until all 50 decision rounds are completed. At the end of the last round, you will be informed that the experiment has ended.

Next, you will have a practice decision round. The choices you make in this practice round will have no impact on the final payoffs to you or to the charity. In this round, you may choose any combination of tokens to Hold (tokens to you) and Pass (tokens to the charity) that are on the line. To choose an allocation, use the mouse to move the cursor on the computer screen to the allocation that you desire. When you are ready to make your practice choice, left-click to enter your chosen allocation. To revise your allocation in the first practice round, click the CANCEL button. To confirm your decision, click on the OK button. You will then be automatically moved to the second practice round. After you complete the practice round, click NEXT to proceed to the next screen.

Payoffs will be determined as follows: At the end of the experiment, the computer will randomly select one of the 50 decisions you made to carry out for real payoffs. You will receive the tokens you held in that round (the tokens allocated to Hold). Your selected charity will receive the tokens that you passed (the tokens allocated to Pass). Note that the charity you selected is not making any allocation decisions. At the end of last round, you will be informed of the round selected for payment, and your choice and payment for the round. At the end of the experiment, the tokens will be converted into money. Each token will be worth 0.50 dollars, and payoffs will be rounded up to the nearest cent. Recall that you are free to stop at any time, and you may return to complete the experimental session at any time between now and 04-01-2016. If you do not complete the experiment between now and 04-01-2016, neither you nor your selected charity will receive any payment.

To review, in every decision problem in this experiment, you will be asked to allocate tokens to Hold and Pass. At the end of the experiment, the computer will randomly select one of the 50 decision problems to carry out for payoffs. The round selected depends solely upon chance. You will then receive the number of tokens you allocated to Hold in the chosen round. The charity you selected will receive the number of tokens you allocated to Pass in the chosen round. Each token will be worth 50 cents. If everything is clear, you are ready to start. Please click NEXT to proceed to the actual experiment.

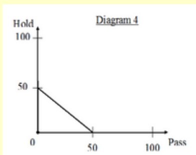
Experiment Screens:

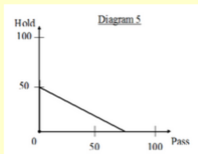


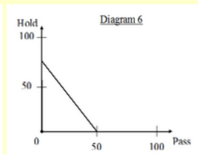
By picking a point on the diagonal line, you choose how many tokens to hold for yourself and how many to pass to the charity. You may select any allocation to **Hold** and **Pass** on that line.

If, for example, the diagonal line runs from 50 tokens on the **Hold** axis to 50 tokens on the **Pass** axis (see Diagram 4), you could choose to hold all 50 tokens for yourself, or pass all 50 tokens to the charity, or anything in between.

However, most of the decision problems will involve flatter or steeper lines: if the line is flatter (see Diagram 5), one less token for yourself means *more than* one additional token is passed to the charity; if the line is steeper (see Diagram 6), one less token held means *less than* one additional token passed to the charity.





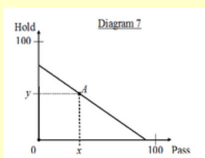


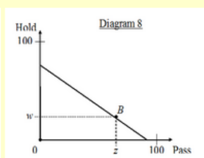
Please click the NEXT button below to proceed to the next screen.

<<Back Next>>

RAND American Life Panel

To further illustrate, in the example below, choice A represents an allocation in which you hold **y** tokens and pass **x** tokens. Thus, if you choose this allocation, you will hold **y** tokens for yourself and you will pass **x** tokens to the charity. Another possible allocation is B, in which you hold **w** tokens and pass **z** tokens to the charity.





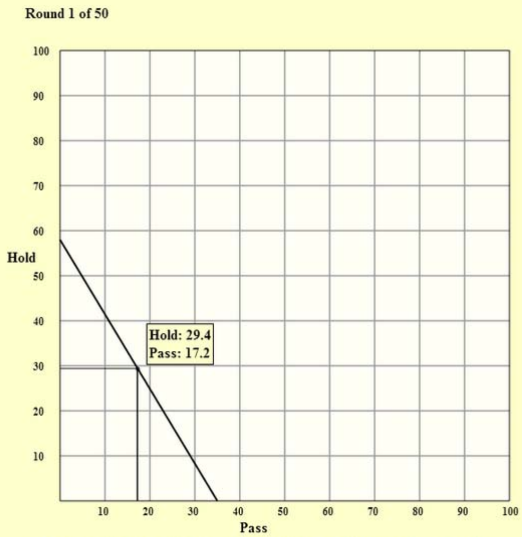
Please click the NEXT button below to proceed to the next screen.

<<Back Next>>

RAND American Life Panel

The computer program dialog window is shown here. In each round, you will choose an allocation by using the mouse to move the pointer on the computer screen to the allocation that you wish to choose (note that the pointer does not need to be precisely *on* the diagonal line to shift the allocation).

When you are ready to make your decision, left-click to enter your chosen allocation. After that, confirm your decision by clicking on the OK button. Note that you can choose only **Hold** and **Pass** combinations that are on the diagonal line. Once you have clicked the OK button, your decision cannot be revised.



After you submit each choice, you will be asked to make another allocation in a different decision problem involving a different diagonal line representing possible allocations. Again, all decision problems are independent of each other. This process will be repeated until all 50 decision rounds are completed. At the end of the last round, you will be informed that the experiment has ended.

Please click the NEXT button below to proceed to the next screen.

<<Back Next>>

RAND American Life Panel

Next, you will have two practice decision rounds. The choices you make in these practice rounds will have no impact on the final payoffs to you or to the other ALP respondent. In each round, you may choose any combination of tokens to **Hold** (tokens to you) and **Pass** (tokens to the charity) that are on the line. To choose an allocation, use the mouse to move the cursor on the computer screen to the allocation that you desire.

When you are ready to make your first practice choice, left-click to enter your chosen allocation. To revise your allocation in the first practice round, click the CANCEL button. To confirm your decision, click on the OK button. You will then be automatically moved to the second practice round. After you complete the two practice rounds, click NEXT to proceed to the next screen.

Please click the NEXT button below to enter the first practice round.

<<Back Next>>

RAND American Life Panel

Bibliography

Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc. 576 U.S. Supreme Court, 2015.

Sidney N Afriat. Efficiency estimation of production functions. *International Economic Review*, pages 568–598, 1972.

Joseph G Altonji and Rebecca M Blank. Race and gender in the labor market. *Handbook of labor economics*, 3:3143–3259, 1999.

Steffen Andersen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutström. Eliciting risk and time preferences. *Econometrica*, 76(3):583–618, 2008.

James Andreoni. Giving with impure altruism: applications to charity and ricardian equivalence. *The Journal of Political Economy*, pages 1447–1458, 1989.

James Andreoni. Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, 100(401):464–477, 1990.

James Andreoni. An experimental test of the public-goods crowding-out hypothesis. *The American Economic Review*, pages 1317–1327, 1993.

James Andreoni and John Miller. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, pages 737–753, 2002.

Kenneth Arrow et al. The theory of discrimination. *Discrimination in labor markets*, 3(10): 3–33, 1973.

Gary S Becker. *The economics of discrimination*. University of Chicago press, 1957.

Gary S Becker. A theory of social interactions. *Journal of Political Economy*, 82(6):1063–1093, 1974.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. *The American Economic Review*, 96(5):1652–1678, 2006.

Marianne Bertrand and Sendhil Mullainathan. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *The American Economic Review*, 94(4):991–1013, 2004.

Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1), 2004.

Marianne Bertrand, Dolly Chugh, and Sendhil Mullainathan. Implicit discrimination. *American Economic Review*, pages 94–98, 2005.

Alison Booth and Patrick Nolen. Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization*, 81(2):542–555, 2012a.

Alison Booth, Lina Cardona Sosa, and Patrick Nolen. Do single-sex classes affect achievement? an experiment in a coeducational university. Technical report, CEPR Discussion Papers, 2014.

Alison L Booth and Patrick Nolen. Gender differences in risk behaviour: does nurture matter?*. *The Economic Journal*, 122(558):F56–F78, 2012b.

Lex Borghans, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel. The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059, 2008.

Tomas Broberg, Tore Ellingsen, and Magnus Johannesson. Is generosity involuntary? *Economics Letters*, 94(1):32–37, 2007.

Alexander L. Brown, Jonathan Meer, and J. Forrest Williams. Social distance and quality ratings in charity choice. Working Paper 20182, National Bureau of Economic Research, May 2014. URL `http://www.nber.org/papers/w20182`.

Jamie Brown-Kruse and David Hummels. Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is? *Journal of Economic Behavior & Organization*, 22(3):255–267, 1993.

Steve Buchheit and Linda M Parsons. An experimental investigation of accounting informations influence on the individual giving process. *Journal of Accounting and Public Policy*, 25(6):666–686, 2006.

Crosby Burns. The costly business of discrimination. *Washington: Center for American Progress*, 2012.

T Buser, M Niederle, H Oosterbeek, et al. Gender, competitiveness and career choices. *The Quarterly Journal of Economics*, 129(3):1409–1447, 2014.

C Bram Cadsby and Elizabeth Maynes. Gender and free riding in a threshold public goods game: experimental evidence. *Journal of economic behavior & organization*, 34(4):603–620, 1998.

Colin Camerer. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2003.

Anne Campbell. *A mind of her own: the evolutionary psychology of women*. Oxford University Press, 2002.

Juan-Camilo Cárdenas, Anna Dreber, Emma Von Essen, and Eva Ranehill. Gender differences in competitiveness and risk taking: Comparing children in colombia and sweden. *Journal of Economic Behavior & Organization*, 83(1):11–23, 2012.

Marco Castillo and Ragan Petrie. Discrimination in the lab: Does information trump appearance? *Games and Economic Behavior*, 68(1):50–59, 2010.

World Giving Index, volume 4, dec 2013. Charities Aid Foundation.

Kerwin Kofi Charles and Jonathan Guryan. Studying discrimination: Fundamental challenges and recent progress. *Annu. Rev. Econ.*, 3(1):479–511, 2011.

Kenneth Y Chay, Michael Greenstone, et al. Does air quality matter? evidence from the housing market. *Journal of Political Economy*, 113(2):376–424, 2005.

Syngjoo Choi, Raymond Fisman, Douglas Gale, and Shachar Kariv. Consistency and heterogeneity of individual behavior under uncertainty. *The American economic review*, 97 (5):1921–1938, 2007.

Syngjoo Choi, Shachar Kariv, Wieland Müller, and Dan Silverman. Who is (more) rational? *The American Economic Review*, 104(6):1518–1550, 2014.

Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

John G Cragg. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, pages 829–844, 1971.

Rachel Croson and Uri Gneezy. Gender differences in preferences. *Journal of Economic literature*, pages 448–474, 2009.

Heidi Crumpler and Philip J Grossman. An experimental test of warm glow giving. *Journal of public Economics*, 92(5):1011–1021, 2008.

Jason Dana, Daylian M Cain, and Robyn M Dawes. What you dont know wont hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2):193–201, 2006.

Nabanita Datta Gupta, Anders Poulsen, and Marie Claire Villeval. Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, 51(1):816–835, 2013.

Douglas D Davis, Edward L Millner, and Robert J Reilly. Subsidy schemes and charitable contributions: a closer look. *Experimental economics*, 8(2):85–106, 2005.

Stefano DellaVigna, John A List, and Ulrike Malmendier. Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, 127(1):1–56, 2012.

Stefano DellaVigna, John A List, Ulrike Malmendier, Gautam Rao, et al. The importance of being marginal: Gender differences in generosity. *American Economic Review*, 103(3): 586–90, 2013.

Catherine C Eckel and Philip J Grossman. Altruism in anonymous dictator games. *Games and economic behavior*, 16(2):181–191, 1996.

Catherine C Eckel and Philip J Grossman. Rebate versus matching: does how we subsidize charitable contributions matter? *Journal of Public Economics*, 87(3):681–701, 2003.

Catherine C Eckel and Ragan Petrie. Face value. *The American Economic Review*, pages 1497–1513, 2011.

Gerald Eisenkopf, Zohal Hessami, Urs Fischbacher, and Heinrich W Ursprung. Academic performance and single-sex schooling: Evidence from a natural experiment in switzerland. *Journal of Economic Behavior & Organization*, 115:123–143, 2015.

Tore Ellingsen and Magnus Johannesson. Pride and prejudice: The human side of incentive theory. *American Economic Review*, 98(3):990–1008, 2008.

Paul J Ferraro and Ronald G Cummings. Cultural diversity, discrimination, and economic outcomes: an experimental analysis. *Economic Inquiry*, 45(2):217–232, 2007.

Chaim Fershtman and Uri Gneezy. Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics*, pages 351–377, 2001.

Klaus Fiedler and Matthias Bluemke. Faking the iat: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology*, 27(4):307–316, 2005.

Raymond Fisman, Shachar Kariv, and Daniel Markovits. Individual preferences for giving. *The American Economic Review*, 97(5):1858–1876, 2007.

Raymond Fisman, Pamela Jakiela, and Shachar Kariv. Distributional preferences and political behavior. 2015.

Jeffrey A Flory, Andreas Leibbrandt, and John A List. Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. *Review of economic studies*, 82(1):122–155, 2015.

Roland Fryer and S. Levitt. An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, pages 210–240, 2010.

Malcolm Gladwell. *Blink: the power of thinking without thinking.* Little, Brown and Co., 2005.

Uri Gneezy and Aldo Rustichini. Gender and competition at a young age. *The American Economic Review*, 94(2):377–381, 2004.

Uri Gneezy, Muriel Niederle, Aldo Rustichini, et al. Performance in competitive environments: Gender differences. *QUARTERLY JOURNAL OF ECONOMICS-CAMBRIDGE MASSACHUSETTS-*, 118(3):1049–1074, 2003.

Uri Gneezy, Kenneth L Leonard, and John A List. Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664, 2009.

Uri Gneezy, John List, and Michael K Price. Toward an understanding of why people discriminate: Evidence from a series of natural field experiments. Technical report, National Bureau of Economic Research, 2012.

Jacob K Goeree, Charles A Holt, and Susan K Laury. Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *Journal of public Economics*, 83(2): 255–276, 2002.

Teresa P Gordon, Cathryn L Knock, and Daniel G Neely. The role of rating agencies in the market for charitable contributions: An empirical test. *Journal of Accounting and Public Policy*, 28(6):469–484, 2009.

Anthony G Greenwald, Debbie E McGhee, and Jordan LK Schwartz. Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6):1464, 1998.

Anthony G Greenwald, Brian A Nosek, and Mahzarin R Banaji. Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of personality and social psychology*, 85(2):197, 2003.

Anthony G Greenwald, T Andrew Poehlman, Eric Luis Uhlmann, and Mahzarin R Banaji. Understanding and using the implicit association test: Iii. meta-analysis of predictive validity. *Journal of personality and social psychology*, 97(1):17, 2009.

David M Grether and Charles R Plott. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, pages 623–638, 1979.

P Grossman, C Eckel, et al. Giving versus taking: a real donation comparison of warm glow and cold prickle in a context-rich environment. Technical report, Monash University, Department of Economics, 2012.

Andrew Hanson and Zackary Hawley. Do landlords discriminate in the rental housing market? evidence from an internet field experiment in us cities. *Journal of Urban Economics*, 70(2):99–114, 2011.

Andrew Hanson, Zackary Hawley, and Aryn Taylor. Subtle discrimination in the rental housing market: Evidence from e-mail correspondence with landlords. *Journal of Housing Economics*, 20(4):276–284, 2011.

Glenn W. Harrison and Laurie T. Johnson. *Identifying Altruism in the Laboratory*, chapter 10, pages 177–223. doi: 10.1016/S0193-2306(06)11008-X.

Glenn W Harrison and John A List. Field experiments. *Journal of Economic literature*, 42 (4):1009–1055, 2004.

Glenn W Harrison and Richard D Phillips. Subjective beliefs and statistical forecasts of financial risks: The chief risk officer project. 2013.

J Heckman. Sample specification bias as a selection error. *Econometrica*, 47(1):153–162, 1979.

James J Heckman. Detecting discrimination. *The Journal of Economic Perspectives*, pages 101–116, 1998.

Charles A Holt and Susan K Laury. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

Steffen Huck, Imran Rasul, and Andrew Shephard. Comparing charitable fundraising schemes: Evidence from a natural field experiment and a structural model. *American Economic Journal: Economic Policy*, 7(2):326–69, May 2015. doi: 10.1257/pol.20120312.

Johnnie EV Johnson and Philip L Powell. Decision making, risk and gender: Are managers different? *British Journal of Management*, 5(2):123–138, 1994.

John T Jost, Laurie A Rudman, Irene V Blair, Dana R Carney, Nilanjana Dasgupta, Jack Glaser, and Curtis D Hardin. The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. *Research in organizational behavior*, 29:39–69, 2009.

Jerry Kang. *Implicit bias: A primer for courts*. National Center for State Courts, 2009.

Peter Kuhn and Marie Claire Villeval. Are women more attracted to co-operation than men? *The Economic Journal*, 125(582):115–140, 2015.

Edward P Lazear, Ulrike Malmendier, and Roberto A Weber. Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, 4(1): 136–163, 2012.

Daniel J. Lee. Risk aversion and implicit bias. Mimeo.

Andreas Leibbrandt and John A List. Do women avoid salary negotiations? evidence from a large-scale natural field experiment. *Management Science*, 61(9):2016–2024, 2014.

John A List. The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics*, pages 49–89, 2004.

Sara Lowes, Nathan Nunn, James A Robinson, and Jonathan Weigel. Understanding ethnic identity in africa: Evidence from the implicit association test (iat). *American Economic Review*, 105(5):340–45, 2015.

Evelyn A McDowell, Wei Li, and Pamela C Smith. An experimental examination of us individual donors information needs and use. *Financial Accountability & Management*, 29 (3):327–347, 2013.

Adam W Meade. Freeiat: An open-source program to administer the implicit association test. *Applied psychological measurement*, 33(8):643–643, 2009.

James Mill. *Analysis of the Phenomena of the Human Mind*, volume 1. London: Longmans, Green, Reader and Dyer, 1869.

David Neumark. Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, 47(4):1128–1157, 2012.

Muriel Niederle and Lise Vesterlund. Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, pages 1067–1101, 2007.

Muriel Niederle and Lise Vesterlund. Gender and competition. *Annu. Rev. Econ.*, 3(1): 601–630, 2011.

Michael I Norton, Malia F Mason, Joseph A Vandello, Andrew Biga, and Rebecca Dyer. An fmri investigation of racial paralysis. *Social cognitive and affective neuroscience*, page nss010, 2012.

Brian A Nosek and Jeffrey J Hansen. The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition & Emotion*, 22(4):553–594, 2008.

Brian A Nosek, Mahzarin Banaji, and Anthony G Greenwald. Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1):101, 2002.

Hessel Oosterbeek and Reyn Van Ewijk. Gender peer effects in university: Evidence from a randomized experiment. *Economics of Education Review*, 38:51–63, 2014.

Hyunjoon Park, Jere R Behrman, and Jaesung Choi. Causal effects of single-sex schools on college entrance exams and college attendance: Random assignment in seoul high schools. *Demography*, 50(2):447–469, 2013.

Ragan Petrie and Carmit Segal. Gender differences in competitiveness: The role of prizes. *Available at SSRN 2520052*, 2014.

Edmund S Phelps. The statistical theory of racism and sexism. *The american economic review*, pages 659–661, 1972.

Solomon William Polachek. Occupational self-selection: A human capital approach to sex differences in occupational structure. *The review of Economics and Statistics*, pages 60–69, 1981.

Joseph Price and Justin Wolfers. Racial discrimination among nba referees. *The Quarterly journal of economics*, 125(4):1859–1887, 2010.

Ernesto Reuben, Paola Sapienza, and Luigi Zingales. How stereotypes impair womens careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408, 2014.

Ernesto Reuben, Matthew Wiswall, and Basit Zafar. Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economic Journal*, 2015.

Motoko Rich. Old tactic gets new use: Public schools separate girls and boys. *New York Times*, page A14, Dec 2014.

Dan-Olof Rooth. Automatic associations and discrimination in hiring: Real world evidence. *Labour Economics*, 17(3):523–534, 2010.

Bradley J Ruffle and Richard Sosis. Cooperation and the in-group-out-group bias: A field test on israeli kibbutz members and city residents. *Journal of Economic Behavior & Organization*, 60(2):147–163, 2006.

Robert Slonim and Pablo Guillen. Gender selection discrimination: Evidence from a trust game. *Journal of Economic Behavior & Organization*, 76(2):385–405, 2010.

John L Solow and Nicole Kirkwood. Group identity and gender in public goods experiments. *Journal of Economic Behavior & Organization*, 48(4):403–412, 2002.

Matthias Sutter and Daniela Glätzle-Rützler. Gender differences in the willingness to compete emerge early in life and persist. *Management Science*, 61(10):2339–23354, 2015.

Mirco Tonin and Michael Vlassopoulos. Experimental evidence of self-image concerns as motivation for giving. *Journal of Economic Behavior & Organization*, 90:19–27, 2013.

Mirco Tonin and Michael Vlassopoulos. An experimental investigation of intrinsic motivations for giving. *Theory and decision*, 76(1):47–67, 2014.

Jennifer Triplett. Racial bias and prosocial behavior. *Sociology Compass*, 6(1):86–96, 2012.

Hal R Varian. The nonparametric approach to demand analysis. *Econometrica: Journal of the Econometric Society*, pages 945–973, 1982.

Lise Vesterlund. Why do people give. *The nonprofit sector: A research handbook*, 2:168–190, 2006.

David Wozniak, William T Harbaugh, and Ulrich Mayr. The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1):161–198, 2014.

John Yinger. Evidence on discrimination in consumer markets. *The Journal of Economic Perspectives*, pages 23–40, 1998.

VITA

Daniel Jonathan Lee was born September 29. 1987 in Evanston, IL. In the fall of 2006 he moved to Atlanta, GA to attend Emory University. He completed a BA in Math and Economics with High Honors in 2010.

In the fall of 2011 he began his doctoral studies in Economics at Georgia State University. During his time at GSU, Daniel was a CEAR Scholar, and was the recipient of a World Institute for Development Economics Research grant from the United Nations University.

Daniel's research interests include Experimental and Behavioral Economics, Public Economics, and Applied Microeconomics. He has joined Rice University as a Postdoctoral Fellow.