NONPARAMETRIC ESTIMATION AND INFERENCE IN THE PRESENCE OF SAMPLE SELECTION BIAS IN EXPERIMENTAL ECONOMICS STUDIES

by

HUIZHEN ZHONG

DANIEL J. HENDERSON, COMMITTEE CHAIR
CARY A. DECK
PAAN JINDAPON
MICHAEL K. PRICE
QIN WANG

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Economics, Finance, and Legal Studies
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2021

ABSTRACT

Experimental economics studies usually involve self-selection behaviors. In this dissertation, we explore the use of nonparametric approaches to estimate the treatment effect in these studies in the presence of sample selection bias.

The first chapter reviews the econometrics literature on nonparametric estimation of treatment effects under sample selection. Specifically, we focus on the Heckman (1979) two-step correction approach, its nonparametric extensions, and three bounding estimation approaches: Horowitz and Manski (2000), Lee (2009), and Behaghel et al. (2015). We also discuss the different estimands and the relative performance in these studies.

The second chapter explores the treatment effect of a higher match ratio on an individual's donation behavior based on evidence from a field experiment using multiple waves of email solicitations. Since donation decisions are observable only for email openers and opening rates differ between treatment and control groups, we apply the nonparametric bounding estimation approaches of Lee (2009) and Behaghel et al. (2015) to correct for selection bias when estimating the treatment effect. A higher match rate significantly increases an email opener's likelihood to give and increases the donation amount for those who contributed to the fund in the past 24 months.

The third chapter investigates whether randomized advertised show-up fees can be used as an exclusion restriction in the Heckman (1979) correction model to correct for bias caused by individuals' self-selection into lab experiment studies. We control for the actual participation fee and study the impact of the advertised show-up fee on an individual's participation decision, subject's decision making, and the treatment effects in three well-studied lab experiment tasks. We estimate these impacts using nonparametric regressions. For the range of show-up fees in our study, we find no impact on individual

participation decisions. Also, the advertised show-up fee does not affect the participant's decision-making or the treatment effect in the tasks related to individuals' social preference and risk attitude. However, the advertised show-up fee impacts subjects' strategic performance under a higher cognitive load. Therefore, caution should be made when we incorporate the randomized advertised show-up fee in the experiment design to correct for participation bias.

DEDICATION

To my dear parents, who always support me to pursue higher education and encourage me to pursue knowledge that benefits society and the world.

LIST OF ABBREVIATIONS AND SYMBOLS

AIC	Akaike information criterion
AICc	Improved Akaike information criterion
ATE	Average treatment effect for the entire population
ATE_{common}	Average treatment effect on comparable respondents
BIC	Bayesian information criterion
CDF	Cumulative distribution function
LCLS	Local-constant-least-squares
LSCV	Least-square-cross-validation

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Daniel J. Henderson, for all the valuable advice and generous support on this dissertation work and my professional development. He has been more than a thesis advisor: He is also a life mentor to me. As a graduate student, I always enjoy working with him, and I always learn how to think better while talking to him.

Besides, I would like to thank my other thesis committee members: Dr. Cary A. Deck, Dr. Paan Jindapon, Dr. Michael K. Price, and Dr. Qin Wang, for their generous support and enduring help in this work and my academic development.

My gratitude also goes out to the many other University of Alabama faculty members and staff, past and present, especially Dr. Amanda Ross, Dr. Laura Razzolini, Dr. Chris Whaley, Dr. Will Walsh, Dr. Robert Reed, Dr. Jun Ma, Dr. Matt Van Essen, Dr. Gregory Givens, Dr. Junsoo Lee, Dr. Tao Chen Yeh, Dr. Traviss Cassidy, Dr. Paul Pecorino, Dr. Byung-Cheol Kim, Ms. Debra Wheatley, Ms. Bobbie Vanlandingham, Ms. RJ Thurman, and Ms. Dixie Hamner, for all their guidance and assistance over the past several years.

Furthermore, I want to extend my gratefulness to my fellow Ph.D. students, past and present, for their friendship and support throughout this entire program, especially Fan Wu, Mia Goodnature, Tina Zhang, Emma Xu, Xiaohu Guo, Boheng Su, Lukai Yang, Andre Harrison, Jay Herndon, Jiahong Gao, Hasan Isomitdinov, Olivia Lu, Michael Soloman, Piyali Banerjee, Matthew Hampton, Alice Sheehan, Charlotte Souto, Kaveh Hasani, Shamar Stewart, Zhe Yang, Wenqiao Zhang, Haoyi Yang, Nicholas Glenn, Rushaine Goulbourne, Md Towidul Islam, Christopher Oconnor, Li Zhang, Zachary Dorobiala and Natalie Millar. In addition, I would like to thank Michael Solomon, Michael Anderson, and Zachary Dorobiala for their help in implementing the experiment.

Finally, I would like to sincerely thank my family, my other friends, and all the other people who have helped me and encourage me in this academic journey. I would not be where I am today without their support and encouragement.

Thank you.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

NONPARAMETRIC ESTIMATION OF TREATMENT EFFECTS UNDER SAMPLE SELECTION

1.1 Introduction

Sample selection is one of the most common problems that empirical economics studies try to overcome in order to identify the causal effect of interest. Even in a randomized controlled field experiment, estimates obtained by simply comparing the outcomes between treatment and control groups can be biased when the treatment can also affect the chance of outcome being observed via the survey. Empirical evidence suggests that data are rarely missing at random. Estimation of treatment effects under sample selection has become a challenge to the empirical research (Behrman et al., 2009; Behaghel and Blau, 2012).

Overcoming selection bias also plays an important role for analyzing policy impact based on program evaluation. For example, one major focus in the literature of labor market program evaluation is to estimate the wage effect of job training programs (Lee, 2009; Blanco et al., 2013). Job training (the treatment) can affect not only their wage (as a reflection of human capital), but also their decision of accepting certain kinds of jobs or not. Even the treatment is randomly assigned in the experiment, simply comparing average wages of those who participated and those who did not gives biased estimates of the average treatment effect on wages as wages are observable only for those who get hired. While it's common to use survey to collect data on program evaluation, sample selection also brings challenge to treatment effect analysis when the treatment affects individuals' likelihood to respond to the survey and the outcome in interest at the same time. Accounting for

selection bias can provide more precise information to interpret program impacts.

In this survey, we contribute to the literature by reviewing the econometric literature on nonparametric estimation methods of treatment effects under sample selection. We review the Heckman (1979) two-step correction approach as well as its nonparametric extensions and discuss the limitations of the Heckman two-step correction approach and its nonparametric extensions. We also focus on recently developed bounding techniques for average treatment effects: Horowitz and Manski (2000), Lee (2009) and Behaghel et al. (2015), as well as other approaches that further tighten the bounds by using additional assumptions or information in the data. Moreover, we discuss the different estimands in these bounding estimation approaches as well as the relative performance of the various estimators, and consider possible extensions based on these studies.

Beginning with the seminal work by Heckman (1976, 1979), there has been a large literature focused on sample selection bias and has become less and less relied on assumptions on the parametric functional form and the distribution of the error terms. Based on the well-known Heckman (1979) model, semiparametric and nonparametric approaches have been developed by replacing the linear function form in the outcome equation and sample-selection equation with unknown functions and using nonparametric estimation methods to estimate these unknown functions (e.g., Newey (2009), Ahn and Powell (1993) and Das et al. (2003)). For decades, statisticians and econometricians mainly focused on strategies that provide a bias-corrected point estimation of the treatment effect. However, these strategies still reply on the exclusion restriction assumption, i.e., the existence of exogenous variables that only determine the selection procedure but not the outcome of interest. In practice, credible instruments are hard to find in experiment data or survey data for program evaluations (Lee, 2009).

To further avoid the reliance on the exclusion restriction, Horowitz and Manski (1995, 1998, 2000) first developed the framework for constructing bounds for the average treatment effect, which also allows for treatment effect heterogeneity. Based on this

nonparametric framework, recent studies have considered different ways obtain sharper bounds on the treatment effect by imposing weaker monotonicity assumptions at the mean outcome level, using a covariate, or the number of prior calls made prior to an individual's response (e.g.,Zhang and Rubin (2003), Lee (2009) and Behaghel et al. (2015)). With less stringent assumptions compared to the parametric settings, these approaches construct bounds for the average treatment effect.

The remaining of this paper is organized as follows. The next section introduces the framework and notation and illustrates the sample selection problem. Section 1.3 reviews the Heckman (1979) model and its nonparametric extensions. Section 1.4 presents three bounding approaches for average treatment effects: Horowitz and Manski (2000), Lee (2009) and Behaghel et al. (2015), as well as other approaches which further tighten the bounds by imposing additional assumptions or using additional information in the data. In Section 1.5, we compare these three bounding estimation methods and consider possible extensions.

1.2 The sample selection problem

1.2.1 Framework and notations

Under the potential-outcome framework, we consider data on m individuals, with index $i = 1, \ldots, m$. For each individual, we observe realized values of the following triples (Y_i, R_i, T_i).

The binary variable $T_i \in \{0, 1\}$ is the treatment status: $T_i = 0$ if individual i is assigned into the control group and $T_i = 1$ if individual i is assigned into the treatment group. The observed self-selection behavior of individual i is given by

$$R_i = R_i(1)T_i + R_i(0)(1 - T_i),$$

where $R_i(0)$ and $R_i(1)$ are the potential sample selection indicators under control ($T_i = 0$)

and under treatment ($T_i = 1$), respectively. For each individual i, $R_i = 0$ if the outcome is missing (or equivalently, if individual i did not reveal or respond with their outcome) while $R_i = 1$ if the outcome is observed.

Similarly, let $Y_i(0)$ and $Y_i(1)$ be the latent outcome variables under control and under treatment respectively. Under sample selection, the observed outcome of individual i is given by

$$Y_i = Y_i(1)T_i + Y_i(0)(1 - T_i) \text{ if } R_i = 1.$$

For each individual i, either the potential outcome under control $Y_i(0)$ or the potential outcome under treatment $Y_i(1)$ will be observed if and only if the individual revealed their outcome, i.e., $R_i = 1$.

A common parameter of interest is the average treatment effect for the entire population (ATE):

$$ATE = E[Y_i(1) - Y_i(0)] = E[Y_i|T_i = 1] - E[Y_i|T_i = 0], \tag{1.1}$$

where the second equality of Equation 1.1 holds under random assignment into treatment, in which case the treatment status T_i is independent of the latent outcome variables $Y_i(0)$ and $Y_i(1)$.

1.2.2 The sample selection problem

Under random assignment, treatment status T_i is independent with the latent outcome variables $Y_i(0)$ and $Y_i(1)$. In the absence of sample selection, the average treatment effect is $ATE = E[Y_i|T_i = 1] - E[Y_i|T_i = 0]$ and can be estimated as the difference between the average outcomes in the treatment group and control group.

When sample selection arises, such as when the outcome in interest is only observable for those who respond to the survey or self-select to participate into the experiment, the observed average outcome in the treatment group and that in the control group measures the mean outcome of the treated ones that reveal their outcome $E[Y_i|R_i = 1, T_i = 1]$ and

that of the untreated counterparts $E[Y_i|R_i = 1, T_i = 0]$. Simply taking difference between the two components results in a biased estimator of the average treatment effect

$$\begin{aligned}\Big\{E[Y_i|R_i = 1, T_i = 1] - E[Y_i|R_i = 1, T_i = 0]\Big\} - E[Y_i(1) - Y_i(0)] \\ = \; E[Y_i(1) - Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(1) - Y_i(0)] \\ + \; E[Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(0)|R_i = 1, T_i = 0]. \qquad (1.2)\end{aligned}$$

The last two terms on the right-hand side of Equation 1.2 captures two types of biases from this naïve estimator. The first term $E\left[Y_i\left(1\right) - Y_i\left(0\right)|R_i = 1, T_i = 1\right] - E\left[Y_i\left(1\right) - Y_i\left(0\right)\right]$ captures treatment effect heterogeneity, which happens when those respondents cannot represent the whole population in terms of the average treatment effect[1]. The second term $E\left[Y_i\left(0\right)|R_i = 1, T_i = 1\right] - E\left[Y_i\left(0\right)|R_i = 1, T_i = 0\right]$ captures selection bias, given that treated respondents and the untreated counterparts are different in the average latent outcome under control. Such selection bias arises when treatment affects the chance to observe the outcome and individuals with better/worse outcomes are more likely to reveal their outcomes.

1.3 The Heckman (1979) model and its nonparametric extensions

1.3.1 The Heckman (1979) model

Suppose for each individual i, we also observe some exogenous pretreatment covariates X_i^o that have an impact on individual i's latent outcome Y_i^* (which equals $Y_i(1)T_i + Y_i(0)\left(1 - T_i\right)$) and some exogenous pretreatment covariates X_i^s that can determines the chance to observe individual i's outcome. In the case where we have a binary treatment status ($T_i = 1$ or 0) and where the treatment effect does not depend on

[1]When we are interested in the treatment effect in the lab experiment, the first term is referred as the participation bias. This bias arises when individual's participation decision and subject's decision making in the experiment are simultaneously affected by their characteristics (e.g., pro-social preference, individual's income level). As result, the lab participants cannot represent the whole population they come from in terms of the average treatment effect.

the exogenous pretreatment covariates, we can use the Heckman (1979) model that includes the following three equations to specify the outcome determinant process with individuals' self-selection behaviors:

$$Y_i^* = \alpha_o + X_i^{o'}\pi_o + T_i\beta + u_i \tag{1.3}$$

$$R_i = 1(\alpha_s + X_i^{s'}\pi_s + T_i\delta + v_i > 0) \tag{1.4}$$

$$Y_i = Y_i^* \text{ if } R_i = 1. \tag{1.5}$$

Equation 1.3 is the population outcome equation in which the latent outcome Y_i^* depends on the treatment status T_i and the pretreatment covariates X_i^o. Equation 1.4 is the sample-selection equation in which the sample selection indicator R_i equals 1 if the latent propensity to observe individual i's outcome $(\alpha_s + X_i^{s'}\pi_s + T_i\delta + v_i)$ is greater than 0. Note that this latent propensity to observe individual i's outcome depends on the treatment status T_i and the pretreatment covariates $X_i^{s'}$. Equation 1.5 indicates that the outcome is observed and equals the latent outcome Y_i^* if the sample selection indicator equals 1. As for the parameters, α_o, π_o and β are the parameters in the outcome equation while α_s, π_s and δ are the parameters in the sample-selection equation, respectively.

Given that $Y_i(1)$ is the latent outcome Y_i^* under treatment (when $T_i = 1$) and $Y_i(0)$ is the latent outcome Y_i^* under control (when $T_i = 0$), we have

$$ATE = E\left[Y_i\left(1\right) - Y_i\left(0\right)\right] = E\left[Y_i^*|T_i = 1 - Y_i^*|T_i = 0\right] = \beta.$$

Thus, the average treatment effect for the population is given by β, which is the parameter of interest in this model.

Simply regressing Y_i on X_i^o and T_i will result in a biased estimate of the ATE (i.e., β) since the truncated mean of the outcome is given by

$$
\begin{aligned}
E\left(Y_i|X_i^o, X_i^s, T_i\right) &= E\left(Y_i^*|X_i^o, X_i^s, T_i, R_i = 1\right) \\
&= \alpha_o + X_i^{o'}\pi_o + T_i\beta + E\left(u_i|X_i^o, X_i^s, T_i, R_i = 1\right) \\
&= \alpha_o + X_i^{o'}\pi_o + T_i\beta + E\left(u_i|v_i > -(\alpha_s + X_i^{s'}\pi_s + T_i\delta)\right),
\end{aligned}
$$

where the error terms (u_i and v_i) in the two equations are usually correlated as they usually both contain common unobservable variables that both affects the latent outcome and the chance that the latent outcome is being observed.

To obtain a consistent estimator of β, Heckman (1979) does not only specify the forms of the outcome equation and sample selection equation as linear functions of the regressors but also imposes the following assumptions on the error terms in these equations:

Assumption 1: [Bivariate normal distribution of the error terms]

The error terms (u_i, v_i) are jointly normal distributed with zero mean and independent of X_i^o, X_i^s and T_i. Moreover, $v_i \sim Normal\left(0, 1\right)$ and $E\left(u_i|v_i\right) = \gamma v_i$.

Under Assumption 1, the probability that Y_i^* is observed conditional on T_i and X_i^s is given by

$$
\Pr(R_i = 1|X_i^s, T_i) = \Phi(\alpha_s + X_i^{s'}\pi_s + T_i\delta), \tag{1.6}
$$

and the truncated mean of the outcome is then given by

$$
\begin{aligned}
E\left(Y_i|X_i^o, X_i^s, T_i\right) &= \alpha_o + X_i^{o'}\pi_o + T_i\beta + E\left(u_i|v_i > -(\alpha_s + X_i^{s'}\pi_s + T_i\delta)\right) \\
&= \alpha_o + X_i^{o'}\pi_o + T_i\beta + \gamma\lambda(\alpha_s + X_i^{s'}\pi_s + T_i\delta),
\end{aligned} \tag{1.7}
$$

where $\lambda\left(\cdot\right) = \frac{\phi(\cdot)}{\Phi(\cdot)}$ is the inverse mills ratio and $\phi\left(\cdot\right)$ and $\Phi(\cdot)$ are the standard normal density function and standard normal cumulative distribution function, respectively.

Based on the models specified in the two equations above, one can apply the following two-step correction approach proposed by Heckman (1979) to obtain a consistent estimator of the average treatment effect for the population:

1. Based on the model specified in Equation 1.6, estimate the parameters α_s, π_s and δ via the Probit regression of R_i on X_i^s and T_i and using all observations in the population. Then obtain the estimated inverse Mills ratios $\hat{\lambda}_i = \lambda(\hat{\alpha}_s + X_i^{s'}\hat{\pi}_s + T_i\hat{\delta})$ given the Probit estimates $\hat{\alpha}_s$, $\hat{\pi}_s$ and $\hat{\delta}$.

2. Based on the model specified in Equation 1.7, obtain estimates of π, β and γ via the OLS regression of Y_i on X_i^o, T_i and $\hat{\lambda}_i$ using all observations with $R_i = 1$.

Then the estimator of the average treatment effect using the two-step correction approach, $\hat{\beta}$, is consistent and asymptotically normal.

Note that X_i^o does not necessarily need to be a strict subset of X_i^s. But in practice, to avoid multicollinearity problems, there should be at least one variable in X_i^s but not in X_i^o so that the variable has an impact on the chance of observe the outcome but not directly affect the latent outcome of interest. Such a condition is usually referred as the exclusion restriction (Cameron and Trivedi, 2005).

1.3.2 The nonparametric extensions of the Heckman (1979) model

A large econometrics literature has been extending the classical two-step Heckman (1979) correction approach by relaxing the parametric model setting and assumptions on the distribution of the error terms (e.g., Newey (2009), Ahn and Powell (1993) and Das et al. (2003)). But few are found to be applicable to the setting where the average treatment effect with a binary treatment status is the main interest.

For example, Ahn and Powell (1993) consider replacing the sample selection equation as in Equation 1.4 with an unknown function form so that $E(R_i\,|T_i, X_i^s) = g(T_i, X_i^s)$ in the case with a binary treatment status, while still keeping the outcome equation with a parametric function form. They propose using the nonparametric kernel estimation method

to estimate this unknown function $g(T_i, X_i^s)$ as in the first step. As for the second step, they provide a consistent estimate of the parameters (including β) in the outcome equation via a weighted instrumental variables estimation approach where the estimated function value of $g(T_i, X_i^s)$ are used in the weighting function and variables constructed from the explanatory variables in the selection equation are used as instruments for the explanatory variables in the outcome equation.

While the Ahn and Powell (1993) approach still remains parametric setting in the outcome equation, Das et al. (2003) further extended Heckman (1979) to an "almost" nonparametric way. They replace the parametric function forms of the continuous variables in both the outcome and sample selection equations with unknown functions and use series estimations to estimate those unknown functions in both steps. Notice that their approach is not yet completely nonparametric since they still leave the discreate variables out of the unknown functions and their estimation depends on the assumption that both the unknown functions in the outcome and sample selection equations are continuously differentiable. In the case where the discrete variable (e.g., the binary treatment status) enter the two equations in a nonlinear way and interact with the continuous variables in the pretreatment covariates, the Das et al. (2003) approach might still come up with an inconsistent estimate of the average treatment effect. Future studies might consider further relax the assumption restrictions imposed by Das et al. (2003) to allow the discrete variables entering the unknown function and obtain a consistent estimator of the average treatment effect without specifying the function forms in the outcome and sample selection equations.

As in the Heckman (1979) model, the nonparametric extensions based on this model still require exclusion restrictions on the explanatory variables in the outcome and sample selection equations. As Lee (2009) indicates, variables that satisfy the exclusion restrictions may not be available in the experiment (or survey) data. For example, in the data for the Job Corps program, many of the pretreatment variables (e.g., race, gender, and education) that are significantly associated with employment (the sample selection indicator) may also

have an impact on the wages (the outcome of interest) or unobservable variables that affect the wage. This practical limitation of the Heckman (1979) model and its nonparametric extensions further motivates another approach to correct for the sample selection bias - the nonparametric bounding estimation approach.

1.4 Bounding estimations for average treatment effects

1.4.1 Horowitz and Manski (1995, 1998, 2000) bounds

Horowitz and Manski (1995, 1998, 2000) first propose a nonparametric framework to construct bounds for the average treatment effect. They consider the worst-case scenario, which is for the missing observations, and allow for treatment effect heterogeneity. Moreover, they do not impose untestable assumptions like data missing-at-random assumptions or other distributional assumptions or exclusion restrictions as in the parametric framework. In the context that only the outcome variable is missing due to sample selection, the key of Horowitz and Manski (1995, 1998, 2000)'s bounding estimation approach is to simply impose a bounded support assumption for the latent outcome $Y_i(0)$ and $Y_i(1)$. Unbiased bounded estimates of the average treatment effect can be obtained by replacing the unobservable mean outcome with the upper and lower bounds of this support.

Based on the Horowitz and Manski (1995, 1998, 2000) bounding estimation approach, one can decompose the ATE into four components conditional on the treatment status T_i and the sample-selection status R_i:

$$\begin{aligned} ATE &= E[Y_i|T_i=1] - E[Y_i|T_i=0] \\ &= E[Y_i|T_i=1, R_i=1]\Pr(R_i=1|T_i=1) + E[Y_i(1)|T_i=1, R_i=0]\Pr(R_i=0|T_i=1) \\ &\quad - E[Y_i|T_i=0, R_i=1]\Pr(R_i=1|T_i=0) - E[Y_i(0)|T_i=0, R_i=0]\Pr(R_i=0|T_i=0), \end{aligned}$$

where the sum of the first two terms is the mean outcome of the treated population $E[Y_i|T_i=1]$, while the sum of the last two terms is the mean outcome of the untreated

population $E[Y_i|T_i = 0]$. From the data, we can identify the proportions of sample-selection in each group that represents the conditional probabilities $\Pr(R_i = r|T_i = t), r, t \in (0, 1)$ and the mean outcome of those respondents (those who reveal their outcomes) $E[Y_i|T_i = 1, R_i = 1]$ and $E[Y_i|T_i = 0, R_i = 1]$. We, unfortunately, are unable to precisely identify the mean outcomes conditional on $R_i = 0$ (i.e., nonrespondents or those who did not reveal their outcomes): $E[Y_i(1)|T_i = 1, R_i = 0]$ and $E[Y_i(0)|T_i = 0, R_i = 0]$, because of sample selection.

Horowitz and Manski (1995, 1998, 2000) solve this problem by imposing the following assumption:

Assumption 2: [Bounded support assumption]

The latent outcomes $Y_i(0)$ and $Y_i(1)$ have common support that lies on the interval $[y^{lb}, y^{ub}]$.

Bounds for ATE can be obtained via replacing the unobservable mean outcomes conditional on $R_i = 0$ with the largest and smallest possible values of the latent outcomes, *i.e.*, y^{lb} and y^{ub}:

$$\begin{aligned}
\Delta^{lb} &= E[Y_i|T_i = 1, R_i = 1] \Pr(R_i = 1|T_i = 1) + y^{lb} \Pr(R_i = 0|T_i = 1) \\
&\quad - E[Y_i|T_i = 0, R_i = 1] \Pr(R_i = 1|T_i = 0) - y^{ub} \Pr(R_i = 0|T_i = 0) \\
\Delta^{ub} &= E[Y_i|T_i = 1, R_i = 1] \Pr(R_i = 1|T_i = 1) + y^{ub} \Pr(R_i = 0|T_i = 1) \\
&\quad - E[Y_i|T_i = 0, R_i = 1] \Pr(R_i = 1|T_i = 0) - y^{lb} \Pr(R_i = 0|T_i = 0).
\end{aligned}$$

The estimators of the lower and upper bounds for the average treatment effect on the entire population consist of replacing population parameters with sample analogs.

Although this nonparametric estimation approach has the advantage that it only requires outcomes are bounded, recent studies (Lee, 2009; Behaghel and Blau, 2012; Behaghel et al., 2015) argue that, the Horowitz and Manski (1995, 1998, 2000) bounds can

be too wide to be informative when the survey attrition rate is high (e.g. above 15%). Even with its flaws, it still works as a benchmark in the nonparametric bounding estimation for treatment effects.

1.4.2 Lee (2009) bounds

The general idea of Lee (2009) is to construct comparable subsamples of respondents in the treatment and control groups when the two groups have unequal response rates. In order to do so, reasonable assumptions should be imposed to identify the marginal respondents. Sharp bounds can be obtained by trimming the upper and lower tails of the outcome for the marginal respondents.

To provide more precise information and tighten the bounds for the average treatment effect, Lee (2009) imposes a monotonicity assumption in addition to the independence assumption:

Assumption 3. [Independence]

$$Y_i(1),\ Y_i(0),\ S_i(1) \text{ and } S_i(0) \text{ are independent of } T_i.$$

Assumption 4a: [Monotonicity on Sample Selection]

$$\text{For all individual } i,\ R_i(1) \geq R_i(0) \text{ with probability 1,}$$

or

Assumption 4b: [Monotonicity on Sample Selection][2]

$$\text{For all individual } i,\ R_i(1) \leq R_i(0) \text{ with probability 1.}$$

Assumption 3 holds under random assignments of the treatment status. Assumption 4a or 4b, the monotonicity assumption, implies that the treatment effect will affect the sample selection behavior of all individuals in the same direction. Suppose the treatment group

[2]Assumption 4a holds when the treatment group has a higher response rate, while Assumption 4b holds when the control group has a higher response rate.

has a higher response rate, Assumption 4a will be adopted and it implies that all treated individuals should have at least the same willingness to reveal the outcome than if they were not treated. Based on the same rationale, Assumption 4b will be imposed in the case where the control group has a higher response rate. Although this monotonicity assumption is fundamentally untestable, it serves as a reasonable assumption when the two groups have unequal response rates. And it's important to note that the Lee (2009) bounding estimation approach does not rely on the bounded support assumption adopted by Horowitz and Manski (1995, 1998, 2000).

Lee (2009) focuses on the average treatment effect on individuals who will reveal their outcomes regardless of their treatment status

$$ATE_{common} = E\left[Y_i\left(1\right) - Y_i\left(0\right)|R_i\left(1\right) = 1, R_i\left(0\right) = 1\right].$$

Notice that this estimand is slightly different from the parameter of interest in Horowitz and Manski (1995, 1998, 2000), where they focus on the average treatment effect on the whole population.

Suppose the treatment has a higher response rate, then

$$\begin{aligned} ATE_{common} &= E[Y_i(1)|R_i(1) = 1, R_i(0) = 1] - E[Y_i(0)|R_i(1) = 1, R_i(0) = 1] \\ &= E[Y_i(1)|R_i(1) = 1, R_i(0) = 1] - E[Y_i|R_i = 1, T_i = 1], \end{aligned}$$

where the focus is the estimation of the term $E\left[Y_i\left(1\right)|R\left(1\right) = 1, R_i\left(0\right) = 1\right]$, the mean outcome of the comparable sub-sample in the group with the higher response rate. Suppose the outcome in interest is a continuous variable, the Lee (2009) bounding approach can be via the following three-step procedure:

1. Calculate the share of marginal respondents as the difference between the share of respondents in the treatment group and that in the control group. Then the share of

marginal respondents among the treated respondents (p) is given by:

$$p = \frac{\Pr(R_i = 1|T_i = 1) - \Pr(R_i = 1|T_i = 0)}{\Pr(R_i = 1|T_i = 0)}.$$

2. Find out the p quantile and $1 - p$ quantile of the treated respondents: $y_p = G_Y^{-1}(p)$ and $y_{1-p} = G_Y^{-1}(1-p)$, where G_Y is the cumulative distribution function (CDF) of Y_i conditional on $T_i = 1$ and $R_i = 1$.

3. The lower (upper) bound of the average treatment effect on the comparable respondents (ATE_{common}) is the difference of the average of the lower (upper) $1 - p$ portion of the outcomes of the treated respondents and the average outcome of the controlled respondents:

$$\Delta_1^{lb} = E(Y_i|R_i = 1, T_i = 1, Y_i \leq y_{1-p}) - E(Y_i|R_i = 1, T_i = 0)$$
$$\Delta_1^{ub} = E\left(Y_i|R_i = 1, T_i = 1, Y_i \geq y_p\right) - E\left(Y_i|R_i = 1, T_i = 0\right).$$

The estimators of Lee (2009) bounds can be obtained by replacing population parameters with sample analogs.

1.4.3 Behaghel et al. (2015) bounds

Behaghel et al. (2015) consider the case where survey data are obtained by multiple attempts (e.g., via phone calls). Here sample selection arises when some individuals do not respond to the survey and leave the outcome missing. The sample selection indicators R_i can be interpreted as the response to the survey.

As in Lee (2009), Behaghel et al. (2015) try to address the selection bias when estimating the average treatment effect for those who will respond to the survey regardless of their treatment status. To tighten the bounds for the average treatment effect, they extend the Lee (2009) bounding estimation approach and construct estimators based on the latent variable threshold-crossing response model. The key insight of their approach is

to use additional information, the number of attempts made before the individual give their response. This results in a narrower range of the outcome of the truncated marginal respondents and tighter bounds on the estimated average treatment effect as compared to those obtained by Lee (2009).

In their framework, sample selection arises when the outcome of interest is correlated with the individual's willingness to respond, which is unobservable in the data. For example, consider a survey on the impact of a job training program on employment status. Those who get employed are less likely to respond to the survey as they are more likely to be occupied by their job and not able to accept the survey via phone calls.

To illustrate their approach, we first introduce additional notation. Define W_i as the total number of attempts made to reach individual i and N_i the number of attempts made before the individual i gives their response. When there exists detailed information of survey on whether the individual has received the attempt and whether they provide a response for each attempt to reach the individuals, we observe realized values of the variables (Y_i, R_i, T_i, N_i) for each individual i.

Denote V_i (unobservable) as the individual's unwillingness to respond. In addition to the monotonicity assumption adopted by Lee (2009), Behaghel et al. (2015) impose the following assumptions:

Assumption 5: [Latent Variable Threshold-Crossing Response Model]

$$R_i = 1(V_i < p(W_i, T_i)), \forall i,$$

where $p(W_i, 0)$ and $p(W_i, 1)$ have nonempty common support as W_i varies.

Assumption 6: [Independence Between Treatment Status and Willingness to Response]

$$V_i \perp T_i, \forall i.$$

Assumption 7: [Same Maximum Potential Attempts to Obtain Individual's Response]

$$W_i = w_{\max}, \text{ where } w_{\max} \text{ is constant}, \forall i.$$

Assumption 8: [Non-decreasing in Response Rate]

$$p(W, Z) \text{ is non-decreasing in } W, \forall Z.$$

Assumption 5 indicates the individual will only respond ($R_i = 1$) when the incentive to respond $p(W_i, T_i)$ exceeds the individual's reluctance to respond V_i. While Assumption 6 implies that the distribution of the individual's willingness to respond should be the same in the control and treatment groups. Thus, the individual from the control group and the one from the treatment group that have the same relative rank of V_i represent comparable pairs. Assumption 7 implies that all individuals in the data should potentially receive the same number of attempts $w_{\max}$. Assumption 8 implies that more attempts to approach the individual will not decrease the response rate.

Suppose the treatment group has a higher response rate and consider the case where the outcome variable is continuous, the Behaghel et al. (2015) bounding approach can be via the following four-step procedure:

1. Find the number of attempts w_1 in the treatment group so that the response rate after the w_1^{th} attempt in the treatment group is closest and higher than the overall response rate in the control group:

$$w_1 = \min_{n \in \{1,\ldots,w_{max}\}} : \Pr(N_i \leq n | T_i = 1) \geq \Pr(R_i = 1 | T_i = 0).$$

2. Identify the share of marginal respondents among the treated individuals that respond after the w_1^{th} attempt but did not respond after the first $w_1 - 1$ attempts:

$$\alpha = \frac{\Pr(N_i \leq w_1 | T_i = 1) - \Pr(R_i = 1 | T_i = 0)}{\Pr(N_i \leq w_1 | T_i = 1) - \Pr(N_i \leq w_1 - 1 | T_i = 1)}.$$

3. Find the α quantile and $1 - \alpha$ quantiles of the treated respondents: $y_\alpha = G_Y^{-1}(\alpha)$ and $y_{1-\alpha} = G_Y^{-1}(1 - \alpha)$, where G_Y is the CDF of Y conditional on $T_i = 1$ and $N_i = w_1$.

4. The lower and upper bounds of the mean outcome of the comparable treated respondents can be derived as the weighted sum of the average outcome of the treated one that respond after the first $w_1 - 1$ attempts ($E\left(Y_i | N_i \leq w_1 - 1, T_i = 1\right)$) and the average of the lower and upper $1 - \alpha$ portion of the outcomes of the treated respondents at the w_1^{th} attempt ($E\left(Y_i | N = w_1, T_i = 1, Y_i \leq y_{1-\alpha}\right)$and $E\left(Y_i | N = w_1, T_i = 1, Y_i \geq y_{\alpha}\right)$), weighted by their shares relative to the controlled response rate.

   If we take the difference between the lower (upper) bound of the mean outcome of the comparable untreated respondents and the average outcome of the treated respondents, the lower bound of the average treatment effect (ATE_{common}) is then given by

$$\begin{aligned}\Delta_2^{ub} = & \frac{\Pr(N_i \leq w_1 | T_i = 1)}{\Pr((R_i = 1 | T_i = 0)} E\left(Y_i | N_i \leq w_1 - 1, T_i = 1\right) \\ & + \left(1 - \frac{\Pr(N_i \leq w_1 | T_i = 1)}{\Pr(R_i = 1 | T_i = 0)}\right) E\left(Y_i | N = w_1, T_i = 1, Y_i \leq y_{1-\alpha}\right) \\ & - E\left(Y_i | R_i = 1, T_i = 0\right),\end{aligned}$$

   and the upper bound of the average treatment effect (ATE_{common}) is given by

$$\begin{aligned}\Delta_2^{lb} = & \frac{\Pr(N_i \leq w_1 | T_i = 1)}{\Pr(R_i = 1 | T_i = 0)} E\left(y | N \leq w_1 - 1, T_i = 1\right) \\ & + \left(1 - \frac{\Pr(N_i \leq w_1 | T_i = 1)}{\Pr(R_i = 1 | T_i = 0)}\right) E\left(Y_i | N = w_1, T_i = 1, Y_i \geq y_{\alpha}\right) \\ & - E(Y_i | R_i = 1, T_i = 0).\end{aligned}$$

The estimators of Behaghel et al. (2015) bounds can be obtained by replacing population parameters with sample analogs.

1.4.4 Other bounding estimation approaches

Other approaches have considered alternative ways or used different assumptions to narrow the bounds. Zhang and Rubin (2003) impose a weak monotonicity assumption of mean potential outcome across strata and Lee (2009) use a pretreatment covariate to construct tighter bounds.

The general idea of Zhang and Rubin (2003) can be illustrated as follows. Suppose the researchers can infer that the outcome in interest is positively (or negatively) correlated with individual's self-selection behavior (e.g., those who have more human capital and thus higher wage are more likely to be employed; those who have better outcomes are more likely to respond to the survey), then we can estimate the bounds for the average treatment effect with a weak monotonicity assumption on the mean potential outcome across strata in addition to Assumption 4a and 4b.

Assumption 9: [Weak Monotonicity of Mean Potential Outcome[3]]

$$E\left[Y_i(1)|\, R_i\left(0\right)=1, R_i\left(1\right)=1\right] \geq E(Y_i(1)|R_i\left(0\right)=0, R_i\left(1\right)=1)$$

What Assumption 9 implies is straightforward: If the treatment group has a higher response rate and suppose those who are more likely to reveal their outcome tend to have better outcomes, it is natural to expect that the mean outcome of those who will respond regardless of treatment status should be higher than the mean outcome of those who will respond only if they are treated. As Blanco et al. (2013) point out, although this assumption is directly untestable as Assumption 4a and 4b, one can tell if it's reasonable to impose this assumption by gauging its implication: If this assumption holds, then those who will respond regardless of treatment status should have better average pretreatment

[3]Here we impose this weak monotonicity assumption in the case where the treatment group has a higher response rate and the outcome in interest is positively correlated with individual's self-selection behavior. Similarly, we can adjust the assumption when the control group has a higher response rate or when the outcome in interest is negatively correlated with individual's self-selection behavior.

characteristics that is positively correlated with the outcome in interest.

To understand how this approach works, recall that:

$$ATE_{common} = E\left[Y_i\left(1\right)|R_i\left(1\right)=1, R_i\left(0\right)=1\right] - E\left[Y_i|R_i=1, T_i=0\right].$$

Under Assumption 9, the mean outcome of the treated respondents ($E\left(Y_i|R_i=1,T_i=1\right)$) is bounded below by the mean outcome of those who will respond regardless of treatment status ($E\left[Y_i\left(1\right)|R_i\left(1\right)=1, R_i\left(0\right)=1\right]$). [4] The upper bound for ATE_{common} remains the same as in Lee (2009). The lower bound for ATE_{common} can be estimated simply by calculating the difference between the observed outcome in the treatment group and those in the control group and it's given by:

$$\Delta_3^{lb} = E\left(Y_i|R_i=1,T_i=1\right) - E\left[Y_i|R_i=1,T_i=0\right].$$

Lee (2009) also shows that it's also practical to use pretreatment covariates to further narrow the bounds for the treatment effect.

The general idea of this approach is straightforward. Consider the case where survey data also contains a pretreatment covariate X_i (e.g., the demographic characteristics of the sample like race, gender, age, etc.). Then we can split the sample into groups based on the value of the pretreatment covariate X_i. Under random assignment, we can assume that $X_i \perp T_i$, $\forall i$. Let x_j, with index $j = 1, \ldots, J$ be the values that X_i can take. Even if X_i is a continuous variable (e.g., age), we can still split the sample based on the interval that the values of X_i fall into. Then x_j implies group J. The proportion of each group among the entire sample is given by $\Pr(X_i = x_j)$.

Suppose the treatment group has a higher response rate, one can apply procedure in the Lee (2009) bounding approach (as in Section 1.4.3) to calculate the share of marginal

[4] Under Assumption 9, we have $E\left[Y_i\left(1\right)|R_i\left(1\right)=1, R_i\left(0\right)=1\right] \geq E\left(Y_i|R_i=1,T_i=1\right)$.

respondents among the treated respondents (p_j)

$$p_j = \frac{\Pr(R_i = 1|T_i = 1, X_i = x_j) - \Pr(R_i = 1|T_i = 0, X_i = x_j)}{\Pr(R_i = 1|T_i = 0, X_i = x_j)}, \tag{1.8}$$

and bounds on the average treatment effect conditional on the comparable respondents for each group j (Δ_j^{lb} and Δ_j^{ub})

$$\begin{aligned}
\Delta_j^{lb} &= E(Y_i|R_i = 1, T_i = 1, Y_i \leq y_{1-p_j}, X_i = x_j) - E(Y_i|R_i = 1, T_i = 0, X_i = x_j \\
\Delta_j^{ub} &= E\left(Y_i|\, R_i = 1, T_i = 1, Y_i \geq y_{1-p_j}, X_i = x_j\right) - E(Y_i\,|R_i = 1, T_i = 0, X_i = x_j)\,,
\end{aligned}$$

where $y_q = G_j^{-1}(q)$ and G_j is the CDF of Y_i conditional on $T_i = 1$ and $X_i = x_j, \forall j = 1, \ldots, J$.

Then sharper bounds for ATE_{common} can then be obtained by averaging the bounds for each group with weights, where the weight for each group equals to the share of the group in terms of the covariate ($\Pr(X_i = x_j)$):

$$\begin{aligned}
\Delta_4^{lb} &= \sum_{j=1}^{J} E(Y_i|R_i = 1, T_i = 1, Y_i \leq y_{1-p_j}, X_i = x_j)P(X_i = x_j) - E(Y_i|R_i = 1, T_i = 0) \\
\Delta_4^{ub} &= \sum_{j=1}^{J} E(Y_i|R_i = 1, T_i = 1, Y_i \geq y_{p_j}, X_i = x_j)P(X_i = x_j) - E(Y_i|R_i = 1, T_i = 0),
\end{aligned}$$

where $\Delta_4^{lb} \geq \Delta_1^{lb}$ and $\Delta_4^{ub} \leq \Delta_1^{ub}$.

1.5 Conclusion

In this paper, we review the Heckman (1979) two-step correction approach, its nonparametric extensions and recent nonparametric bounding approaches for estimating average treatment effects under sample selection. Without imposing parametric functional form assumptions or assumptions on the distribution of missing data, these nonparametric bounding approaches can still provide informative analysis on the treatment effect if the

estimated bounds are narrow enough. As sample selection problem is a prevalent challenge to policy analysis and program evaluation, recent empirical studies have noticed the importance of these approaches and applied these approaches to correct for sample selection in their context. For example, Lee (2009) bounds have been widely applied to social experiments, policy analysis and program evaluations, for robust estimation or robustness check (Glewwe et al. (2010); De Mel et al. (2008); Finkelstein et al. (2012).

Though Lee (2009) bounds are narrower than the Horowitz and Manski (1995, 1998, 2000) bounds, they estimate different objects. Caution should be paid when trying to interpret the results based on these bounding estimation approaches, especially when the sample selection rate is high. In this case, the estimated average treatment effect for those who will reveal their outcomes regardless of the treatment status may not well represent the average treatment effect for the entire population (which can be what policy makers actually care about). That being said, estimates from the Lee (2009) bounding approach are typically more informative than the ones proposed by Horowitz and Manski (1995, 1998, 2000) as sharper bounds can be achieved.

Extensions can be made based on these existing bounding approaches. For example, we can also use the pretreatment covariate to further narrow the Behaghel et al. (2015) bounds in a way that's similar to the Lee (2009) grouping approach. Future empirical analysis will benefit from these nonparametric bounding estimation approaches and the extension based upon them for estimating the average treatment effect of policy programs under sample selection.

CHAPTER 2

MATCH RATE AND CHARITABLE GIVING: EVIDENCE FROM A FIELD EXPERIMENT USING EMAIL SOLICITATIONS (WITH DANIEL J. HENDERSON AND MICHAEL K. PRICE)

2.1 Introduction

Charitable giving has been an important component in the US economy. In 2019, individuals and other organizations in the US contributed to various charity causes an estimated 450 billion dollars, which exceeded 2% of GDP over the same year. Based on a report from the Giving USA Foundation, the majority, as large as 69%, of overall charitable dollars came from giving by individuals (Giving USA Foundation (2020)). Still, relatively little is known about what motivates people to give. From the perspective of charitable organizations, understanding how individual donors respond to donation incentives can help them better understand the effectiveness of various fundraising mechanisms and raise the most for each dollar spent.

As one of the most popular fundraising strategies used by practitioners, matching gifts [1] has been the focus of many studies in the charitable giving literature. The theoretical warm glow model proposed by Andreoni (1989, 1990) suggests that providing matching gifts lowers the price of providing donations as a public good and thus increases demand of the public good. Increasing the match rate should lead to increased donation participation as it lowers the cost of giving the first dollar. However, little is known about the effect on dollars given by donors.

[1] An example of matching gift can be providing seed money or dollar matching with a fixed match ratio.

The existing empirical literature has explored two common types of fund-raising designs. Natural field experiments examined the effect of providing seed money (List and Lucking-Reiley (2002), Frey and Meier (2004), Rondeau and List (2008)). Results in those studies suggest increases in a matching grant greatly increase both the donation rate and the average donation amount from donors. Studies focusing on dollar matching mechanisms with fixed match ratios (Karlan and List (2007) Huck and Rasul (2011), Karlan et al. (2011), Huck et al. (2015)) found that matching grants work largely along the extensive margin and draw in new donors. However, in those studies, donors do not seem to respond to the change in match ratios (whether up or below the baseline 1:1).

In this digital era, e-mail solicitations and social media have become a common way to raise funds as compared to traditional direct mail solicitations. Sending e-mails to potential donors lowers costs and allows fundraisers to make more attempts to reach potential donors. Previous studies have used direct mail solicitation to examine the price effect of the dollar matching mechanism in fundraising. It remains to be seen whether the conclusions still hold when we use online tools for fundraising.

In this paper, we re-examine whether a higher match ratio affects individuals' donation participation and their donation amount. To answer this question, we provide evidence from a field experiment driven by five waves of e-mail solicitation for the Environmental Defense Fund (EDF) in 2013. Besides the intent-to-treat effect, we also examine the causal impact of a higher match rate on the intensive margin of giving average treatment effect when the match rate increases from 1:1 to 2:1.

Our contribution to the literature is twofold. We contribute to the literature on charitable giving by providing evidence from a field experiment with a different solicitation method. Our paper is also the first to apply the Behaghel et al. (2015) bounding estimation approach to the case where the outcome variable is continuous. It aims to reduce sample selection bias caused by non-response. In our dataset, e-mail response rates differ between treatment and control groups because the subject line of the e-mail included the match

rate. Hence, in order to reduce the sample selection bias caused by differing e-mail response rates (not to mention the fact that donation decisions are observable only for those who opened at least one e-mail), we apply the Lee (2009) and Behaghel et al. (2015) bounding estimation approaches to bound the average treatment effect of the higher match rate.

In contrast to the findings in previous studies on the effect of a higher match rate of the dollar matching mechanism, we find that increasing the match rate brings more donors and more dollars given on average. We also find that increasing the match rate reduced the e-mail open rate. The estimated Behaghel et al. (2015) bounds show that the estimated treatment effect on likelihood to give and donation amount are bounded between 1.37 and 2.58 percent, and 38 and 67 cents, respectively. By exploiting the additional information of the number of e-mails sent, the Behaghel et al. (2015) bounds are able to provide a more narrowly estimated interval as compared to the Lee (2009) bounds.

The reminder of the paper proceeds as follows. Sections 2.2 and 2.3 illustrate the sample selection problem and bounding estimation approaches, respectively. Section 2.4 presents the field experiment. Section 2.5 provides the results and the final section concludes. Appendix A includes a formal proof as well as additional results.

2.2 Sample selection

2.2.1 Framework and notations

Under the potential-outcome framework, we consider data on m individuals, with index $i = 1, \ldots, m$. For each individual i, we observe realized values of the triples (Y_i, R_i, T_i).

The binary variable $T_i \in \{0, 1\}$ is the treatment status: $T_i = 0$ if individual i is randomly assigned to the control group and $T_i = 1$ if individual i is randomly assigned to the treatment group that provides a higher match rate as compared to the control group. To capture the fact that not all individuals opened at least one e-mail, and hence by our definition, gave a response to the charitable giving campaign, the observed self-selection

behavior of individual i is given by

$$R_i = R_i(1)T_i + R_i(0)(1 - T_i),$$

where $R_i(0)$ and $R_i(1)$ are the indicators for whether or not they opened at least one e-mail under control ($T_i = 0$) and under treatment ($T_i = 1$), respectively. In other words, for each individual i, $R_i = 0$ when individual i did not open a single e-mail, while $R_i = 1$ if the individual "responded" by opening at least one e-mail.

Let $Y_i(0)$ and $Y_i(1)$ are the latent outcome variables regarding individual i's final donation decision after multiple e-mails under control ($T_i = 0$) and under treatment ($T_i = 1$), respectively. The observed final donation decision of individual i after 5 waves of e-mails is given by

$$Y_i = Y_i(1)T_i + Y_i(0)\,(1 - T_i) \text{ if } R_i = 1.$$

Note for each individual i, either the potential outcome under control, $Y_i(0)$, or the potential outcome under treatment, $Y_i(1)$, will be observed if and only if the individual opened at least one e-mail and gave a response ($R_i = 1$).

In this setting, a common parameter of interest is the average treatment effect for the entire population (ATE):

$$ATE = E\big[Y_i(1) - Y_i(0)\big] = E\big[Y_i|T_i = 1\big] - E\big[Y_i|T_i = 0\big], \tag{2.1}$$

where, the second equality of Equation 2.1 holds if the treatment status T_i is independent of the latent outcome variables $Y_i(0)$ and $Y_i(1)$. This is a reasonable assumption under random assignment.

2.2.2 Biases

What Equation 2.1 implies is straightforward. If all potential donors opened at least one e-mail (i.e., $R_i = 1, \forall i$), the average treatment effect (ATE) can be estimated as the

difference between the average outcomes in the treatment and control groups. However, the final donation decisions are only observable for those who opened at least one e-mail. The average observed outcome in the treatment group measures the mean outcome of the treated respondents $E[Y_i|R_i = 1, T_i = 1]$ and the average observed outcome in the control group measures the mean outcome of their untreated counterparts $E[Y_i|R_i = 1, T_i = 0]$.

Simply taking the difference between these two components results in a biased estimator of the average treatment effect, where the bias equals

$$\begin{aligned}
&\Big\{E[Y_i|R_i = 1, T_i = 1] - E[Y_i|R_i = 1, T_i = 0]\Big\} - E[Y_i(1) - Y_i(0)] \\
&\qquad = \; E[Y_i(1) - Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(1) - Y_i(0)] \\
&\qquad + \; E[Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(0)|R_i = 1, T_i = 0]. \qquad (2.2)
\end{aligned}$$

The last two terms on the right-hand side of Equation 2.2 capture two separate biases from this naive estimator. The first term $E[Y_i(1) - Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(1) - Y_i(0)]$ captures treatment effect heterogeneity, which happens when those e-mail respondents cannot represent the entire population in terms of the average treatment effect. The second term $E[Y_i(0)|R_i = 1, T_i = 1] - E[Y_i(0)|R_i = 1, T_i = 0]$ captures selection bias, given that the treated and untreated e-mail respondents have different average latent outcomes under control.

Such selection bias could arise in this context when treatment affects an individual's chance of revealing their donation decision via opening at least one e-mail and when an individual's donation decision (likelihood to give and how much to give) is correlated with their willingness to open any e-mail to reveal their donation decision. In our application, the subject line of the solicitation e-mails included the match rate and led to a difference in the e-mail open rates between treatment and control groups.

2.3 Bounding

Selection bias in our setting makes point identification impossible. This requires us to consider the bounding estimation approaches of Lee (2009) and Behaghel et al. (2015). The general idea behind these bounding estimation approaches is to construct the largest comparable sub-samples of respondents in the treatment and control groups when the two groups have different response rates. The two sub-samples will then be considered comparable in the sense that they contain individuals who will respond regardless of their treatment status. Given that we are comparing sub-samples, our the parameter of interest will be restricted to be the average treatment effect on comparable respondents[2]

$$ATE_{common} = E[Y_i(1) - Y_i(0)|R_i(1) = 1, R_i(0) = 1].$$

The comparable sub-samples are illustrated in Figure 1. In our application, the control group has a higher response rate. In this case, the parameter of interest measures the difference between the mean outcome of the treated respondents ($E[Y_i|R_i = 1, T_i = 1]$) and the mean outcome of the comparable respondents in the control group who would also respond if they were in the treated group ($E[Y_i(0)|R_i(1) = 1, R_i(0) = 1]$):

$$\begin{aligned} ATE_{common} &= E[Y_i(1)|R_i(1) = 1, R_i(0) = 1] - E[Y_i(0)|R_i(1) = 1, R_i(0) = 1] \\ &= E[Y_i|R_i = 1, T_i = 1] - E[Y_i(0)|R_i(1) = 1, R_i(0) = 1]. \end{aligned}$$

Our goal is to derive the mean outcome of the comparable sub-sample in the group with the higher response rate ($E[Y_i(0)|R_i(1) = 1, R_i(0) = 1]$ in this case). The group with the higher response rate (i.e., the control group in our application) consists of two types of respondents: comparable respondents and marginal respondents. Comparable respondents are those who would respond regardless of their treatment status; while marginal

[2]Note that this estimand is different from the parameter of interest in Horowitz and Manski (2000), where they focus on the average treatment effect for the entire population.

respondents are those who would not respond if they were assigned to the other group. Bounds arise as the marginal respondents cannot be perfectly identified and trimmed away. For the remainder of the paper, we refer to ATE_{common} as the average treatment effect of a higher match rate on the donation decision (donation participation or donation amount), conditional on comparable respondents.

2.3.1 Lee (2009) bounds

In the presence of sample selection bias, Lee (2009) imposes a monotonicity assumption to identify the share of comparable respondents (or equivalently, the share of marginal respondents) and tighten bounds on ATE_{common}:

Assumption 1. [Independence]

$Y_i(1)$, $Y_i(0)$, $R_i(1)$ and $R_i(0)$ are independent of T_i.

Assumption 2a. [Monotonicity on Sample Selection]

For each individual i, $R_i(1) \leq R_i(0)$ with probability 1,

or

Assumption 2b. [Monotonicity on Sample Selection]

For each individual i, $R_i(1) \leq R_i(0)$ with probability 1.

Assumption 1 holds under random assignments in experiments. Assumption 2a is imposed in our application since the control group has higher response rate, while Assumption 2b will be imposed in the case where the treatment group is observed with a higher response rate. Either assumption implies that the treatment will affect the self-selection behavior of all individuals in the same direction. In the context of our application, this implication means any individual in the control group should have at least the same willingness to respond if they were in the treatment group. Although this

monotonicity assumption is not testable, it is a reasonable assumption when the two groups have different response rates.[3]

When employing the monotonicity assumption, the share of comparable respondents equals the share of respondents in the group with the lower response rate. The marginal respondents can thus be identified as the portion of the respondents in the group with a higher response rate (again, see Figure 1). By trimming away the upper or lower tail of the outcome by this portion for the group with the higher response rate, we can obtain the mean outcome of the comparable sub-sample in the group with a higher response rate and obtain sharp bounds on ATE_{common}.

2.3.1.1 Parameter of interest

When we are interested in the average treatment effect on the donation amount given by individuals, the outcome variable Y_i is a continuous variable that indicates the dollar amount given by individual i. Suppose the monotonicity assumption holds in our application where the control group has a higher response rate, we can follow the Lee (2009) bounding approach to derive the bounds for the average treatment effect via the following three-step procedure:

1. Calculate the share of marginal respondents as the difference between the share of respondents in the control group and that in the treatment group. Then the share of marginal respondents among the untreated respondents (p) is given by:

$$p = \frac{\Pr(R_i = 1|T_i = 0) - \Pr(R_i = 1|T_i = 1)}{\Pr(R_i = 1|T_i = 0)}.$$

2. Find the p quantile and $1 - p$ quantile of the control respondents' outcome distribution: $y_p = G_Y^{-1}(p)$ and $y_{1-p} = G_Y^{-1}(1 - p)$, where G_Y is the cumulative

[3]It is important to note that the Lee (2009) bounding estimation approach does not rely on the bounded support assumption (see, Horowitz and Manski (1995, 1998, 2000)).

distribution function (CDF) of Y_i conditional on $T_i = 0$ and $R_i = 1$.

3. Denote $\underline{y}_{CR}$ ($\overline{y}_{CR}$) as the lower (upper) bound of the mean outcome of the comparable treated respondents $E[Y_i(0)|R_i(1) = 1, R_i(0) = 1]$. It can be identified as the mean of the lower (upper) $1 - p$ portion of the outcomes of the control respondents:

$$\underline{y}_{CR} = E[Y_i|R_i = 1, T_i = 0, Y_i \leq y_{1-p}]$$
$$\overline{y}_{CR} = E[Y_i|R_i = 1, T_i = 0, Y_i \geq y_p].$$

The lower and upper bounds of the average treatment effect on the comparable respondents (ATE_{common}) are then given as

$$\Delta_1^{lb} = E[Y_i|R_i = 1, T_i = 1] - E[Y_i|R_i = 1, T_i = 0, Y_i \geq y_p].$$
$$\Delta_1^{ub} = E[Y_i|R_i = 1, T_i = 1] - E[Y_i|R_i = 1, T_i = 0, Y_i \leq y_{1-p}]$$

When we are interested in the average treatment effect of an individual's likelihood to give, the outcome variable Y_i is a binary variable that equals 1 if individual i donates and 0 otherwise. Similar to the continuous case, we can follow a three-step procedure to derive bounds for the average treatment effect. The first step is the same as in the continuous case, but we must modify Steps 2 and 3:

2. Calculate the share of donors among the untreated respondents(γ)

$$\gamma = \Pr(Y_i = 1|R_i = 1, T_i = 0).$$

3. The lower and upper bounds of the mean outcome of the comparable untreated respondents ($\underline{y}_{CR}$ and $\overline{y}_{CR}$) depend on the share of the untreated respondents who donated (γ) and the share of marginal respondents among the untreated respondents

(p). The bounds are given as

$$\underline{y}_{CR} = \max\left\{0,\ \frac{\gamma - p}{1-p}\right\}$$
$$\overline{y}_{CR} = \min\left\{1,\ \frac{\gamma}{1-p}\right\}.^{4}$$

The lower and upper bounds of the average treatment effect on the comparable respondents are then given by:

$$\Delta_1^{lb} = E[Y_i|R_i = 1, T_i = 1] - \min\left\{1,\ \frac{\gamma}{1-p}\right\}$$
$$\Delta_1^{ub} = E[Y_i|R_i = 1, T_i = 1] - \max\left\{0,\ \frac{\gamma - p}{1-p}\right\}.$$

2.3.1.2 Estimation and inference

Our estimators replace population parameters with sample analogs. We obtain estimators for the bounds of ATE_{common} on the amount of dollars given via

$$\widehat{\Delta}_1^{lb} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \frac{\sum Y_i R_i (1-T_i) 1(Y_i \geq \widehat{y}_{\widehat{p}})}{\sum R_i (1-T_i) 1(Y_i \geq \widehat{y}_{\widehat{p}})}$$
$$\widehat{\Delta}_1^{ub} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \frac{\sum Y_i R_i (1-T_i) 1(Y_i \leq \widehat{y}_{1-\widehat{p}})}{\sum R_i (1-T_i) 1(Y_i \leq \widehat{y}_{1-\widehat{p}})},$$

where

$$\widehat{y}_q = \min\left\{y : \frac{\sum R_i (1-T_i) 1(Y_i \leq y)}{\sum R_i (1-T_i)} \geq q\right\}, \quad \text{for } q = \widehat{p} \text{ or } q = 1 - \widehat{p}$$

where

$$\widehat{p} = 1 - \frac{\sum R_i T_i / \sum T_i}{\sum R_i (1-T_i) / \sum (1-T_i)}.$$

[4]Note that, $\underline{y}_{CR} = 0$ if $\gamma < p$ and $\underline{y}_{CR} = \frac{\gamma - p}{1-p}$ otherwise; while $\overline{y}_{CR} = 1$ if $\gamma > 1 - p$ and $\overline{y}_{CR} = \frac{\gamma}{1-p}$ otherwise.

For the case where Y_i is continuous, Lee (2009) shows consistency and asymptotic normality of the estimator in Propositions 2 and 3 of his paper. As we extend this for the binary case below, we restate his two propositions. Suppose Y_i has bounded support, $\Pr(R_i = 1|T_i = 1) > 0$ and $p > 0$, then $\widehat{\Delta}_1^{lb}$ and $\widehat{\Delta}_1^{ub}$ converge in probability to Δ_1^{lb} and Δ_1^{ub}, respectively. Additionally, as $\Pr(R_i = 1|T_i = 1) < \Pr(R_i = 1|T_i = 0) < 1$, then $\sqrt{n}\Big(\widehat{\Delta}_1^{lb} - \Delta_1^{lb})\Big) \xrightarrow{d} N(0, V^{lb})$ and $\sqrt{n}\Big(\widehat{\Delta}_1^{ub} - \Delta_1^{ub})\Big) \xrightarrow{d} N(0, V^{ub})$.[5]

In the binary case, we can obtain the estimators for the bounds on ATE_{common} on the probability to give via

$$\widehat{\Delta}_1^{lb} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \min\Big\{1,\ \frac{\widehat{\gamma}}{1-\widehat{p}}\Big\},$$

and

$$\widehat{\Delta}_1^{ub} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \max\Big\{0,\ \frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}\Big\},$$

where

$$\widehat{\gamma} = \frac{\sum Y_i R_i (1-T_i)}{\sum R_i (1-T_i)}$$

and

$$\widehat{p} = 1 - \frac{\sum R_i T_i / \sum T_i}{\sum R_i (1-T_i) / \sum (1-T_i)}.$$

Note that Lee (2002) provides the procedure to construct the lower and upper bounds when the outcome variable is binary, but does not formally show consistency or the asymptotic distribution. For completeness, we formally provide these results. Analogous to Lee (2009), we apply Theorems 2.6 and 7.2 of Newey and McFadden (1994) and show

[5]The asymptotic variances are given by $V^{lb} = \frac{Var[Y_i|R_i=1,T_i=0,Y_i\geq y_p]}{E[R_i(1-T_i)](1-p)} + \frac{(y_p-\overline{y}_{CR})^2 p}{E[R_i(1-T_i)](1-p)} + (y_p - \overline{y}_{CR})^2\Big(\frac{1/\Pr(R_i=1|T_i=1)-1}{E[T_i]} + \frac{1/\Pr(R_i=1|T_i=0)-1}{1-E[T_i]}\Big) + \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_i T_i]}$, and $V^{ub} = \frac{Var[Y_i|R_i=1,T_i=0,Y_i\leq y_{1-p}]}{E[R_i(1-T_i)](1-p)} + \frac{(y_{1-p}-\underline{y}_{CR}])^2 p}{E[R_i(1-T_i)](1-p)} + (y_{1-p} - \underline{y}_{CR})^2\Big(\frac{1/\Pr(R_i=1|T_i=1)-1}{E[T_i]} + \frac{1/\Pr(R_i=1|T_i=0)-1}{1-E[T_i]}\Big) + \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_i T_i]}$, respectively.

consistency and the asymptotic normality of the estimators in the binary outcome case.

Proposition 1. Suppose $0 < \Pr(R_i = 1|T_i = 1) < \Pr(R_i = 1|T_i = 0) < 1$, then $\widehat{\Delta}_1^{lb}$ and $\widehat{\Delta}_1^{ub}$ converge in probability to Δ_1^{lb} and Δ_1^{ub}, respectively. Additionally, suppose we know which minimum of the estimator $\widehat{\overline{y}}_{CR} = \min\left\{1,\ \frac{\widehat{\gamma}}{1-\widehat{p}}\right\}$ and which maximum of the estimator $\widehat{\underline{y}}_{CR} = \max\left\{0,\ \frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}\right\}$ are attained, then we have $\sqrt{n}\left(\widehat{\Delta}_1^{lb} - \Delta_1^{lb})\right) \xrightarrow{d} N(0, V^{lb})$ and $\sqrt{n}\left(\widehat{\Delta}_1^{ub} - \Delta_1^{ub})\right) \xrightarrow{d} N(0, V^{ub})$, where

$$V^{lb} = \begin{cases} \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_iT_i]}, & \text{if } \widehat{\overline{y}}_{CR} = 1, \\ \frac{Var[Y_i|R_i=1,T_i=0]}{E[R_i(1-T_i)](1-p)^2} + \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_iT_i]} & \\ +(\frac{\gamma}{1-p})^2\left(\frac{1/\Pr(R_i=1|T_i=1)-1}{E[T_i]} + \frac{1/\Pr(R_i=1|T_i=0)-1}{1-E[T_i]}\right), & \text{if } \widehat{\overline{y}}_{CR} = \frac{\widehat{\gamma}}{1-\widehat{p}}, \end{cases}$$

and

$$V^{ub} = \begin{cases} \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_iT_i]}, & \text{if } \widehat{\underline{y}}_{CR} = 0, \\ \frac{Var[Y_i|R_i=1,T_i=0]}{E[R_i(1-T_i)](1-p)^2} + \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_iT_i]} & \\ +(\frac{1-\gamma}{1-p})^2\left(\frac{1/\Pr(R_i=1|T_i=1)-1}{E[T_i]} + \frac{1/\Pr(R_i=1|T_i=0)-1}{1-E[T_i]}\right), & \text{if } \widehat{\underline{y}}_{CR} = \frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}. \end{cases}$$

The Appendix A.1 formally provides the proof of Proposition 1.

2.3.2 Behaghel et al. (2015) bounds

If we have information on the number of attempts that are made to collect survey information, Behaghel et al. (2015) show that this information can be used to tighten bounds. In other words, their idea is to exploit additional information: the number of attempts made before the individual gives a response. When the gap between response rates is large enough, this additional information results in a narrower range of the outcome of the truncated marginal respondents and hence tighter bounds as compared to those obtained by Lee (2009).

To more formally describe their approach, we first introduce additional notation. Define

W_i as the total number of attempts made to reach individual i and N_i the number of attempts made before individual i responds with their final donation decision. We observe realized values of the following variables (Y_i, R_i, T_i, N_i) for each individual i.

Denote V_i (unobservable) as the individual's unwillingness to respond. Behaghel et al. (2015) impose the following assumptions (in addition to the independence and monotonicity assumption adopted by Lee (2009)):

Assumption 3: [Latent Variable Threshold-Crossing Response Model]

$$R_i = 1(V_i < p(W_i, T_i)), \forall i,$$

where $p(W_i, 0)$ and $p(W_i, 1)$ have nonempty common support as W_i varies.

Assumption 4: [Independence Between Treatment Status and Willingness to Response]

$$V_i \perp T_i, \forall i.$$

Assumption 5: [Same Maximum Potential Attempts to Obtain Individual's Response]

$$W_i = w_{\max}, \text{ where } w_{\max} \text{ is constant, } \forall i.$$

Assumption 6: [Non-decreasing in Response Rate]

$$p(W, Z) \text{ is non-decreasing in } W, \forall Z.$$

Assumption 3 indicates the individual i will only respond ($R_i = 1$) when the incentive to respond $p(W_i, T_i)$ exceeds the individual's unwillingness to respond V_i. Assumption 4 implies that the distribution of the individual's willingness to respond should be the same in the control and treatment groups under random assignment. Thus, an individual from the control group and one from the treatment group that have the same values of V_i represent comparable pairs. And the comparable subsamples in each group should represent the same population with the lowest $V_i's$ Assumption 5 implies that all

individuals in the data should potentially receive the same amount of attempts w_{max}.[6] Assumption 6 implies that more attempts to approach the individual will not decrease the response rate. This is reasonable in our case as none of the subjects were informed about the maximum number of e-mails sent in advance.

2.3.2.1 Parameter of interest

When our interest lies in the average treatment effect of an individual's donation amount, the outcome Y_i is continuous. Suppose the control group has a lower attrition rate, we can follow the Behaghel et al. (2015) bounding approach to estimate the bounds of the ATE_{common} on the donation amount, via the following four-step procedure:

1. Find the number of e-mails w_1 in the control group such that the response rate after the w_1^{th} e-mail in the control group is closest and higher than the overall response rate in the treatment group[7]:

$$w_1 = \min_{n \in \{1,\ldots,w_{max}\}} : \Pr(N_i \leq n | T_i = 0) \geq \Pr(R_i = 1 | T_i = 1).$$

2. Identify the share of marginal untreated respondents (those who will not respond if they were in the treatment group, but respond with their final donation decision after the w_1^{th} e-mail) among the untreated individuals that respond with their final donation decision after the w_1^{th} e-mail:

$$\alpha = \frac{\Pr(N_i \leq w_1 | T_i = 0) - \Pr(R_i = 1 | T_i = 1))}{\Pr(N_i \leq w_1 | T_i = 0) - \Pr(N_i \leq w_1 - 1 | T_i = 0)}.$$

   Then the share of the comparable untreated respondents among the untreated ones that respond to the w_1^{th} e-mail equals $1 - \alpha$.

[6] In our setting, $w_{max} = 5$ as 5 waves of e-mail solicitations are sent to raise funds for the EDF.

[7] In our field experiment, given the difference in response rates, w_1 is the same as the maximal number of e-mails sent w_{max}.

3. Find the α and $1-\alpha$ quantiles of the treated respondents' outcome distribution: $y_\alpha = G_Y^{-1}(\alpha)$ and $y_{1-\alpha} = G_Y^{-1}(1-\alpha)$, where G_Y is the CDF of Y conditional on $T_i = 0$ and $N_i = w_1$ (i.e., the CDF of the outcomes of untreated respondents who gave their final donation decision after the w_1^{th} e-mail).

4. The mean outcome of the comparable untreated respondents can be derived as the weighted sum of the mean outcome of the untreated individuals who gave their final donation decision by the first $w_1 - 1$ e-mails ($E[Y_i|N_i \leq w_1 - 1, T_i = 0]$) and the mean outcome of the comparable untreated respondents who gave their final donation decision after the w_1^{th} e-mail (y_{w_1}), weighted by their relative shares to the overall response rate of the treatment group.

   As illustrated in Figure 2, $\underline{y}$ and $\overline{y}$, the lower and upper bounds of y_{w_1}, are given as the mean of the lower and upper $1-\alpha$ portion of the outcomes of the untreated respondents who gave their final donation decision after the w_1^{th} e-mail, respectively:

$$\underline{y} = E\big[Y_i|N = w_1, T_i = 0, Y_i \leq y_{1-\alpha}\big] \tag{2.3}$$

$$\overline{y} = E\big[Y_i|N = w_1, T_i = 0, Y_i \geq y_\alpha\big]. \tag{2.4}$$

   If we take the difference between the average outcome of the treated respondents and the lower (upper) bound of the mean outcome of the comparable untreated respondents and the average outcome of the treated respondents, the lower bound on the average treatment effect (ATE_{common}) is then given by

$$\Delta_2^{lb} = E[Y_i|R_i = 1, T_i = 1] - \Bigg[\frac{\Pr(N_i \leq w_1 - 1|T_i = 0)}{\Pr(R_i = 1|T_i = 1)} E[Y_i|N_i \leq w_1 - 1, T_i = 0] + \Big(1 - \frac{\Pr(N_i \leq w_1 - 1|T_i = 0)}{\Pr(R_i = 1|T_i = 1)}\Big)\,\overline{y}\Bigg], \tag{2.5}$$

and the upper bound is given by

$$\Delta_2^{ub} = E[Y_i|R_i = 1, T_i = 1] - \Bigg[\frac{\Pr(N_i \leq w_1 - 1|T_i = 0)}{\Pr(R_i = 1|T_i = 1)}$$
$$E[Y_i|N_i \leq w_1 - 1, T_i = 0] + \Big(1 - \frac{\Pr(N_i \leq w_1 - 1|T_i = 0)}{\Pr(R_i = 1|T_i = 1)}\Big)\, \underline{y}\Bigg], \qquad (2.6)$$

where $\underline{y}$ and $\bar{y}$ are given in Equations 2.3 and 2.4, respectively.

When we are interested in the average treatment effect on an individual's probability to give, the outcome Y_i is binary. We again follow a four-step procedure to obtain the bounds of ATE_{common} on an individual's donation decision. The first two steps are the same, but differ in Steps 3 and 4:

3. Calculate the share of untreated respondents who donated among those who gave their final donation decision after the w_1^{th} e-mail (γ):

$$\gamma = \Pr(Y_i = 1|N_i = w_1, T_i = 0).$$

4. As illustrated in Figure 3, when y is binary, the bounds of mean outcome of comparable untreated respondents who respond after the w_1^{th} e-mail ($\underline{y}$ and $\overline{y}$) depend on the share of untreated respondents who donated among those who gave their final donation decision after the w_1^{th} e-mail (γ) and the share of marginal untreated respondents among the untreated respondents that respond after the w_1^{th} e-mail (α):

$$\underline{y} = \max\Big\{0,\ \frac{\gamma - \alpha}{1 - \alpha}\Big\} \qquad (2.7)$$
$$\overline{y} = \min\Big\{1,\ \frac{\gamma}{1 - \alpha}\Big\}.^{8} \qquad (2.8)$$

[8]Note that, $\underline{y} = 0$ if $\gamma < \alpha$, then $\underline{y} = 0$ and $\underline{y} = \frac{\gamma - \alpha}{1 - \alpha}$ otherwise; while $\overline{y} = 1$ if $\gamma > 1 - \alpha$ and $\overline{y} = \frac{\gamma}{1 - \alpha}$ otherwise.

The lower and upper bounds of the average treatment effect on the comparable respondents are then given by Equation 5 and 6 respectively, where $\underline{y}$ and $\overline{y}$ are given in Equation 7 and 8 instead for the binary case.

2.3.2.2 Estimation and inference

To estimate the bounds of the average treatment effect on the amount of dollars given, we replace the population parameters in Equations 2.5 and 2.6 with sample analogs. The response rates $\Pr(R_i = 1|T_i = 1)$ and $\Pr(N_i \leq w_1 - 1|T_i = 0)$ are respectively replaced by

$$\widehat{p_a} = \frac{\sum R_i T_i}{\sum T_i}$$

and

$$\widehat{p_c} = \frac{\sum 1(N_i \leq \widehat{w}_i - 1)(1 - T_i)}{\sum (1 - T_i)},$$

where

$$\widehat{w}_1 = \min_{n \in \{1,\ldots,w_{max}\}} : \frac{\sum 1(N_i \leq n)}{\sum T_i} \geq \widehat{p}_a.$$

Similarly, the mean outcomes $E[Y_i|N_i \leq w_1 - 1, T_i = 0]$ and $E[Y_i|R_i = 1, T_i = 1]$ are then replaced by

$$\widehat{\mu} = \frac{\sum Y_i 1(N_i \leq w_i - 1)(1 - T_i)}{\sum 1(N_i \leq w_i - 1)(1 - T_i)}$$

and

$$\widehat{v} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i},$$

respectively. $\underline{y}$ and $\overline{y}$ can be estimated via their sample analogs

$$\widehat{\underline{y}} = \frac{\sum Y_i 1(Y_i \leq \widehat{y}_{1-\widehat{\alpha}}) 1(N_i = \widehat{w}_1)(1 - T_i)}{\sum 1(Y_i \leq \widehat{y}_{1-\widehat{\alpha}}) 1(N_i = \widehat{w}_1)(1 - T_i)}$$

and

$$\widehat{\overline{y}} = \frac{\sum Y_i 1(Y_i \geq \widehat{y}_{\widehat{\alpha}}) 1(N_i = \widehat{w}_1)(1 - T_i)}{\sum 1(Y_i \geq \widehat{y}_{\widehat{\alpha}}) 1(N_i = \widehat{w}_1)(1 - T_i)},$$

respectively, where

$$\widehat{y}_q = \min\Big\{y : \frac{\sum 1(N_i = \widehat{w}_1)(1-T_i)1(Y_i \leq y)}{\sum 1(N_i = \widehat{w}_1)(1-T_i)} \geq q\Big\},$$

for $q = \widehat{\alpha}$ and $1 - \widehat{\alpha}$, respectively. Note that

$$\widehat{\alpha} = \frac{\widehat{p}_b - \widehat{p}_a}{\widehat{p}_b - \widehat{p}_c}$$

is the estimator of α, where $\widehat{p}_b$ is given by

$$\widehat{p}_b = \frac{\sum 1(N_i \leq \widehat{w}_i)(1-T_i)}{\sum(1-T_i)}.$$

The bounding estimators $\widehat{\Delta}_2^{lb}$ and $\widehat{\Delta}_2^{ub}$ are thus given by

$$\widehat{\Delta}_2^{lb} = \widehat{v} - \Big[\ \frac{\widehat{p}_c}{\widehat{p}_a}\ \widehat{\mu} + (1 - \frac{\widehat{p}_c}{\widehat{p}_a})\ \widehat{\overline{y}}\ \Big].$$
$$\widehat{\Delta}_2^{ub} = \widehat{v} - \Big[\ \frac{\widehat{p}_c}{\widehat{p}_a}\ \widehat{\mu} + (1 - \frac{\widehat{p}_c}{\widehat{p}_a})\ \widehat{\underline{y}}\ \Big]$$

Similarly, the estimator for the bounds of ATE_{common} on the probability to give are nearly identical to those in the continuous case. The difference being that the estimators of $\underline{y}$ and $\overline{y}$ are now given by

$$\widehat{\underline{y}} = \max\Big\{0,\ \frac{\widehat{\gamma} - \widehat{\alpha}}{1 - \widehat{\alpha}}\Big\}$$

and

$$\widehat{\overline{y}} = \max\Big\{1,\ \frac{\widehat{\gamma}}{1 - \widehat{\alpha}}\Big\},$$

where

$$\widehat{\gamma} = \frac{\sum Y_i 1(N_i = \widehat{w_1})(1-T_i)}{\sum 1(N_i = \widehat{w_1})(1-T_i)}$$

is the sample analog of γ. Appendix A.2 of Behaghel et al. (2015) shows consistency and asymptotic normality of the estimators mentioned above, the proofs of which are similar to

that in Lee (2009).

2.4 The field experiment

Our field experiment is an online fundraising campaign driven by five waves of e-mail solicitation for the Environmental Defense Fund (EDF). Based on the last two digits of their membership number, potential donors, those who have donated to the EDF before, are randomly assigned into one of the four assignment groups.

To examine whether a higher match rate would lead to changes in contributions, we focus on the two groups that have different match ratios. The control group, the 1:1 unlimited group, has a match rate that is 1:1. The 1:1 match rate implies that: for every dollar the individual donates, a leading donor donates a dollar (hence the fund receives two dollars). The treatment group, the 2:1 unlimited group, has a match rate that is 2:1. For every dollar the individual donates, a leading donor donates two dollars (hence the fund receives three dollars).

When asking for donations, e-mails were sent in five waves between the end of June and the first week of July in 2013. Once individuals gave or if they opted-out of future solicitations, they stopped receiving additional e-mails. By the end of the experiment, we collected data on the member number, the donation amount given (if any) and when they gave. We also recorded whether the e-mail was received and whether the e-mail was opened in each wave.

For our analysis, we focus on low donors, those who have lower ability to give (to rule out large values potentially skewing our results). Within the group of low donors, we separate them into two categories. The warm list donors are ones which contributed to the fund in the past 24 months, while the cold list donors are ones that did not make a contribution in the previous 24 months.

2.5 Empirical results

2.5.1 Summary statistics

Table 2.1 summarizes the donation rate, average donation amount and total amount of dollars raised for all groups. Note that for both groups, the wave 1 e-mail opening information is missing. For the 1:1 unlimited group, the wave 5 e-mail receiving and opening information are missing. As a result, the average number of e-mails received are calculated based on wave 1 to wave 4 for the 2:1 unlimited group and the 1:1 unlimited group. Also, the average number of e-mails opened are calculated based on wave 2 to wave 4 for the 2:1 unlimited group and the 1:1 unlimited group. We find that the 2:1 unlimited group has a lower average e-mail open rate. In Appendix A.2, we outline the methods used to account for the missing data. The main take from the appendix is that the results of the paper are robust to alternative approaches to accounting for the missing data.

Even though each group contains over 40,000 potential donors, the giving rate in each group is very low (below 1% on average). The 2:1 unlimited group has a higher donation rate (0.331%) compared to the 1:1 unlimited group (0.218%). For each group, the cold list has a much lower donation rate as compared to the warm list (1.02% versus 0.10% in the treatment group and 0.64% versus 0.09% in the control group). The 2:1 unlimited group raised more dollars (given) on average ($0.09 versus $0.06). This holds true when we restrict our analysis to the warm list donors ($0.2837 versus $0.1797). Among the cold list donors, there is almost no difference between the two groups in terms of the average gift.

As for the donation amount, the warm list donors have higher donation amounts on average as compared to the cold list ($0.2837 versus $0.0228 in the treatment group and $0.1797 versus $0.0228 in the control group). In terms of the total dollar amounts being raised, the 2:1 unlimited group raised more both out of pocket and with matched dollars compared to the 1:1 unlimited group ($3,956 versus $2,600 and $11,868 versus $5,200 respectively).

In this study, we are interested in estimating the average treatment effect of a higher

match rate on the donation amount and donation rate. The naive estimator can be obtained by simply calculating the difference of the average donation amount (or donation rate) between the two groups. However, due to sample selection, the naive estimate is biased as the treatment affects both the outcome (donation amount or donation rate) and whether the potential donor responded via opening an e-mail. In what follows, we present the estimated bounds of the average treatment effect on the donation amount and the donation rate after correcting for sample selection bias.

2.5.2 Bounds on the effect of a higher match rate

Tables 2.2 and 2.3 present the estimated treatment effect of a higher match rate on the donation decision and donation amount. The first two columns of each table give the results from the naive regressions (for all observations and e-mail respondents only, respectively). Column 1 shows the intent-to-treat effect. For the warm list donors, increasing the match rate significantly increases the likelihood to donate by 0.38% and the donation amount by around 10 cents on average. By contrast, there is almost no effect on the cold list donors. Column 2 gives the results for the naive estimator conditional on those who opened at least one e-mail, without correcting for selection bias. The number of observations who opened at least one e-mail are significantly smaller (down from 66,322 and 21,551 to 10,896 and 5,657 for the cold and warm list, respectively). This resulted in essentially no difference for cold list donors, but increases in the treatment effect for warm list donors. Increasing the match rate for this group increases the likelihood to donate by 1.11% and the donation amount by around 30 cents (the latter being insignificant).

The last two columns of Tables 2.2 and 2.3 present our bounding estimates. We begin with the Lee (2009) bounds (Column 3). The empirical bounds for the estimated average treatment effect on the likelihood to give are $[0.0008, 0.0018]$ for the cold list donors and $[0.0137^{***}, 0.0258^{***}]$ for the warm list donors (where three asterisks represent statistical significance at the one-percent level). The empirical bounds for the estimated average treatment effect on the donation amount are given by $[-0.0005, 0.0551]$ for the cold list

donors and $[0.3765^{*}, 0.8479^{***}]$ for the warm list donors. The estimates of the lower and upper bounds are both significantly different from zero for the warm list donors. There is essentially no effect on the cold list donors.

The Behaghel et al. (2015) bounds are given in the fourth column. The estimated bounds for the average treatment effect on the likelihood to give are $[0.0008, 0.0018]$ for the cold list donors and $[0.0137^{***}, 0.0232^{***}]$ for the warm list donors. The empirical bounds for the estimated average treatment effect on the donation amount are given by $[-0.0382, 0.0308]$ for the cold list donors and $[0.3840^{*}, 0.6650^{***}]$ for the warm list donors. Similar to the Lee (2009) bounds, the estimated lower and upper bounds of the Behaghel et al. (2015) bounds are both significantly different from zero for the warm list donors and there is essentially no effect on the cold list donors. Although the estimates of the lower bound are close, the Behaghel et al. (2015) bounds provide a more narrowly estimated interval compared to the Lee (2009) bounds.

In summary, the Lee (2009) and Behaghel et al. (2015) bounds imply that, for those warm list donors who would open at least one e-mail regardless of their treatment status, increasing the match rate (from 1:1 to 2:1) significantly increases the likelihood to donate by 1.37% to 2.32% (2.58% based on the Lee (2009) bounds) and the donation amount by around 38 cents to 67 cents (85 cents based on the Lee (2009) bounds).

2.6 Conclusion

In this paper, we explored the average treatment effect of a larger match ratio on an individual's donation behavior based on evidence from a field experiment using multiple waves of e-mail solicitation. Due to sample selection, estimation of this effect is not straightforward as donation decisions are observable only for those who opened at least one e-mail. In this case, naive estimation can result in biased estimates as our treatment impacts the response rate and the donation decision simultaneously. We apply nonparametric bounding estimation approaches (Lee (2009) and Behaghel et al. (2015)) to

correct for selection bias.

Based on the results of both approaches, we find that a higher match rate significantly increases an e-mail opener's likelihood to give and the donation amount for those who had contributed to the fund in the past 24 months. The Behaghel et al. (2015) bounds show that the estimated treatment effect on the likelihood to give and the donation amount are bounded between 1.37 and 2.58 percent, and 38 and 67 cents, respectively. Note that previous studies on matching gifts in charitable fund-raising based on evidence from direct mail solicitations (Karlan and List (2007), Eckel and Grossman (2008), Huck and Rasul (2011)) find that a higher match ratio or more matching grants are not more effective than smaller matches.

It is worth mentioning that, in our experiment, individuals are more likely to open e-mails when the match ratio is lower (1:1), especially warm list donors, based on the e-mail open information. As noted in Gee and Schreck (2018), an individual's belief of their own importance in providing funds and belief about other potential donors' donations matter in their donation decision. When individuals receive subsequent e-mails, they may consider themselves less important in providing funds when the match rate is high (say, 2:1) and thus less likely to open the e-mail to reconsider donation decisions, compared to the case with a lower match rate. Overall, we hope that these results shed light on theoretical and empirical studies of charitable giving and the mechanisms that drive people to give. We also hope that it helps charities develop optimal fund-raising strategies.

Figure 2.1: Identification under Behaghel et al. (2015) and Lee (2009) bounds

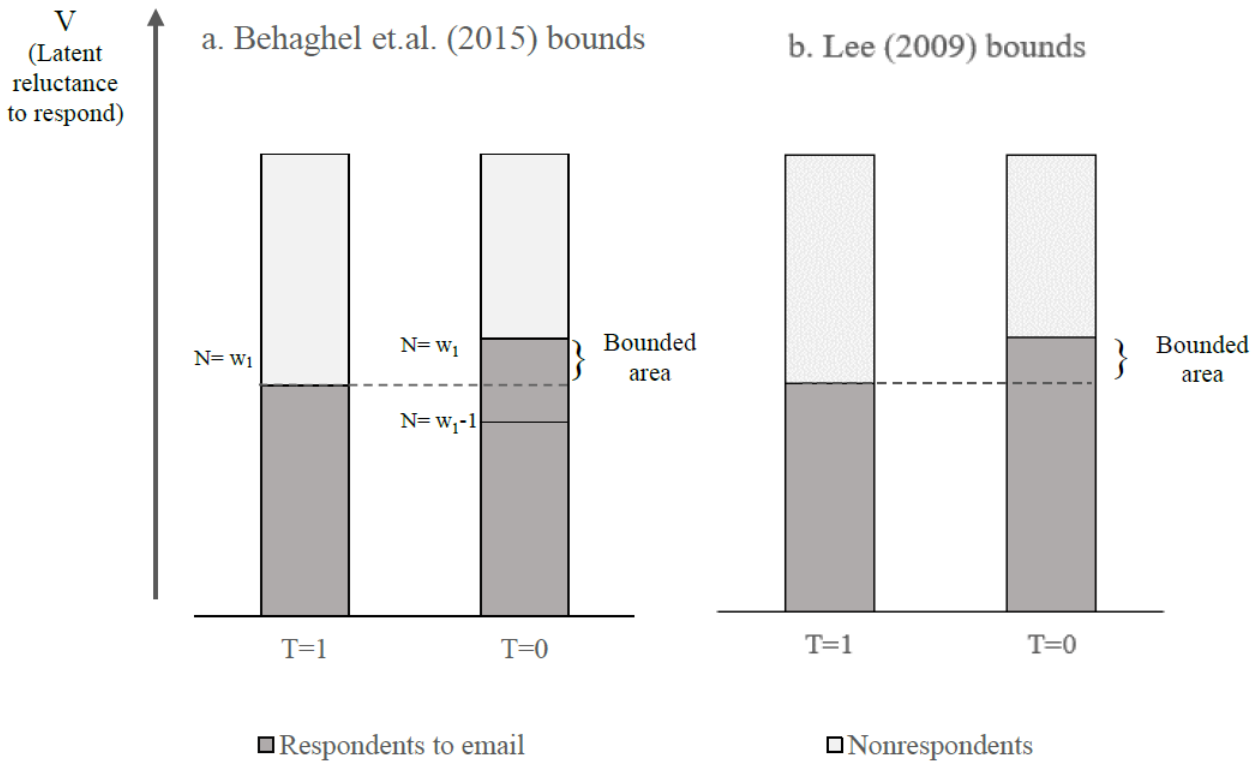


Figure 2.2: Bounding the mean outcome of comparable untreated respondents who gave their final donation decision to the w_1^{th} e-mail (Behaghel et al. (2015) bounds): continuous case

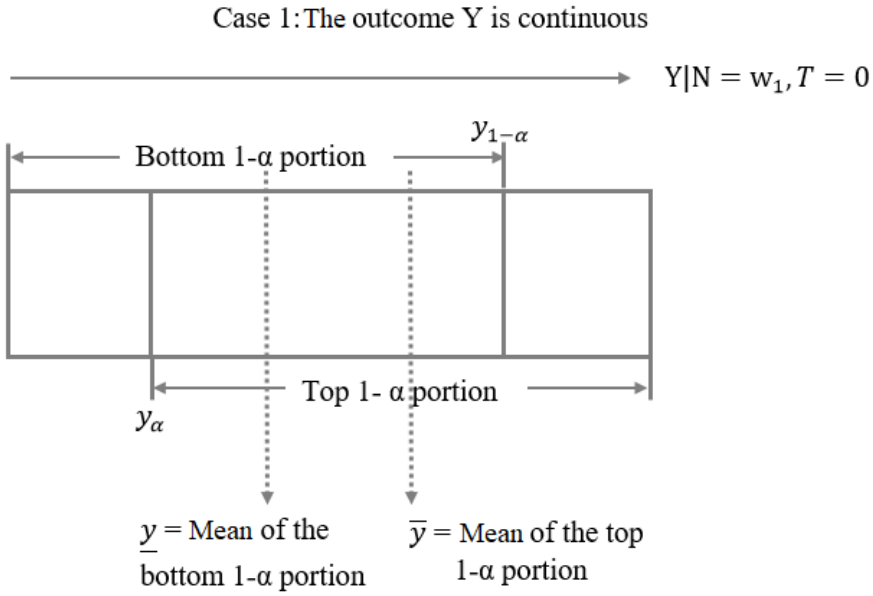


Figure 2.3: Bounding the mean outcome of comparable untreated respondents who gave their final donation decision to the w_1^{th} e-mail (Behaghel et al. (2015) bounds): discrete case

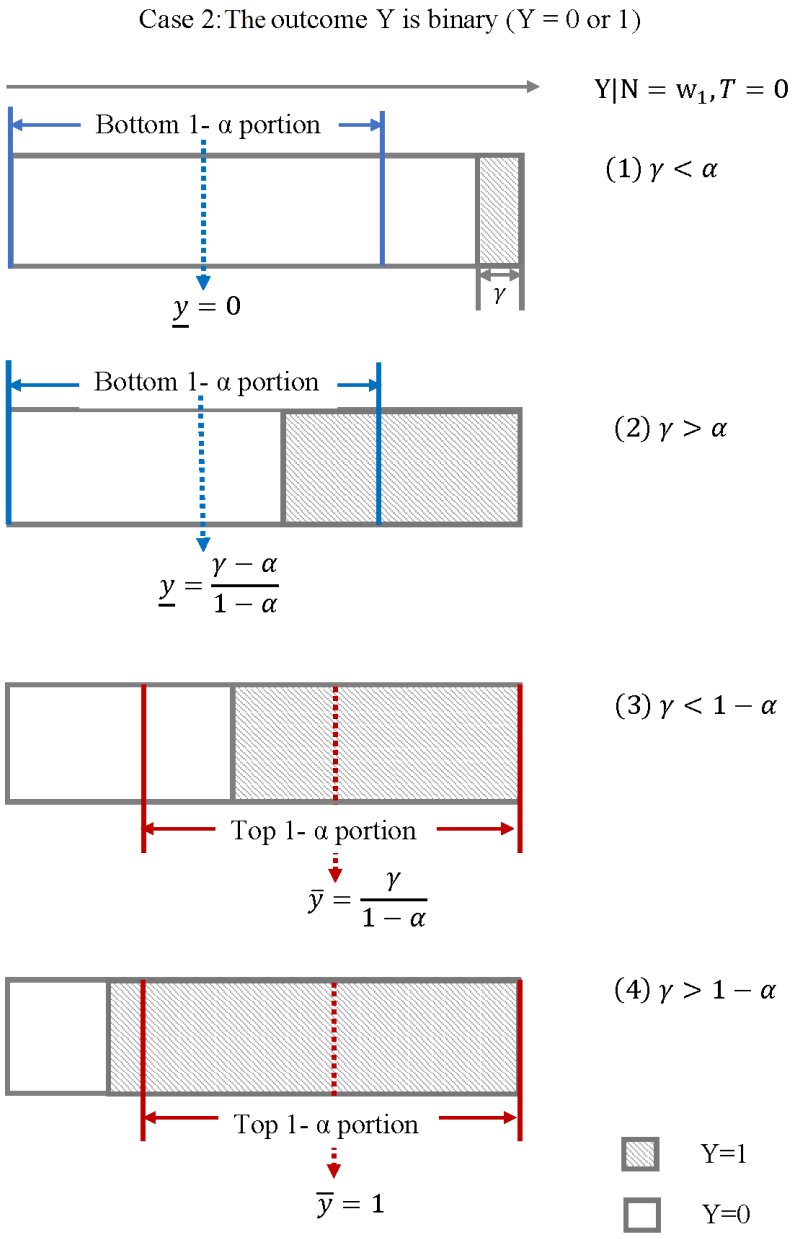


Table 2.1: Summary statistics (wave 1 to wave 5)

	2:1 unlimited (1)	1:1 unlimited (2)
Low donors - all		
Donation rate	0.3314%	0.2175%
Average gift	0.0886	0.0602
	(1.8599)	(1.5609)
Number of observations	44,664	43,209
Dollars raised - out of pocket	3,956	2,600
Dollars raised - with match	11,868	5,200
Dollars raised per match dollar	0.5000	1
Low donors - warm list		
Donation rate	1.0207%	0.6418%
Average gift	0.2837	0.1797
	(3.4229)	(2.7485)
Number of observations	11,267	10,284
Average number of e-mails received	3.2374	3.3116
Average number of e-mails opened	0.4231	0.4513
Low donors - cold list		
Donation rate	0.0988%	0.0850%
Average gift	0.0228	0.0228
	(0.8104)	(0.9122)
Number of observations	33,397	32,925
Average number of e-mails received	3.4789	3.4400
Average number of e-mails opened	0.2584	0.2706

Note: For both groups, the wave 1 e-mail opening information is missing. For the 1:1 unlimited group, the wave 5 e-mail receiving and opening information is missing. As result, (1) the average number of e-mails received are calculated based on wave 1 to wave 4 for both groups; (2) the average number of e-mails opened are calculated based on wave 2 to wave 4 for both groups.

Table 2.2: Estimated impact of increasing match rate on donation decision

	Naive regression	Naive regression (Respondents only)	Behaghel et al. (2015) bounds	Lee (2009) bounds
	(1)	(2)	(3)	(4)
Low donors - Cold list				
Treatment	0.0001	0.0001	[0.0008, 0.0018]	[0.0008, 0.0018]
	(0.0002)	(0.0014)	(0.0014, 0.0169)	(0.0013, 0.0013)
Number of observations	66,322	10,896	66,322	66,322
Number of observations based on comparable subsamples			11,601	11,601
The group with a higher e-mail open rate			1:1 unlimited	1:1 unlimited
Low donors - Warm list				
Treatment	0.0038**	0.0111*	[0.0137***, 0.0232***]	[0.0137*** , 0.0258***]
	(0.0012)	(0.0047)	(0.0044, 0.0040)	(0.0039, 0.0044)
Number of observations	21,551	5,657	21,551	21,551
Number of observations based on comparable subsamples			5,948	5,948
The group with a higher e-mail open rate			1:1 unlimited	1:1 unlimited

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2.3: Estimated impact of increasing match rate on donation amount

	Naive regression	Naive regression (Respondents only)	Behaghel et al. (2015) bounds	Lee (2009) bounds
	(1)	(2)	(3)	(4)
Low donors - Cold list				
Treatment	-0.0001	−0.0188	[−0.0382, 0.0308]	[−0.0005, 0.0551]
	(0.0067)	(0.0407)	(0.0383, 0.2890)	(0.0383, 0.0304)
Number of observations	66,322	10,896	66,322	66,322
Number of observations based on comparable subsamples			11,601	11,601
The group with a higher e-mail open rate			1:1 unlimited	1:1 unlimited
Low donors - Warm list				
Treatment	0.1040*	0.3021	[0.3840*, 0.6650***]	[0.3765*, 0.8479***]
	(0.0425)	(0.1614)	(0.1520, 0.1330)	(0.1515, 0.1211)
Number of observations	21,551	5,657	21,551	21,551
Number of observations based on comparable subsamples			5,948	5,948
The group with a higher e-mail open rate			1:1 unlimited	1:1 unlimited

Note: *p<0.1; **p<0.05; ***p<0.01

CHAPTER 3

ADVERTISED SHOW-UP FEES: SELF-SELECTION INTO EXPERIMENTS AND TREATMENT EFFECTS (WITH CARY A. DECK AND DANIEL J. HENDERSON)

3.1 Introduction

While random assignment of participants into treatments facilitates the internal validity of experiment evidence, the fact that lab experiments involve voluntary participation decisions has raised the concern about the external validity of the experiment results. Recent studies based on university student subjects have shown that those who have lower income, more leisure time, more interest in economics, more pro-social in terms of volunteering are more likely to participate into the lab experiment (Al-Ubaydli and List, 2013; Slonim et al., 2013). When subject's decision-making in the experiment and individual's participation decision are simultaneously affected by their characteristics (e.g., pro-social preference, individual's income level), participation bias arises - the estimated treatment effect based on the lab experiment evidence would be a biased estimate of the population treatment effect of interest and the inference from the lab experiment might not be able to be used on the broader population (Czibor et al., 2019).

As participation bias has become a rising concern in the lab and artefactual field experiment studies, one possible solution is to increase the advertised show-up fee (and thus the anticipated earning) when we recruit subjects. This solution helps reduce the participation bias that comes from the possibility that lower-income and more pro-social individuals more likely to attend lab experiments. Moreover, suppose the advertised participation fee only promotes participation and does not directly affect the participant's

decision in the experiment. In that case, we can utilize the randomly assigned advertised participation fees as the exclusion restriction and use the Heckman (1979) correction approach to account for participation bias that may arise in experiments related to individual preferences. However, the very few studies that have employed variation in the show-up fee in the estimation model to correct for the participation bias look at very high show-up fee levels (starting from around $19) and do not control for the actual show-up fee (Harrison et al., 2009, 2020)[1]. Failing to control for the actual show-up fee in the lab and artefactual field experiments will result in different endowments in the experiment and create another bias.

To fill this gap in the literature, in this paper, we study the impact of the advertised show-up fee on individual's participation decisions, their decision-making in three well-studied lab experiment tasks that respectively relate to an individual's risk preference, pro-social preference, and strategic sophistication. We also go a further step by investigating whether the advertised show-up fee has a direct impact on the treatment effect in these experiment tasks by holding the actual payment for participation to be the same for all participants. We design a lab experiment and incorporate these features in our recruitment procedure and experiment instruction, with the following three experiment tasks: (1) the risk bomb task where the treatment group faces higher stakes, (2) the donation task where the donation amount from subjects in the treatment group are matched one-for-one (versus no match in the control group), and (3) the p-beauty contest under cognitive load task where the treatment group is subject to a higher cognitive load. The impact of the advertised show-up fee was explored in a range from $5 that commonly used in lab experiment studies up to 3 times of this amount.

[1]Harrison et al. (2009) show that lab participants recruited with a higher show-up fee (250 kroner ≈ US$48 versus 100 kroner ≈ US$19) and the same price scale are observed to be more risk averse in the experiment. In another study based on artefactual field experiment evidence (Harrison et al., 2020), the participation rate increased from 18.1% to 24.1% when the individuals are recruited with a higher show-up fee (500 kroner ≈ US$96 versus 300 kroner ≈ US$58). However, both studies did not control for the actual show-up fee, it remains a question to see whether the difference in the measured risk attitude is coming from the difference in the characteristics of the participants or it's the result of the difference in the endowment (i.e., the actual show-up fee) that impacts participant's decision-making in the tasks.

In contrast to our hypothesis, we do not find any impact of the advertised show-up fee on an individual's decision to sign up or show up for the lab experiment study in a range of show-up fee that experimenters usually use to encourage participation. The results of the average treatment effect from the lab experiment tasks are mainly consistent with the findings from the previous literature, except that we do not find any significant treatment effect of a one-for-one match on subject's donation amount.

Consistent with the hypothesis that the advertised show-up fee should have no direct impact on the subject's decision-making and the treatment effect in the experiment tasks, we do not find any significant effect of the advertised show-up fee in the risk bomb task and the donation task. However, to our surprise, both nonparametric and parametric regression results suggest the treatment effect of a higher cognitive load on the guess in the p-beauty contest might fade away as the advertised show-up fee goes up. We also find a significant impact of the advertised show-up fee on the subject's guess for those under a high cognitive load but not for those under a low cognitive load based on the parametric regression results.

Our paper contributes to the literature on experiment design (e.g., Slonim et al. (2013), Harrison et al. (2009), and Schulz et al. (2019)) by providing evidence on how the advertised show-up fee as an important recruitment condition affect individual's participation decision and their performance in the experiment. The results of this paper also extend the literature on participation bias correction (e.g., Harrison et al. (2009, 2020)) by providing lab experiment evidence about whether the advertised show-up fee directly affects the treatment effect observed from various lab experiment tasks.

Our findings call for caution in the experiment design and econometric estimation of the treatment effect. Researchers should be careful when applying the methods that empiricists have used to correct for the participation bias by varying the show-up fee and using it as the exclusion restriction in the Heckman (1979) correction method (Harrison et al., 2009, 2020). Even experimenters can hold the actual show-up fee the same to control for the endowment effect and only vary the amount of advertised show-up fee when

recruiting subjects, the advertised show-up fee might still have an impact on subject's decision-making and even the treatment effect in the task (e.g., the p-beauty contest under cognitive load in our study). In that case, the use of randomized advertised show-up fees will not satisfy the exclusion restriction required in the Heckman (1979) correction approach.

Before going into the details of our study, we discuss why participation bias can be an issue in generalizing the inference from lab experiments to the broader population in interest. And we demonstrate how the randomly assigned advertised show-up fees can be used as an exclusion restriction in the Heckman (1979) two-step correction approach to correct for participation bias.

3.1.1 Why participation bias can be an issue in generalizing the inference from lab experiments

Consider the case where participation bias arises when both the impact of the treatment on the outcome of interest and the individual's participation decision in the lab experiment depend on individual characteristics. This can be illustrated in the Heckman (1979) selection framework. For simplicity, we use parametric function forms to illustrate why participation bias can be an issue in estimating the population average treatment effect. We consider data on m individuals, with index $i = 1, \ldots, m$.

Let P_i be individual i's participation decision in the lab experiment. Suppose the individual i's characteristic X_i also has an impact on their willingness to participate in the lab experiment and thus their participation decision P_i, then the sample selection equation is given by

$$P_i = 1(\pi + X_i\phi + v_i > 0), \tag{3.1}$$

where v_i is the error term that is independent of X_i. In addition, π is the intercept coefficient, and ϕ is the slope parameter for the variable X_i in the sample selection equation.

Let Y_i^* be individual i's outcome or economic behavior of interest. Suppose Y_i^* does not only depend on the randomly assigned treatment status T_i ($T_i = 1$ if treated and $T_i = 0$ if untreated) but also depends on the interaction term of the treatment status T_i and the individual's characteristic X_i (e.g., income level), then the population outcome equation is given by

$$Y_i^* = \beta_0 + T_i\beta_1 + T_iX_i\beta_2 + u_i, \tag{3.2}$$

where u_i is the error term that is independent of T_i and X_i. In addition, β_0 is the intercept coefficient, β_1 is the slope parameter for the variable T_i, and β_2 is the slope parameter for the interaction term T_iX_i in the population outcome equation.

From Equation 3.2, we can obtain the population average treatment effect (ATE), which is given by

$$ATE = E\left(Y_i^*|T_i = 1\right) - E\left(Y_i^*|T_i = 0\right) = \beta_1 + E(X_i)\beta_2. \tag{3.3}$$

If we would like to obtain inference of the treatment effect by conducting a lab experiment study, the outcome Y_i^* is only observed if the individual i attends the lab experiment. In other words, the observed outcome $Y_i = Y_i^*$ if $P_i = 1$.

For example, one maybe interested in estimating the impact of the provision of matching grant (the treatment, T_i) on an individual's donation behavior (Y_i^*) based on lab experiment evidence. It can be possible that the true treatment effect (i.e., how providing the matching grant affects the individual's donation) also depends on the individual's income level (X_i). The findings from Slonim et al. (2013) suggest lower-income individuals are more likely to attend a lab experiment (i.e., X_i is negatively correlated with P_i). In this case, when both the impact of the treatment on the outcome of interest and the individual's participation decision into the lab experiment depend on individual characteristics (as specified in Equations 3.1 and 3.2), simply comparing the average outcomes between the treatment group and the control group in the lab experiment give us

an estimate of $\beta_1 + E(X_i|P_i = 1)\beta_2$. This will result in a biased estimate of the population average treatment effect (ATE) since (1) ATE depends on the mean of the individual characteristic X_i and (2) the distributions of the individual characteristic X_i differ between the lab participants and the population they are recruited from[2].

In addition, simply regressing the observable outcome Y_i on T_i and T_iX_i will also result in a biased estimate of β_2 (and ATE) since

$$\begin{aligned} E\left(Y_i|T_i, X_i\right) &= E\left(Y_i^*|T_i, X_i, P_i = 1\right) \\ &= \beta_0 + T_i\beta_1 + T_iX_i\beta_2 + E(u_i\,|v_i > -\pi - X_i\delta) \\ &= \beta_0 + T_i\beta_1 + T_iX_i\beta_2 + m\left(X_i\right), \end{aligned}$$

where the last equality holds since the error terms u_i and v_i are correlated and both are independent with T_i and X_i.

3.1.2 Using Heckman (1979) model with randomly assigned advertised show-up fees as an exclusion restriction to correct for participation bias

While participation bias can limit the inference of the average treatment effect from lab experiment studies, one possible solution is to use the randomization and variation of the advertised show-up fees as an exclusion restriction and use the Heckman (1979) two-step correction approach to correct for this bias.

Let S_i be the randomly assigned advertised show-up fee for individual i. Suppose the randomly assigned advertised show-up fee only affects individual i's participation decision but has no direct impact on their decision-making in the lab experiment, then we have:

$$P_i = 1(\pi + X_i\delta + S_i\theta + v_i > 0) \tag{3.4}$$

$$Y_i^* = \beta_0 + T_i\beta_1 + T_iX_i\beta_2 + u_i, \tag{3.5}$$

[2]Note that $ATE = \beta_1 + E(X_i)\beta_2$ from Equation 3.3.

where Equation 3.4 is the sample selection equation modified from Equation 3.1 if we use the randomly assigned advertised show-up fee in the experiment recruitment and Equation 3.5 is the same population outcome equation as Equation 3.2 because we assume that the advertised show-up fee has no direct impact on the outcome of interest.

Once we randomly assigned the advertised show-up fee with variation, the estimation of the population average treatment effect is straight-forward: Following the estimation procedures in Heckman (1979), we can obtain a consistent estimate of the parameters β_1 and β_2. And we can also obtain the average mean of the covariate X_i which gives us an unbiased estimate of $E(X_i)$. In this way, we are able to correct for participation bias in estimating the population average treatment effect and obtain a consistent estimate of ATE, which equals $\beta_1 + E(X_i)\beta_2$.

The remainder of this paper is organized as follows: Section 3.2 presents the hypotheses we make on the impact of the advertised show-up fee. Section 3.3 introduces the experiment design including the recruitment procedure and the three lab experiment tasks. Sections 3.4 and 3.5 provide the experiment results and conclusion of our study, respectively.

3.2 Hypotheses

In this study, we are interested in whether a higher advertised show-up fee will attract additional participants with different characteristics and thus result in different performance and treatment effects in the three experiment tasks that relate to individual's risk preference, pro-social preference and strategic sophistication, respectively. Specifically, we try to answer the following two questions: (1) Does the advertised show-up fee have any impact on individual's decision to sign up for or participate in the lab experiment study? (2) Does it have any impact on participant's decision-making and the treatment effect in the lab experiment tasks when we hold the actual show-up fee constant for all participants?

Accordingly, we come up with the following three hypotheses:

Hypothesis 1: A higher advertised show-up fee will draw more individuals to sign up for and participate in the lab experiment study since it induces a higher expected payoff from the experiment.

Hypothesis 2: A higher advertised show-up fee should be more likely to attract those who are risk-averse and those who are less pro-social to participate the experiment but should have no impact on attracting those who are more sophisticated or less sophisticated as compared to the regular show-up fee indicated in the recruitment email.

Hypothesis 3: The advertised show-up fee should have no direct impact on the treatment effect in the experiment tasks, given the actual show-up fees are the same for all participants.

- **Hypothesis 3a:** If the participation rate does not increase with the advertised show-up fee at a specific range of the show-up fee levels, the treatment effect should remain the same at these show-up fee levels.

- **Hypothesis 3b:** If the participation rate increases with the advertised show-up fee at a specific range of the show-up fee levels, the treatment effect might vary across different show-up fee levels for the tasks related to individual's risk and pro-social preferences but not for one related to individual's strategic sophistication.

3.3 The experiment

The study was conducted in April 2021 and it contained two parts: (1) the recruitment procedure where potential participants were randomly noticed about advertised show-up fee and (2) the lab experiment where participants made their decisions in three individual decision-making tasks and answered survey questions about their demographic characteristics via the Qualtrics survey on the computers.

3.3.1 The recruitment procedure

The lab experiment was conducted in the TIDE Lab at the University of Alabama on April 15^{th} and 16^{th} in 2021. On the Sunday evening before the experiment, 6400 potential participants (those who have participated in the TIDE Lab experiment no more than once before the study) were invited through e-mails using the online recruitment software SONA[3]. In the recruitment email, they were told that: (1) they would receive a specific amount of money for showing up on time (i.e., the advertised show-up fee) and there would be the potential to earn more money based on their decisions in the study; (2) the study would last about 30 minutes; (3) they could choose one of the multiple time slots every half hour from 9:30-4:00 on Thursday, April 15^{th} and every 30 minutes from 9:30-2:30 on Friday, April 16^{th}[4]; (4) they must sign up for a session in SONA by midnight Wednesday to participate in the study; (5) there was a link to (sign up for) this study. All individuals received an identical recruitment email except the specific amount of the advertised show-up fee and the signup link assigned to them.

To investigate the impact of advertised show-up fee on individual's participation decisions, subject's decision-making and the treatment effects in the experiment tasks, we consider the range of show-up fee starting from \$5 (the amount that is commonly used in lab experiment studies in U.S.) up to three times of this amount. In order to randomly assign the advertised show-up fee among the 6400 potential participants, we randomly drew 100 levels of advertised show-up fee with two-digit decimals from the uniform distribution between 5 and 15. We then created 100 courses in SONA with 64 potential participants randomly assigned to each course. And each potential participant in the same course was noticed about the same amount of advertised show-up fee and the same signup link in the recruitment email.

[3]SONA database should have the contact information of all junior and senior students in the Culverhouse College of Business and students from other colleges and other class standings as those students have signed up to participate in lab experiment studies (either for course credits or monetary payments) through SONA before the experiment starts.

[4]To make it convenient and flexible to attend the study, we created multiple sessions every 30 minutes during the time periods on those two days when most of the college students would be available.

3.3.2 The lab experiment procedure

In each session of the lab experiment, each participant was given a slip of paper with a randomly assigned Lab ID on after showing up and signing the consent form. The first two digits of the Lab ID was the course number they were assigned. Since the course number is exclusively related to their advertised show-up fee, we can verify the advertised show-up fee assigned to the participant via the Lab ID they entered in the Qualtrics survey. The third digit is a letter to verify they correctly entered the Lab ID they assigned. And the last 4 digits are random number and letters to protect subject's anonymity within the same group. To protect subject's anonymity, the recruitment data in SONA was deleted after the experiment and there is no more linkage between subject's performance in the experiment and their personal information in the SONA database.

After they received the slip of paper with the Lab ID, each participant was seated to read the instruction on the computer screen and enter the Lab ID on the slip of paper to proceed to the experiment tasks. As shown in the experiment instruction, all participants were notified that: (1) the actual show-up fee was $15, which included the amount that they were told that they would receive when they registered for this study, (2) they might also receive an additional payment based upon the decisions that they make in this study, and (3) after they completed all three tasks, one of the tasks would be randomly selected to determine their final payment with equal chance to be chosen. The order of the three tasks was presented randomly to each participant. After finishing the three tasks, the subjects were asked to finish a survey about their demographic characteristics, including gender, race, class standing/university affiliation, household income of their family and their work status. When they finished the survey, they were notified about the randomly selected task that determined their final payment. At the end of the experiment, they were notified about their final payment on the screen. They were all paid in cash and left the study.

3.3.3 The three lab experiment tasks

For the lab experiment, we consider three individual decision-making tasks that have been well-studied in the experimental economics literature: the risk bomb task, the donation task and the p-beauty contest under cognitive load task. Each of these three tasks has a specific treatment. For each task, the subjects were randomly assigned into the treatment group and the control group. We are interested in whether the advertised show-up fee will have a direct impact on the experiment outcomes and the treatment effect in any of these three tasks.

3.3.3.1 The risk bomb task

The risk bomb task is a modified version of the Bomb Risk Elicitation Task (BRET) initially used in the lab by Crosetto and Filippin (2013). In this task, subjects were asked about how many boxes (b) they would like to collect out of 100 boxes. They were also notified that: (1) a time bomb was randomly hidden inside one of these boxes, (2) the boxes would be collected in a numerical order that starts with box number 1, number $2, \ldots,$ and ends with box number 100, (3) they would gain a specific amount of money (m dollars) for each box they collected if the bomb was not inside one of the boxes they collected and gain nothing if the opposite case happened. With all else the same, $m = 0.5$ in the treatment group (with higher stakes) and $m = 0.1$ in the control group (with lower stakes).

We use this risk bomb task to elicit participants' risk preference and to test if the treatment effect, the impact of a higher stake on participant's decision in the task, will vary with the advertised show-up fee. The payoff scale effect, i.e., higher stakes induce higher average measured degree of risk aversion, was first found by Holt and Laury (2002) using the paired lottery-choice tasks (henceforth, HL task) and was also found by Crosetto and Filippin (2013) using BRET. As compared to the HL task that were commonly used in the literature, the (modified) BRET makes it easier for subject to understand and make decisions. Given the outcome (b) is an integer between 0 and 100, it also allows more

continuous and more precise measurement of individual's risk preference. This measurement is also excepted from the degree of loss aversion or violating the Reduction Axiom (Crosetto and Filippin, 2013).

3.3.3.2 The donation task

The donation task is adapted version of the dictator game initially used by Eckel and Grossman (1996) where subjects were asked to split a certain amount of money between himself/herself and an established charity. In this task, subjects were asked to decide how much of the $15 participation payment they would like to keep and how much they would like to donate to The United Way of West Alabama. In addition to this information, the participants in the treatment group were notified that any amount they donate would be matched one-for-one so that the charity would receive two dollars for every dollar they donate.

We use this donation task to elicit participants' pro-social preference and to test if the treatment effect, the impact of a one-for-one donation match on participant's binary decision to give and donation amount in the task, will vary with the advertised show-up fee. The match effect, i.e., matching participant's donation (with a 100% match rate) lead to higher donation amount on average, was found in the lab experiment studies by Eckel and Grossman (2003, 2006) and the lab experiment study by Blumenthal et al. (2012) using different endowment amounts and with-in subject design[5]. As the former one adopted much lower endowments and the latter one used a much higher endowment, it would be interesting to see if the match effect also holds using different endowment. It would be also interesting to see if the match effect holds under between-subject design.

[5]The experiment in Eckel and Grossman (2006) used $4, $6 and $7.50 as endowment while the experiment in Blumenthal et al. (2012) used $100 as endowment. In both previous studies as well as the Eckel and Grossman (2003) study, researchers provided a list of ten charities for experiment participants to choose among while our study only provides one charity for experiment participants to consider for donation.

3.3.3.3 The p-beauty contest under cognitive load task

The p-beauty contest under cognitive load task is a modified version of the one used by Carpenter et al. (2013). In this task, subjects were told that the task would proceed in three stages. In stage 1, they would memorize a number that would be shown on their screen for 6 seconds after they read instruction without any assistance to record the number. In stage 2, they would play the guessing game where they would pick a number between 0 and 100 (inclusive). In stage 3, they would recall the number that were shown to them before the guessing game. With all else the same, the number to memorize is a randomly picked 7-digit number without repetition in the treatment group and a randomly picked 1-digit number in the control group.

At the beginning of the task, they were also notified that (1) they would be paid by either correctly recalling the number or winning the guessing game with equal chance, (2) they would earn $20 for successfully recalling the number or $10 by winning the guessing game, (3) a group of people in a previous were asked to play this same guessing game and 15 guesses from those people would be used with the subject's guess to see if the subject win the guessing game, and (4) the winning rule (i.e., that the winning guess is the guess closest to one half of the average guess) would be shown when they play the guessing game in stage 2. To adapt the p-beauty contest under an individual task setting, we randomly picked 15 guesses from the lab experiment study by Nagel (1995) for the guesses used to pin down the winning guess in this experiment task.

We use this p-beauty contest under cognitive load task to test if the treatment effect, the impact of a higher cognitive load on participant's guess in the p-beauty contest game, will vary with the advertised show-up fee. The cognitive load effect, i.e., a higher cognitive load leads to lower strategic sophistication (with higher expected guesses from other and higher guess away from their best response to their expectation), was first found in the lab experiment study by Carpenter et al. (2013).

3.4 Results

Though 6400 recruitment emails were sent to potential participants in the SONA database, we ended up with 1734 individuals that actually received the emails[6]. In total, 125 individuals signed up for the study and 105 of them showed up and participated in different sessions of the experiment. Note that both the registration rate and participation rate are low (7.21 and 6.06%, respectively) but still consistent with the normal sign-up and show-up rates in the Tide lab. And the participation rate conditional on registration is 84%. On average, participants received $18.99 for attending the experiment (standard deviation = 8.94), including the $15 actual show-up fee. Notice that median payment is $15, which equals the actual show-up fee.

3.4.1 The impact of the advertised show-up fee on individual's registration and participation decisions

We begin by examining the effect of the advertised show-up fee on individual's registration decision, participation decision as well their participation decision conditional on registration based on the observed decisions from the 1734 individuals who received the recruitment email.

For each of these outcomes in interest, we estimate the outcome on the advertised show-up fee using the local-constant-least-squares (LCLS) regressions[7]. This should give us a hint about the relationship between the outcome in interest and the advertised show-up fee. For these LCLS regressions, we specify the kernel function to be the Gaussian kernel and use the least-square-cross-validation (LSCV) method to determine the bandwidth

[6]There is a slight and statistically significant upward trend that more people are being asked for higher dollar amounts. If individuals with higher advertised show-up fee are more likely to sign up for or participate in the study, we should observe more individuals with higher advertised show-up fee to sign up for or participate due to this upward trend. However, this upward trend should have very little impact on our result since we do not observe any impact of the advertised show-up fee in individual's registration and participation decisions based on our experiment data.

[7]The general idea of the LCLS is to find the function value at each data point by using the weighted average of the nearby data points, with the further data point giving the lower weight. The weight is determined by the kernel function while the amount of nearby data points used is determined by the smoothing parameter – the bandwidth.

within the range from 0.1 to 100[8]. Notice that the upper bound of the search range for the bandwidth is larger than three times of the standard deviation of the regressor (i.e., the advertised show-up fee). Even in the case where we have a small sample size, the LSCV bandwidth selection method works well in removing irrelevant variables in the regression model: if the selected bandwidth is very large (larger than three times the standard deviation of the regressor), the estimated function will be approximately equal to the average outcome[9] and the plot of the fitted values will be almost a straight line. This suggests that the regressor has no impact on smoothing the function and implies that the regressor has essentially no impact on the dependent variable Henderson and Parmeter (2015). In addition, we use the wild bootstrap with 10,000 replications to obtain the 95% confidence intervals of the estimated conditional mean and estimated gradient from the LCLS regression. This also helps us to pin down the correct function forms for the parametric regressions of the outcome in interest on the advertised show-up fee.

We first present the results of the three LCLS regressions using the advertised participation fee as the regressor and using individual's registration decision, participation decision and participation decision conditional on registration as the dependent variable, respectively. As shown in Figures 3.1, 3.4 and 3.7, for all the three LCLS regressions, the bandwidth selected via the LSCV method is 100[10], which implies that the advertised show-up fee has almost no impact on smoothing the LCLS function. As for the impact of the show-up fee on registration decision, the plot of the fitted values from the LCLS regression in Figure 3.2 is almost a straight line and the gradient plot and Figure 3.3 suggests that the marginal effect of the advertised show-up fee is not statistically different from zero. Similar patterns and conclusions can be obtained from results of the LCLS

[8]We also consider using another popular cross-validation method, the AICc criterion, to determine the bandwidth for the LCLS regressions. All results from the LCLS regressions using the AICc bandwidths are similar to the results from the LCLS regressions using the LSCV bandwidths. For simplicity, we only report the results using the LSCV bandwidths.

[9]As the bandwidth approaches infinity, the estimated function in the LCLS regression is just the average outcome.

[10]This selected bandwidth exceeds three times the standard error of the regressor, i.e., advertised show-up fee.

regression of participation decision shown in Figures 3.5 and 3.6 as well as from results of the LCLS regression of participation decision conditional on registration shown in Figures 3.8 and 3.9.

We then consider using inference from the parametric regression models to verify our conjecture. As in Table 3.1, we show the regression results of registration decision on advertised show-up fee based on the Probit model and the linear probability model. For each model, the estimated coefficient for the advertised show-up fee is not statistically different from zero when the model includes the show-up fee as explanatory variable. Moreover, both AIC and BIC suggest selecting the model without the show-up fee as explanatory variable. For the other two outcomes in interest, we present the parametric regression results in Tables 3.2 and 3.3 and we find similar results as in Table 3.1: the estimated coefficient for the advertised show-up fee is not statistically different from zero and both AIC and BIC favor the model without the show-up fee as explanatory variable.

Result 1. In contrast to our hypothesis (**Hypothesis 1**), both the nonparametric and parametric regression results suggest that the advertised participation fee (in the range between \$5 and \$15) has essentially no impact on individual's decision to sign up for or participate in the lab study. They also suggest that the advertised show-up fee has no impact on individual's decision to show up in the lab experiment once they sign up for the study. And given the show-up fee in the recruitment email has no impact on the participation rate or the registration rate, we are unable to verify **Hypothesis 2** within the range of the advertised participation fee considered in this study.

3.4.2 The impact of the advertised show-up fee on participant's decision and treatment effect in the tasks

We then followed by examining the impact of the advertised show-up fee on participant's decision as well as the treatment effect in each of the three lab experiment tasks: the risk bomb task, the donation task and the p-beauty contest under cognitive load task. For each task, we also estimate participant's decision made in the task on the

advertised show-up fee using the LCLS regressions with Gaussian kernel and LSCV bandwidth using the same search range from 0.1 to 100, separately for the treatment group and the control group. This should help us to find out whether the advertised show-up fee has an impact on participant's decision in the task and give us a hint about whether the treatment effect vary across different show-up fee levels.

3.4.2.1 Results from the risk bomb task

In the risk bomb task, 53 participants are randomly assigned to the treatment group with higher stakes and 52 participants are randomly assigned to the control group with lower stakes. Figure 3.10 shows the distribution of outcome (the boxes collected in the risk bomb task) in the treatment group with respect to the control group. We can see that the distribution shapes from the two groups are largely the same, but the distribution of the control group seems to shift to the right of that of the treatment group. As shown in Table 3.4, the average number of boxes collected by the subjects is lower in the treatment group with higher stakes (mean = 34.792 vs. 41.288). Though results from the Kolmogorov-Smirnov test fail to reject the null that the distributions of the task outcome in both groups are the same ($p = 0.524$ for two-sided test and $p = 0.267$ for the one-sided test with H_1: the distribution of the outcome in the control group stochastically dominates that in the treatment group), suggesting that the distribution shapes from the two groups are largely the same. Both results from the t-test of equal means ($p = 0.027$ with H_1: control group has a higher mean) and result from the Mann-Whitney test of equality in distribution ($p = 0.060$ with H_1: true location shift is less than 0) suggest that higher stakes significantly increase the average degree of risk aversion elicited from the lab experiment task, but not by a large value. Our finding about the average treatment effect of the higher stakes on the measured risk attitude is consistent with the findings in risk elicitation literature (Holt and Laury, 2002; Crosetto and Filippin, 2013).

We then present the results of the LCLS regressions of the risk bomb task outcome on

the advertised participation fee for the treatment group and the control group, respectively. As shown in Figure 3.11, the LSCV bandwidth selected for the treatment group is 1.77, while the LSCV bandwidth selected for the control group is 100. For the control group, the bandwidth selected via the LSCV approach (i.e., the LSCV bandwidth) suggests that the advertised show-up fee has almost no impact on smoothing the LCLS function. As result, Figure 3.12a shows the plot of the fitted values from the LCLS regressions using LSCV bandwidth separately for the treatment and the control group: the plot for the treatment group has a slight dip while the plot for the control group is almost flat across all show-up fee levels. And there is almost no overlap in the 95% confidence intervals when the show-up fee is between 8 and 12. This suggests there might be significant treatment effect on this task outcome. As in the gradient plots in Figure 3.13a, we find no significant marginal effect of the show-up fee on the risk bomb task outcome either for the treatment group or for the control group. There seems to be no significant difference in the marginal effect of the advertised show-up fee as well.

To make the treatment group and control group more comparable, we also show the plot of the fitted values and the plot of the gradients from the LCLS regressions using the same bandwidth (the LSCV bandwidth for the treatment group = 1.77)[11] . As result, Figure 3.12b show that the treatment effect seems to vanish as the advertised show-up fee exceeds $13 since this smaller bandwidth gives less smoothing to the LCLS regression function for the control group. However, the gradient plot in Figure 3.13b still suggests insignificant marginal effect of the show-up fee on the risk bomb task outcome for both groups as well as no evidence on significant impact on the treatment effect.

We then consider using inference from the parametric regression models to verify our conjecture based on the nonparametric regression results. As in Table 3.4, we show the results from the linear regression model, separately for the treatment group and the control

[11]We can also plot the fitted values from the LCLS regressions using the same bandwidth (the LSCV bandwidth for the control group = 100). In that case, both the plots of the fitted values should be almost straight lines.

group. For each group, the coefficient of the advertised participation fee is not statistically different from zero and both AIC and BIC favor the model without the explanatory variable (i.e., the advertised participation fee). This suggests that the advertised participation fee has essentially no impact on lab participant's decision-making in the risk bomb task. Using the data of both groups in Table 3.6, we also present the results from the linear regression models with the interaction term to test for the effect of the advertised show-up fee on the treatment effect. The results suggest a marginally statistically significant treatment effect of the higher stakes on the risk bomb task outcome ($p = 0.053$) and no statistically significant impact of the advertised show-up fee on the treatment effect. Note that both AIC and BIC favor the models without the interaction term of the treatment and the advertised show-up fee.

Notice that the outcome in the risk bomb task, the number of boxes collected, is a count variable, we also consider inference from the negative binomial regression model. Based the results from the negative binomial regressions in Tables 3.7 and 3.8, we can obtain almost the same conclusion since those results are very similar to the results based on the linear regression model.

Result 2. Consistent to our hypothesis (**Hypothesis 3a**), we found no evidence on the impact of the advertised show-up fee on participant's decision or impact on the treatment effect in the risk bomb task. Our findings about the average treatment effect (higher stakes) are consistent with previous lab experiment evidence.

3.4.2.2 Results from the donation task

In the donation task, 55 participants are randomly assigned to the treatment group with a one-for-one donation match and 50 participants are randomly assigned to the control group with no match. As compared to the control group, the treatment group has a slightly lower donation amount on average (mean = 3.066 vs. 3.273) and a lower donation rate (63.64% vs. 72.00%). But the donation amount conditional on giving is slightly higher

in the treatment group (mean = 4.818 vs. 4.545).

In contrast to the previous lab experiments (Eckel and Grossman (2006); Blumenthal et al. (2012)) that found a positive impact of a 100% match on donation amount, we do not find any evidence on the average treatment effect of the 1:1 match on donation amount[12]. As shown in Figures 3.14 and 3.15, the distributions of donation amount (unconditional and conditional on giving, respectively) from the treatment group and the control group suggest that the distribution of the donation amount from the two groups are largely the same. This conjecture is also supported by the results from the tests of equal means and equal distributions in Tables 3.9 and 3.10: all the one-sided and two-sided tests fail to reject the null of equal means or equal distributions of donation amount (unconditional and conditional on giving) from the two groups ($p > 0.100$). Although our result seems to be different from the previous lab experiment studies, the evidence of no significant difference in the average dollars given conditional on giving with or without match is consistent with some of the field experiment evidence (Eckel and Grossman (2006); Karlan and List (2007)).

(a) The impact of the advertised show-up fee on donation amount and the related treatment effect

To find out whether the advertised participation fee has an impact on the donation amount and the treatment effect on donation amount, we first present the results of the LCLS regressions for the treatment group and the control group, respectively. As shown in Figure 3.16, the LSCV bandwidth for the treatment group is 4.27 and the LSCV bandwidth for the control group is 100. As a result, the plot of the fitted values from the LCLS regressions in Figure 3.17a is a slightly upward curve for the treatment group while

[12]Such a difference may come from the facts that: (1) we are using college students from different universities and Eckel and Grossman (2006) showed different results from college students from different universities based on a lower match rate and a lower endowment level, (2) we have a between-subject design in our experiment while Eckel and Grossman (2006) and Blumenthal et al. (2012) both use within-subject design to compare match (with different match rates) and no match, and (3) both studies provided a list of six or ten charities for experiment participants to choose among while our study only provides one charity for experiment participants to consider for donation. We might have a lower donation rate as compared to the previous lab experiment studies.

the plot for the control group is almost flat across all show-up fee levels. The plot shows very little difference in the fitted values of donation amount from the LCLS regressions for these two groups and the differences vanish as the show-up fee increases. The almost overlapping 95% confidence intervals of the fitted values for the two groups suggest no evidence on a significant treatment effect[13]. And the gradient plots in Figure 3.18 also suggest relatively small and insignificant marginal effect of the advertised show-up fee on the donation amount for both groups. And the overlap in the confidence intervals of the marginal effect from the two groups suggests no evidence on significant difference in the marginal effect of the advertised show-up fee as well.

We then consider using inference from the parametric regression models (the linear regression model, the linear regression model with a log-level function) to verify our conjecture based on the nonparametric regression results. We perform parametric regressions separately for the treatment group and the control group and also jointly using the whole sample with the interaction term to test for existence of the effect of the advertised show-up fee on the treatment effect. We present the results from the linear regression model and the linear regression model using the log of donation amount as dependent variable[14] in Tables 3.11, 3.12, 3.13 and 3.15: there is relatively small and insignificant impact of the advertised participation fee on the donation amount or on the related treatment effect on donation amount. Both AIC and BIC suggest the model with only a constant term more likely to be the correct model.

As the donation amount is a nonnegative number and around 30% of the donation amounts are zeros, we also consider the two-limit Tobit model to estimate the effect on donation amount without the lab environment constraint (i.e., without censoring from below at 0 and censoring from above at the \$15 participation fee). We present the results in Appendix B.1. The results in Table B.1 from the two-limit Tobit regressions suggests

[13]Similar patterns and conclusions can be reached if we use the same bandwidth (the LSCV bandwidth selected for the treatment group) for the two groups as in Figure 3.17b.

[14]Notice that around 30% of the donation amounts are zeros, we add \$0.1 to the zero donation amount for the linear regression model with the log of donation amount as the dependent variable.

that the advertised show-up fee does not have any significant marginal effect on donation amount for either group. Based on the whole sample regression in Table B.2, the two-limit Tobit regressions implies a significant negative treatment effect if we include the interaction term of the advertised participation fee and treatment status. However, both AIC and BIC favor the model without the interaction term and with a constant term only in the joint regression.

Since there are a few subjects who decided to donate nothing in the donation task, the relationship between the advertised show-up fee and the donation amount might be better analyzed in a two-part framework. We then look at the impact of the advertised show-up fee on subject's binary decision to give and the donation amount conditional on giving respectively.

(b) The impact of the advertised show-up fee on the binary decision to give and the related treatment effect

To find out whether the advertised participation fee has an impact on subject's decision to give or not and the related treatment effect, we first present the results of the LCLS regressions for the treatment group and the control group, respectively. As shown in Figure 3.19, the LSCV bandwidth for the treatment group is 1.40 and the LSCV bandwidth for the control group is 100. As a result, the plot of the fitted values from the LCLS regressions in Figure 3.20a is a curve that bends upward in the middle for the treatment group while the plot for the control group is almost flat across all show-up fee levels. Although the plot shows large difference in the fitted values of the likelihood to give when the show-up fee level is relatively small or relatively large, the 95% confidence intervals of the fitted values for the two groups overlap along all show-up fee levels. Note that this can result from the fact that the nonparametric regressions perform relatively poor around the lower bound or upper bound of the regressors when there are not enough data points for smoothing in the neighborhood. And the gradient plots in Figure 3.21a also suggest relatively small and insignificant marginal effect of the advertised show-up fee on the

decision to give for both groups. These plots suggest no evidence on statistically significant treatment effect on subject's decision to give or not[15].

We also present the results from the parametric regressions (the linear regression model and the Probit regression model) separately for the treatment group and the control group and also jointly using the whole sample with the interaction term to test for existence of the effect of the advertised show-up fee on the treatment effect. As in Tables 3.15, 3.16, 3.17 and 3.18, there is relatively small and insignificant impact of the advertised participation fee on the binary decision to give or on the related treatment effect. Both AIC and BIC suggest the model with only a constant term more likely to be the correct model.

(c) The impact of the advertised show-up fee on the donation amount conditional on giving and the related treatment effect

We then conclude the analysis on the donation task by examining whether the advertised participation fee has an impact on subject's donation amount conditional on giving and the related treatment effect. We first present the results of the LCLS regressions for the treatment group and the control group, respectively. As shown in Figure 3.22, the LSCV bandwidth for the treatment group is 3.19 and the LSCV bandwidth for the control group is 9.61 (which is larger than three times the standard deviation of the advertised show-up fee). As a result, the plot of the fitted values from the LCLS regressions in Figure 3.23a is a relatively flat curve with a slightly upward slope for the treatment group while the plot for the control group is almost flat across all show-up fee levels. The almost overlapping 95% confidence intervals of the fitted values for the two groups suggest no evidence on a significant treatment effect. And the gradient plots in Figure 3.24a also suggest relatively small and insignificant marginal effect of the advertised show-up fee on the donation amount conditional on giving for both groups. These plots suggest no evidence on statistically significant impact of the advertised show-up fee on the treatment

[15]Similar patterns and conclusions can be reached if we use the same bandwidth (the LSCV bandwidth selected for the treatment group) for the two groups as in Figures 3.20b and 3.21b.

effect[16].

We also present the results from the parametric regressions (the linear regression model and the linear regression model with a log-level function) separately for the treatment group and the control group and also jointly using the whole sample with the interaction term to test for existence of the effect of the advertised show-up fee on the treatment effect. As in Tables 3.19, 3.20, 3.21 and 3.22, there is a relatively small and insignificant impact of the advertised participation fee on the donation amount conditional on giving or on the related treatment effect. Both AIC and BIC suggest the model with only a constant term more likely to be the correct model. To estimate the effect of the advertised show-up fee on the donation amount conditional on giving without the lab environment constraint (i.e., without censoring from above at the $15 participation fee), we also consider the Type I Tobit model with right-censoring and using a log-level function form. We present the results in Appendix B.1. Similar conclusions can be reached from the results in Tables B.3 and B.4 from the Type I Tobit model regressions.

Result 3. Consistent with our hypothesis (**Hypothesis 3a**), we found no evidence on the impact of the advertised show-up fee on participant's donation behavior or impact on the treatment effect in the donation task. In contrast to the previous lab experiment evidence, we do not find any significant average treatment effect (a one-for-one donation match as compared to no match).

3.4.2.3 Results from the p-beauty contest under cognitive load task

In the p-beauty contest under cognitive load task, 53 participants are randomly assigned to the treatment group and 52 participants are randomly assigned to the control group. The cognitive load is imposed successfully as 81.13% treated participants successfully recalled the 7-digit number and all control participants successfully recalled the 1-digit number. Figure 3.25 shows the distribution of the guess in the p-beauty contest

[16]Similar patterns and conclusions can be reached if we use the same bandwidth (the LSCV bandwidth selected for the treatment group) for the two groups as in Figures 3.23b and 3.24b.

from the treatment group with respect to the control group. We can see that the distribution shapes from the two groups are very different: the density from the treatment group is highly centered at around 22 while the density from the control group looks relatively platykurtic. And the empirical CDF plot suggests the distribution of guess from the control group seems to stochastically dominate that from the treatment group.

This conjecture is also supported by the results in Table 3.23: the average guess is higher (and the average level of strategic sophistication is lower) in the treatment group with higher cognitive load (mean = 38.270 vs. 29.679). Moreover, all the one-sided and two-sided tests fail to reject the null of equal mean or equal distribution of the guess from the two groups ($p < 0.01$ for the t-test, $p < 0.05$ for the Mann-Whitney test as well as the Kolmogorov-Smirnov test). Our finding about the average treatment effect of the higher cognitive load on the guess is consistent with the findings in strategic sophistication literature (Carpenter et al., 2013; Deck and Jahedi, 2015). Moreover, we find the guess from the treatment group stochastically dominates those from the control group in a statistically significant way.

To find out whether the advertised participation fee has an impact on subject's guess in this task and the related treatment effect, we first present the results of the LCLS regressions for the treatment group and the control group, respectively. As shown in Figure 3.26, the LSCV bandwidth selected for the treatment group is 1.75, while the LSCV bandwidth selected for the control group is 100. As result, the plot of the fitted values for the treatment group in Figure 3.27a has a slight curve-up when the show-up fee is less than around \$13 while the plot for the control group is almost flat across all show-up fee levels. Moreover, there is not much overlap of the 95% confidence intervals from the two groups and no overlap when the show-up fee level is below \$13. If we hold the bandwidth for the two groups to be the same (both use the LSCV bandwidth selected for the treatment group), the plots of the fitted values from the two groups become relatively parallel as in Figure 3.27b. As in the gradient plots in Figure 3.28, we find no evidence on significant

marginal effect of the show-up fee on the guess for either group or evidence on significant difference in the marginal effect between the two groups. From the plots of the fitted values and the gradients from the LCLS regression, we can infer that: (1) there might be significant treatment effect on the guess, and (2) the treatment effect might vary along different show-up fee levels.

We then consider using inference from the parametric regressions to verify our conjecture based on the nonparametric regression results. We present results separately for the treatment group and the control group and also jointly using the whole sample with the interaction term to test for existence of the effect of the advertised show-up fee on the treatment effect. The regression results in Table 3.24 from the linear regression model suggest that there is relatively small and insignificant impact of the advertised participation fee on the guess respectively for the two groups. The results in Table 3.25 using the whole sample show a significant treatment effect and marginally significant negative impact of the advertised show-up fee on the treatment effect: on average, imposing a higher cognitive load (memorizing a 7-digit number rather than a 1-digit number) results in a much higher guess by around 24.5, but this difference in guess will fade away as the subject are recruited with a higher advertised show-up fee (the difference in the guess will go down by 1.5 if the advertised show-up fee go up by \$1). However, while the AIC favors the model with the interaction term in Column 3, BIC suggests the model without the interaction term of treatment status and the show-up fee level might be more likely the correct model.

Since the dependent variable, the guess in this task, is always a positive variable, we then consider another parametric regression model – the linear regression model with log-level function forms. By taking the log of the guess, we obtain the results as in Tables 3.26 and 3.27. The results from regressions separately for the treatment and control group suggests that the advertised show-up fee has a negative and marginally statistically significant impact on the guess for the treated participants but not for the control participants: if the advertised show-up fee increased by \$1, the guess from the treated

subject will go down by 4.7 And the results from the regression using the whole sample in Table 3.27 suggest a significant treatment effect and marginally significant negative impact of the advertised show-up fee on the treatment effect: if the advertised show-up fee increases by $1, the treatment effect on the guess will go down by around 4.7%. Notice that while the AIC favors the model with the interaction term in Column 3, BIC suggests the model that only considers the treatment effect might be more likely the correct model.

Result 4. Based on the nonparametric regression results, we find no evidence of a statistically significant impact of the advertised show-up fee on the guess for either group based on the gradient plot though the LCLS regression results suggests treatment effect might vary with the advertised show-up fee. However, in contrast to our hypothesis (**Hypothesis 3a**) and some of the nonparametric regression results[17], the parametric regression results suggest that subjects who are recruited with a higher advertised participation fee perform less sophisticated (i.e., giving a lower guess) under a higher cognitive load treatment as compared to those who recruited with a lower show-up fee level[18]. In addition, our findings about the average treatment effect (higher cognitive load) are consistent with the previous lab experiment evidence.

3.5 Conclusion

Lab experiments have been an important tool in understanding individuals' economic behaviors and testing the predictive power of economic theory. To generalize the experiment results, researchers have designed strategies to reduce participation bias by varying the show-up fee and using it as exclusion restriction in the sample selection correction approach. In this study, we design a lab experiment by focusing on the impact

[17]Note that the rate of convergence is slower for the gradient estimates than the conditional mean of outcome. The gradients are estimated less precisely (as shown in the gradient plots) than the conditional mean of outcome (as shown in the fitted value plots).

[18]One might come up with a plausible reason to explain this phenomenon: those who recruited via a lower advertised participation fee consider the difference in the actual participation fee and the advertised participation fee as windfall money. As a result, those participants might have less incentive and put less effort in the guessing game when under higher cognitive load.

of the advertised show-up fee and holding the actual show-up fee the same for all lab participants. We investigate if the variation in the advertised show-up fee does always satisfy the exclusion restriction: (1) whether a higher advertised show-up fee will encourage more individuals to participate in the lab experiments, and (2) whether the advertised show-up fee will have a direct impact on the lab participant's decision-making and treatment effect observed in the experiment.

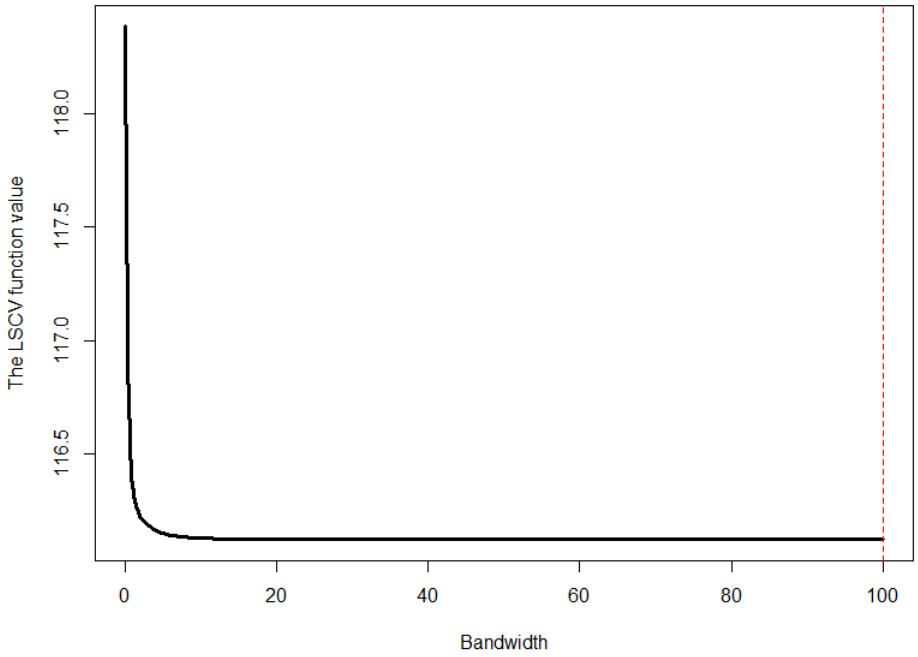
The implications of our findings are two-fold. First, we find no impact of the variation of show-up fee on individual's registration decision or participation decision in the range of show-up fee levels that are commonly used in the lab experiment study. This suggests that a higher show-up fee level might not results in variation in participant's characteristics and the experiment evidence should be robust at these show-up fee levels. Secondly, our experiment results show that the advertised show-up fee has no direct impact on lab participant's decision-making or treatment effect in the lab experiment tasks that relates to individual's social preference and risk attitude. However, our results suggest that those recruited with a higher advertised participation fee perform less sophisticated under a higher cognitive load. This suggests that caution should be made when we incorporate the randomization and variation of advertised show-up fee (or show-up fee) in the experiment design to correct for participation bias. For example, we might be able to use the randomization and variation in the advertised show-up fee as exclusion restriction for studies that relates to individual's social preference and risk attitude but not for studies that requires a lot of cognitive effort from the participants or those that imposing a high cognitive load.

Based on the evidence in our study, future studies should also look at broader range of participation fee levels to see whether and when individual's participation decisions are sensitive to the variation in the show-up fee. This will be crucial to the experiment design that aims to reduce the participation bias (e.g., the bias caused by the possibility that lower-income/more pro-social individuals are more likely to participate in the study).

Moreover, future studies should also examine for which kinds of experiments the results are subject to participation bias and for which kinds of experiments the results will not be directly affected by the advertised show-up fee[19]. This will give us a better understanding about when and how we should use the randomization and variation in the advertised show-up fee as the exclusion restriction to correct for participation bias.

[19]For some studies, we can just make the actual show-up fee the same as the advertised show-up fee if the results are robust to different endowment in the experiment.

Figure 3.1: LSCV bandwidth selection for the LCLS regression (registration decision on advertised show-up fee)



Note: The LSCV bandwidth, i.e., the bandwidth for the LCLS regression that minimize the LSCV function, is chosen from 0.1 to 100 with step = 0.01. The LSCV bandwidth might go infinity if we keep increasing the upper bound of the search range.

Figure 3.2: Scatter plot of the advertised show-up fee versus registration decision with the fitted values from the LCLS regression using LSCV bandwidth

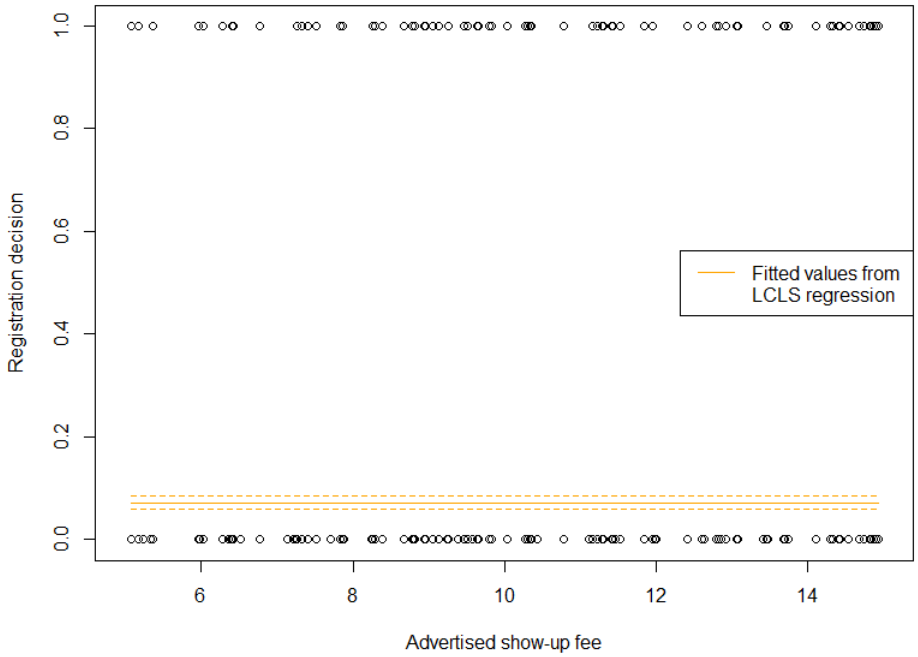


Figure 3.3: Gradient plots from the LCLS regression (registration decision on advertised show-up fee) using LSCV bandwidth

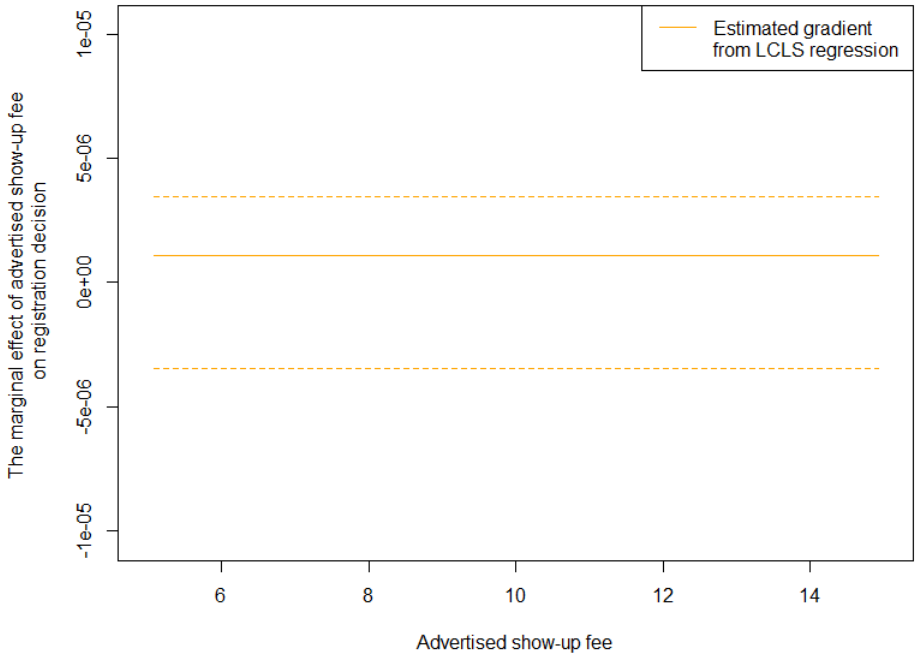
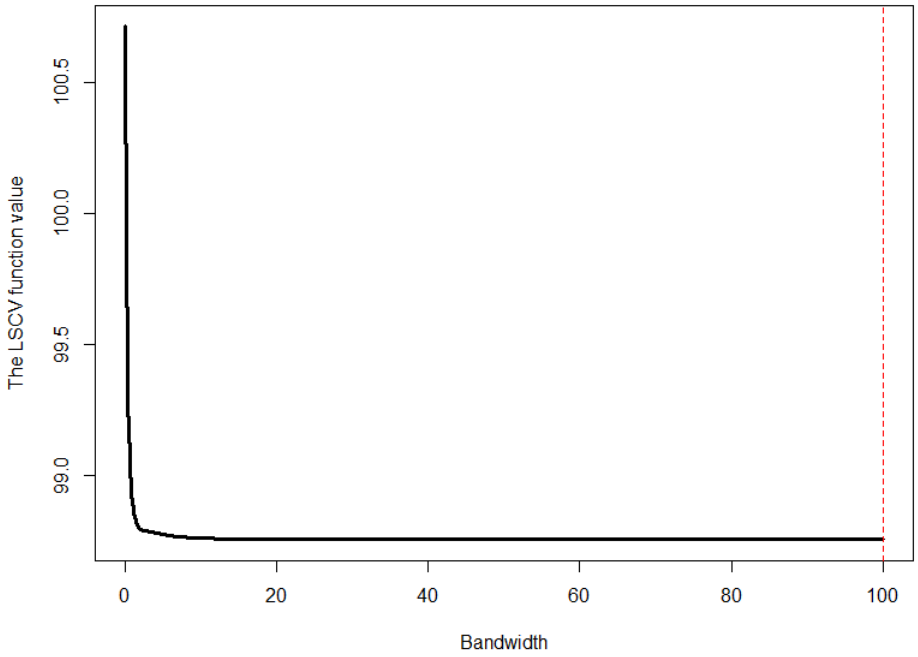


Table 3.1: Parametric regression results of registration decision on advertised show-up fee

	Dependent variable:			
	Registration decision			
	Probit model		Linear probability model	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.009		0.001
		(0.015)		(0.002)
Constant	−1.460***	−1.559***	0.072***	0.059**
	(0.045)	(0.168)	(0.006)	(0.023)
Observations	1,734	1,734	1,734	1,734
Log likelihood	−449.116	−448.932	−116.473	−116.291
AIC	900.232	901.863	234.946	236.581
BIC	905.691	912.780	245.862	252.956
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.4: LSCV bandwidth selection for the LCLS regression (participation decision on advertised show-up fee)



Note: The LSCV bandwidth, i.e., the bandwidth for the LCLS regression that minimize the LSCV function, is chosen from 0.1 to 100 with step = 0.01. The LSCV bandwidth might go infinity if we keep increasing the upper bound of the search range.

Figure 3.5: Scatter plot of the advertised show-up fee versus participation decision with the fitted values from the LCLS regression using LSCV bandwidth

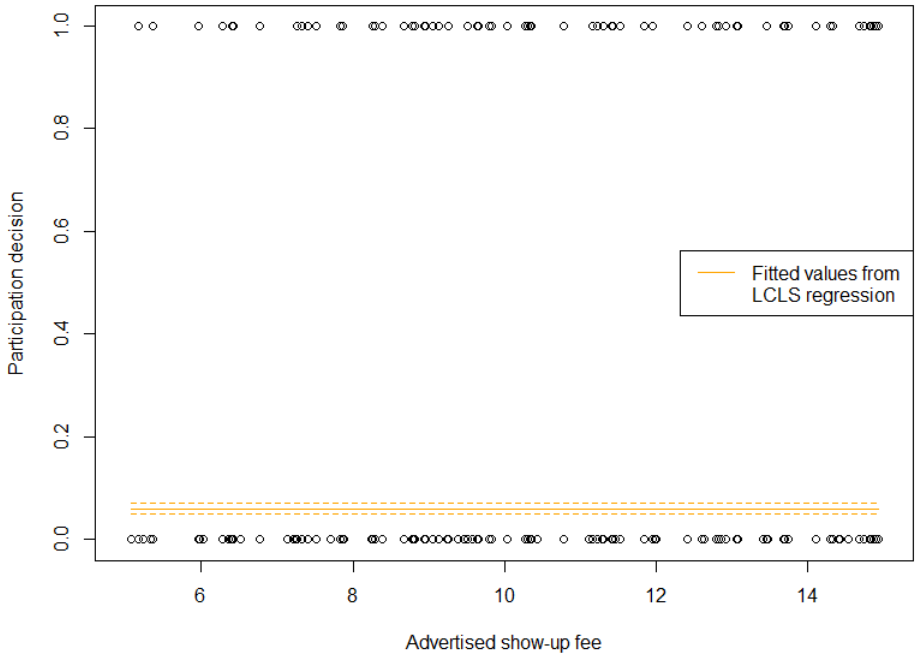


Figure 3.6: Gradient plots from the LCLS regression (participation decision on advertised show-up fee) using LSCV bandwidth

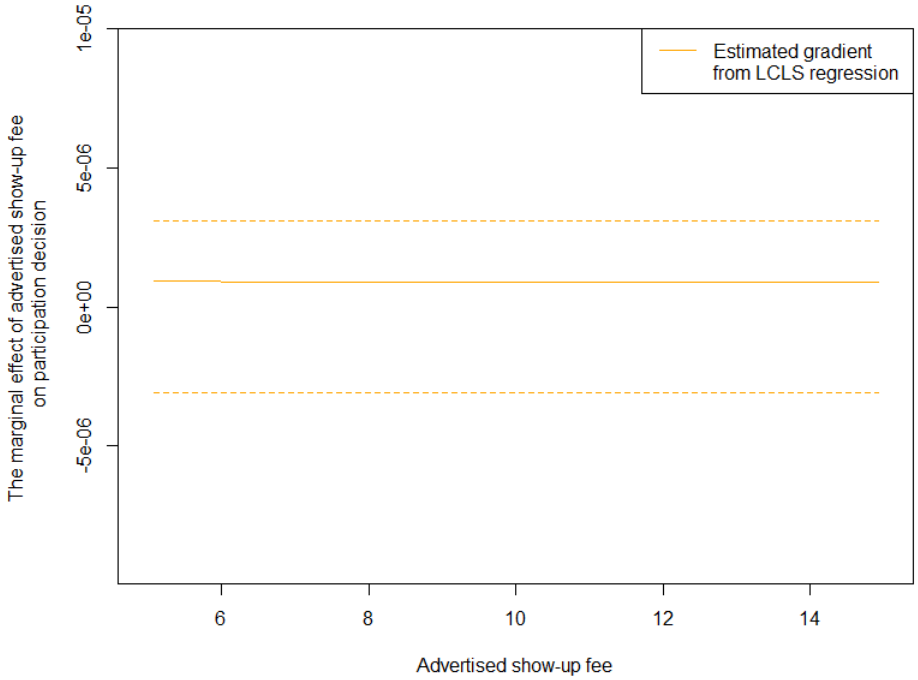
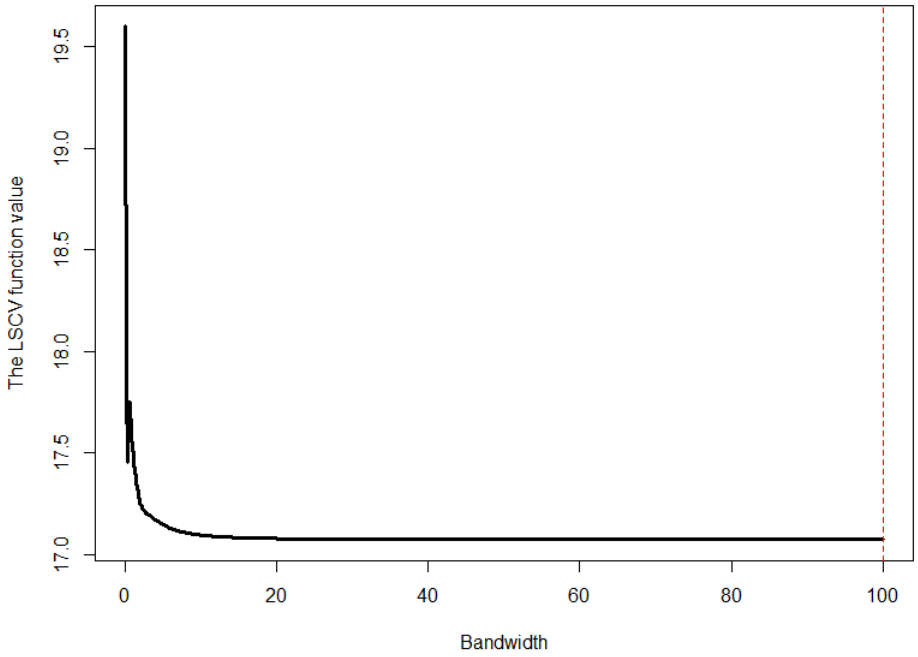


Table 3.2: Parametric regression results of participation decision on advertised show-up fee

	Dependent variable:			
	Participation decision			
	Probit model		Linear probability model	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.009		0.001
		(0.016)		(0.002)
Constant	−1.550***	−1.644***	0.061***	0.050**
	(0.048)	(0.177)	(0.006)	(0.021)
Observations	1,734	1,734	1,734	1,734
Log likelihood	−396.198	−396.048	23.981	24.128
AIC	794.397	796.097	−45.962	−44.257
BIC	799.855	807.013	-35.046	-27.882
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.7: LSCV bandwidth selection for the LCLS regression (participation decision conditional on registration on advertised show-up fee)



Note: The LSCV bandwidth, i.e., the bandwidth for the LCLS regression that minimize the LSCV function, is chosen from 0.1 to 100 with step = 0.01. The LSCV bandwidth might go infinity if we keep increasing the upper bound of the search range.

Figure 3.8: Scatter plot of the advertised show-up fee versus participation decision conditional on registration with the fitted values from LCLS regression

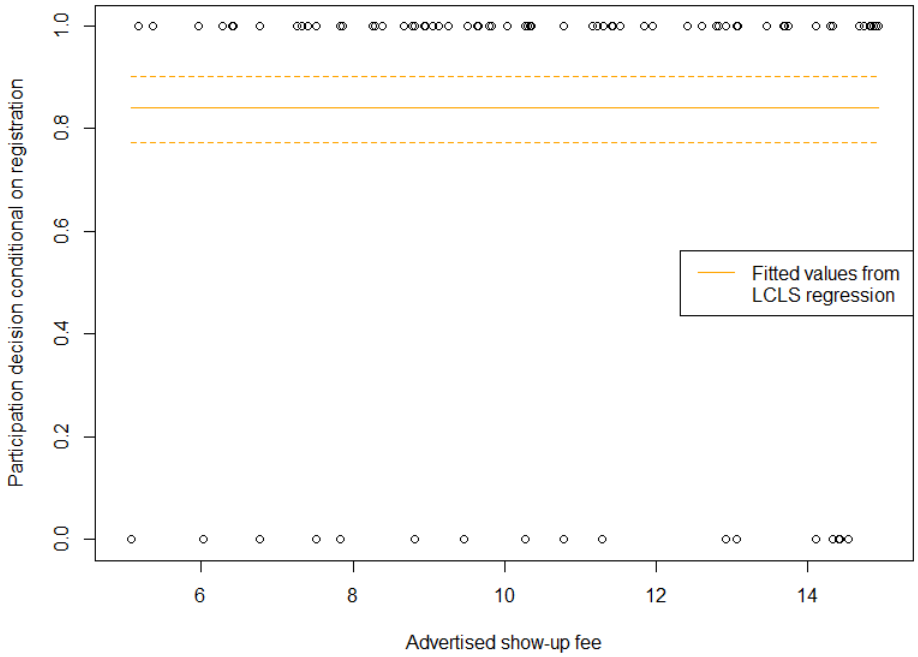


Figure 3.9: Gradient plots from the LCLS regression (participation decision conditional on registration on advertised show-up fee)

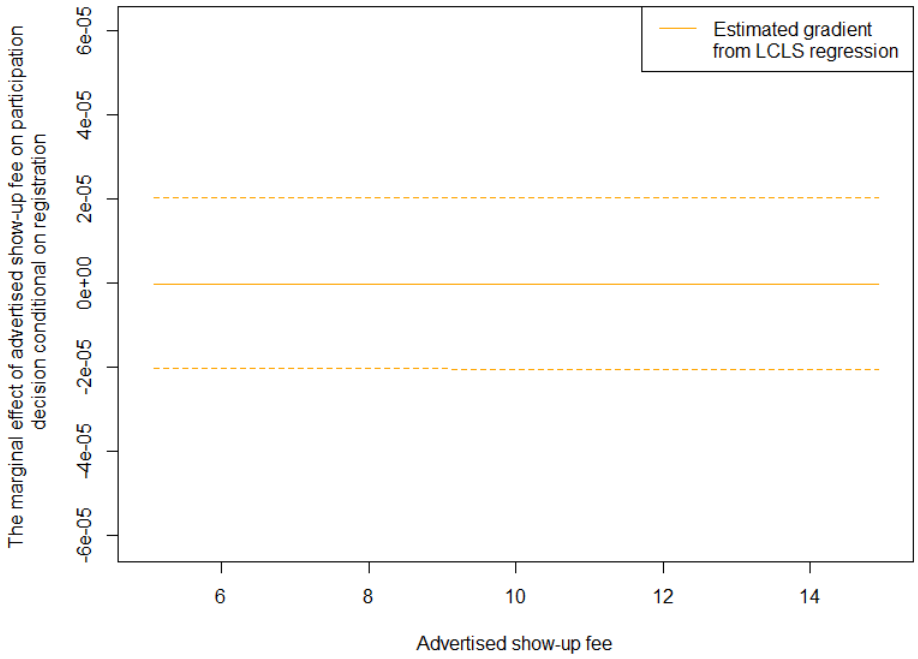
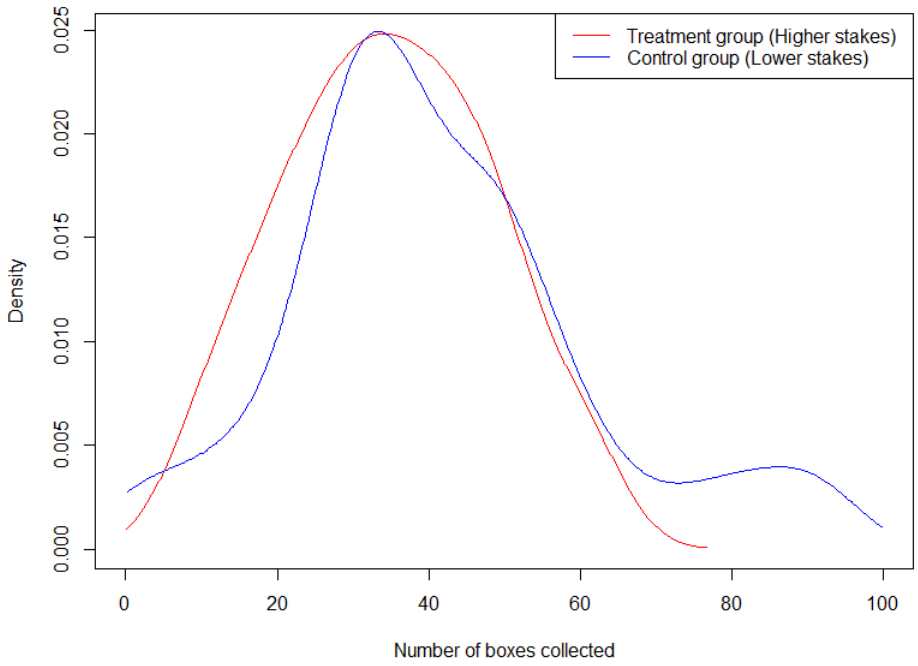


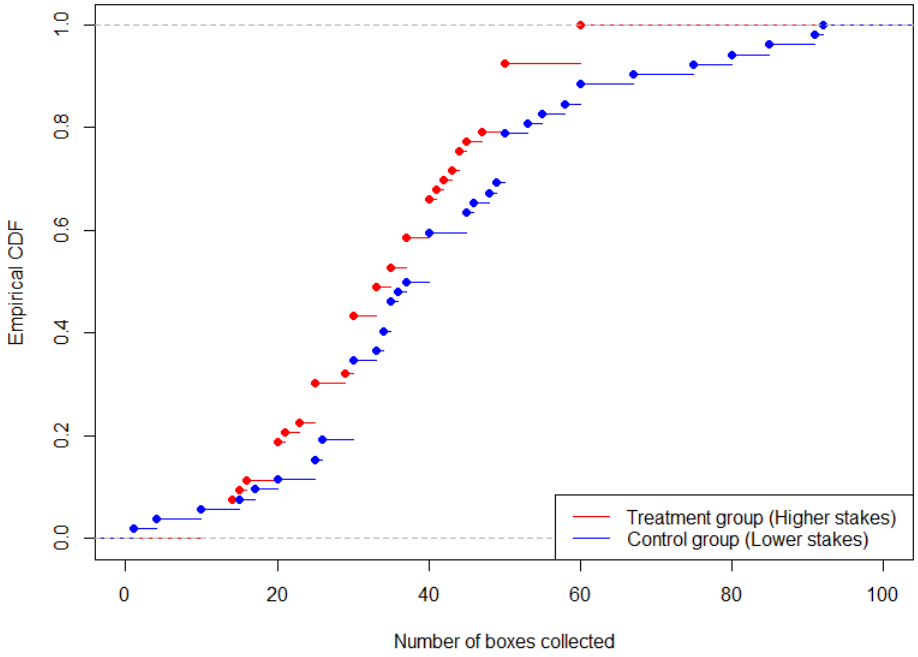
Table 3.3: Parametric regression results of participation decision conditional on registration on advertised show-up fee

	Dependent variable:			
	Participation decision conditional on registration			
	Probit model		Linear probability model	
	(1)	(2)	(3)	(4)
Advertised show-up fee		−0.001		−0.000
		(0.048)		(0.012)
Constant	0.994***	1.003*	0.840***	0.842***
	(0.135)	(0.521)	(0.033)	(0.128)
Observations	125	125	125	125
Log likelihood	−54.959	−54.959	−52.934	−52.934
AIC	111.917	113.917	107.868	109.867
BIC	114.746	119.574	113.5244	118.352
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.10: Kernel density and empirical CDF plots of the risk bomb task outcome by treatment



(a) Kernel density



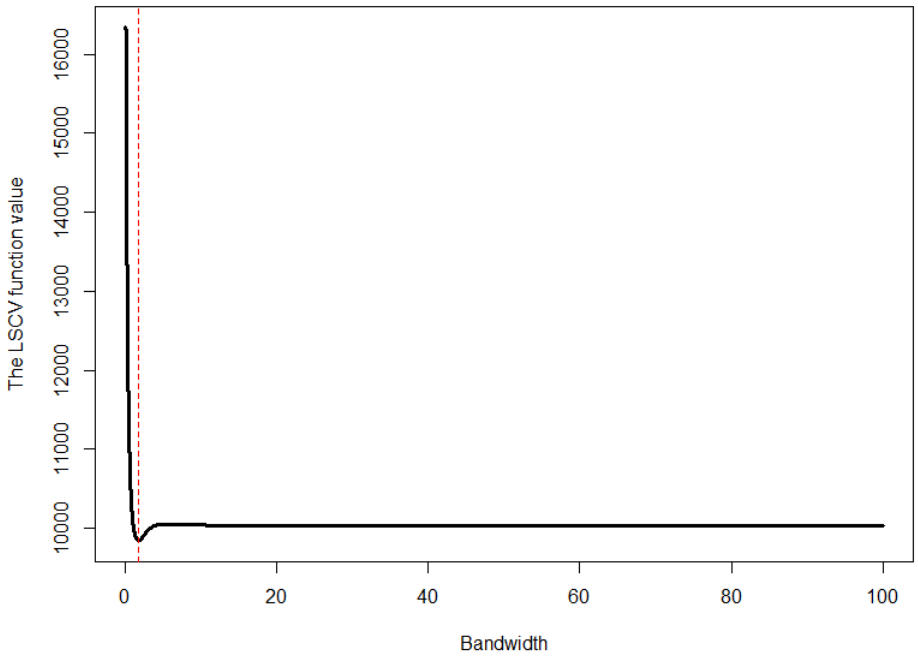
(b) Empirical CDF

Table 3.4: Results of the risk bomb task outcome by treatment

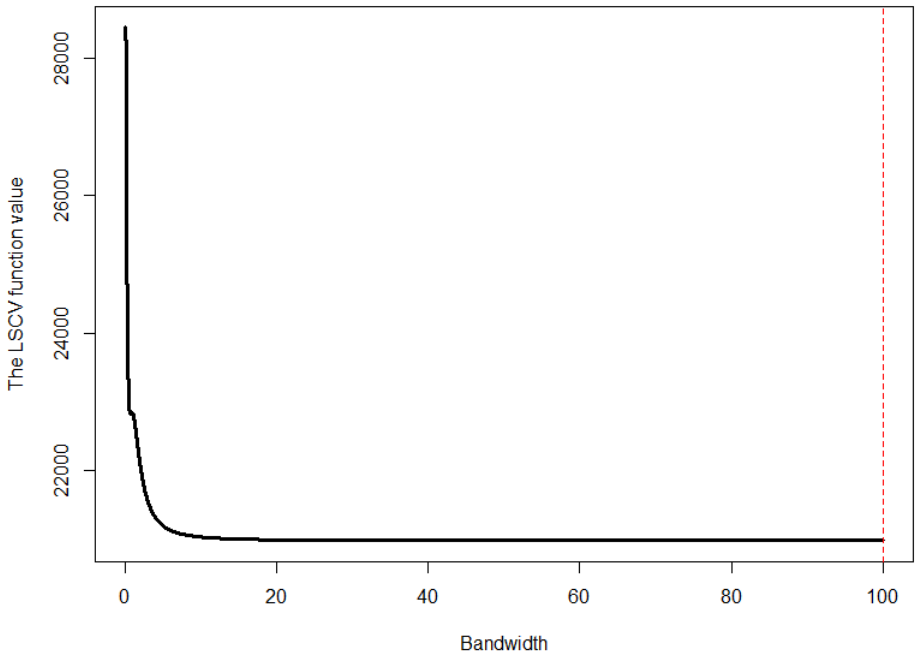
	N	Mean	T-test for equal means	Mann-Whitney test	Kolmogorov-Smirnov test
Treatment group (Higher stakes)	53	34.792 (13.626)	p = 0.054 (H_1: unequal mean) $p = 0.027$	p = 0.120 (H_1: unequal mean) $p = 0.060$	$p = 0.524$ (H_1: unequal distribution)
Control group (Lower stakes)	52	41.288 (19.892)	(H_1: control group has a higher mean)	(H_1: control group has a higher mean)	$p = 0.267$ (H_1: the distribution in the control group stochastically dominates that in the treatment group)

Note: Standard errors are shown in the parentheses.

Figure 3.11: LSCV bandwidth selection for the LCLS regression (risk bomb task outcome on advertised show-up fee)



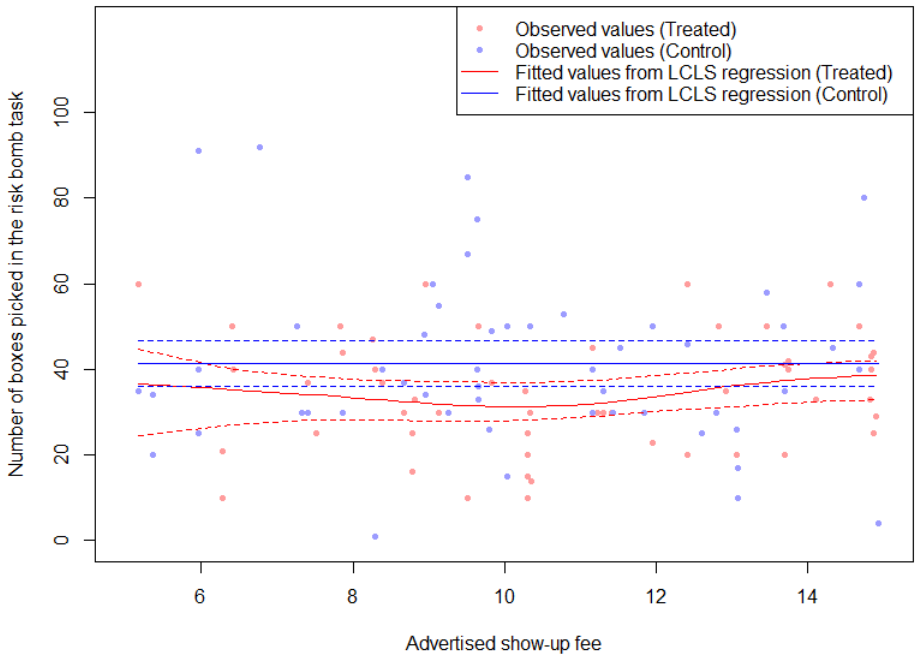
(a) LSCV bandwidth selection for the treatment group



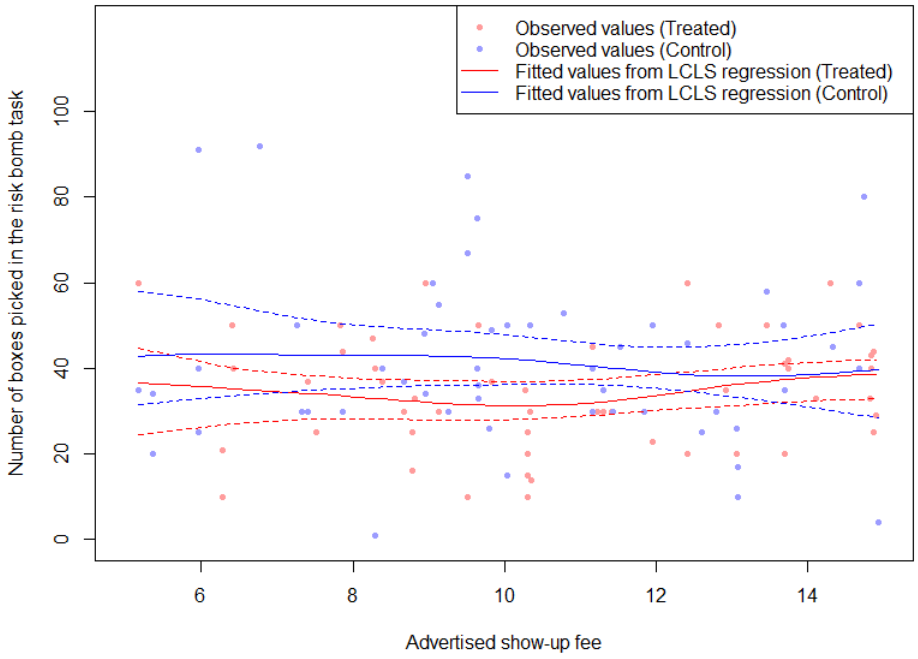
(b) LSCV bandwidth selection for the control group

Note: The LSCV bandwidth for the treatment group is 1.77 based on the LCLS regression for the treatment group, while the LSCV bandwidth for the control group is 100 based on the LCLS regression for the control group. The LSCV bandwidth for the control group might go infinity if we keep increasing the upper bound of the search range.

Figure 3.12: Scatter plots with the fitted values from LCLS regressions (risk bomb task outcome on advertised show-up fee) using LSCV bandwidths

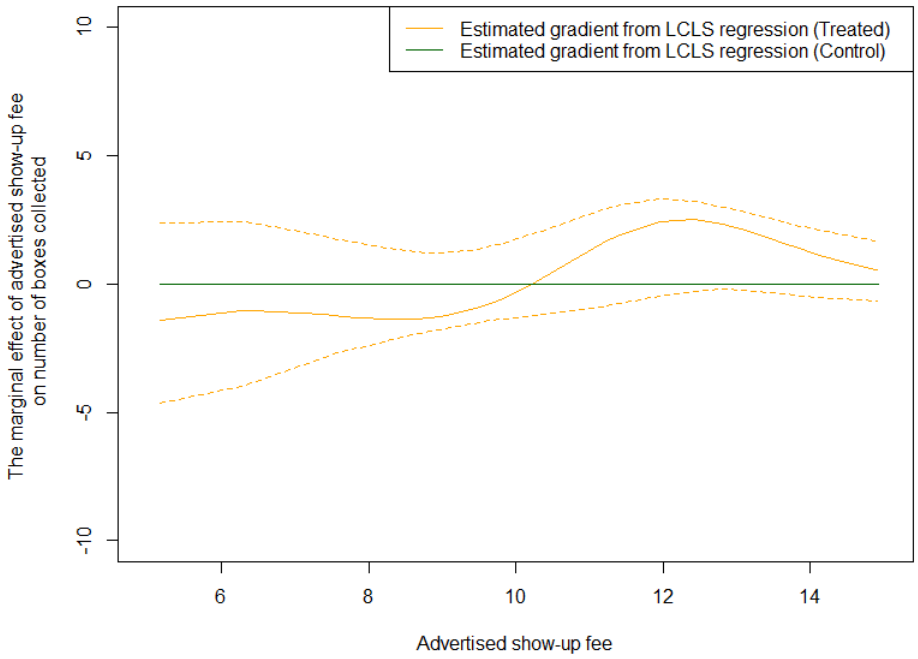


(a) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.77$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)

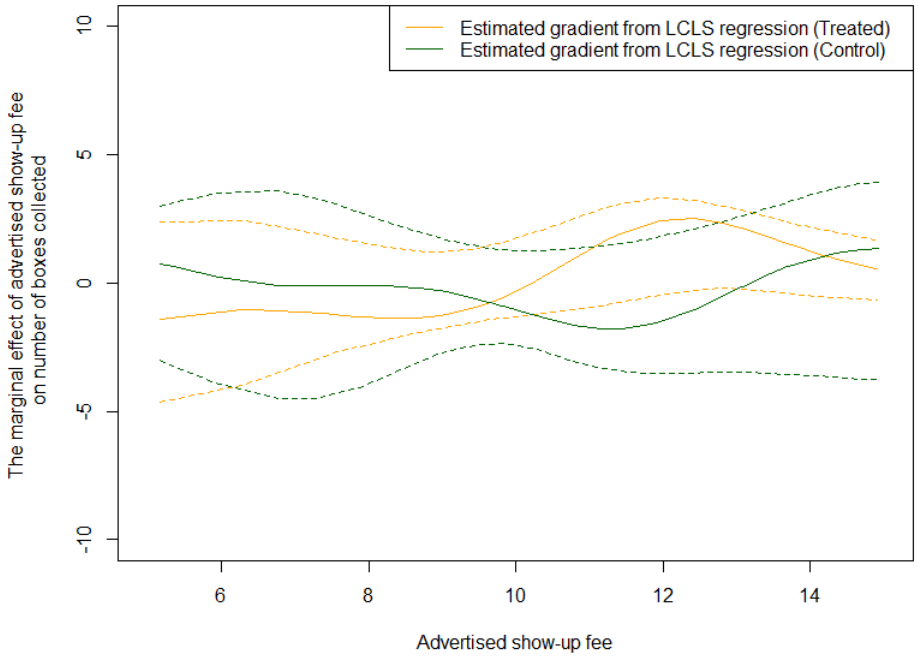


(b) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.77$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.77$)

Figure 3.13: Gradient plots from LCLS regressions (risk bomb task outcome on advertised show-up fee) using LSCV bandwidths



(a) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.77$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)



(b) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.77$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.77$)

Table 3.5: Results of the risk bomb task based on the linear regression model, separately for the treatment and control groups

	Dependent variable:			
	Number of boxes collected			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.475		−0.521
		(0.680)		(1.032)
Constant	34.792***	29.638***	41.288***	46.606**
	(1.872)	(7.618)	(2.759)	(10.907)
Observations	53	53	52	52
Log likelihood	−214.134	−213.882	−229.777	−229.645
AIC	430.268	431.763	461.554	463.290
BIC	434.208	437.674	465.456	469.144
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.6: Results of the risk bomb task based on the linear regression model using the whole sample

	Dependent variable:			
	Number of boxes collected			
	(1)	(2)	(3)	(4)
Treatment		−6.496*	−11.650	−16.968
		(3.322)	(9.791)	(13.333)
Treatment × advertised show-up fee			0.475	0.995
			(0.848)	(1.226)
Advertised show-up fee				−0.521
				(0.883)
Constant	38.010***	41.288***	41.288***	46.606***
	(1.683)	(2.360)	(2.368)	(9.324)
Observations	105	105	105	105
Log likelihood	−448.500	−446.587	−446.426	−446.245
AIC	899.001	897.174	898.851	900.490
BIC	904.309	905.135	909.467	913.760
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.7: Results of the risk bomb task based on the negative binomial regression model, separately for the treatment and control groups

	Dependent variable:			
	Number of boxes collected			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.013		−0.013
		(0.020)		(0.027)
Constant	3.549***	3.412***	3.721***	3.850***
	(0.057)	(0.230)	(0.073)	(0.286)
Observations	53	53	52	52
Log likelihood	−215.134	−214.929	−231.054	−230.945
AIC	432.268	433.859	464.107	465.890
BIC	436.209	439.770	468.010	471.743
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.8: Results of the risk bomb task based on the negative binomial regression model using the whole sample

	Dependent variable:			
	Number of boxes collected			
	(1)	(2)	(3)	(4)
Treatment		−0.171*	−0.308	−0.437
		(0.092)	(0.272)	(0.366)
Treatment × advertised show-up fee			0.013	0.025
			(0.024)	(0.034)
Advertised show-up fee				−0.013
				(0.024)
Constant	3.638***	3.721***	3.721***	3.850***
	(0.047)	(0.065)	(0.065)	(0.254)
Observations	105	105	105	105
Log likelihood	−448.492	−446.793	−446.638	−446.501
θ	4.912***	5.101***	5.119***	5.135***
	(0.776)	(0.812)	(0.816)	(0.819)
AIC	898.984	897.586	899.276	901.001
BIC	904.292	905.548	909.892	914.271
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.14: Empirical CDF plot of donation amount by treatment

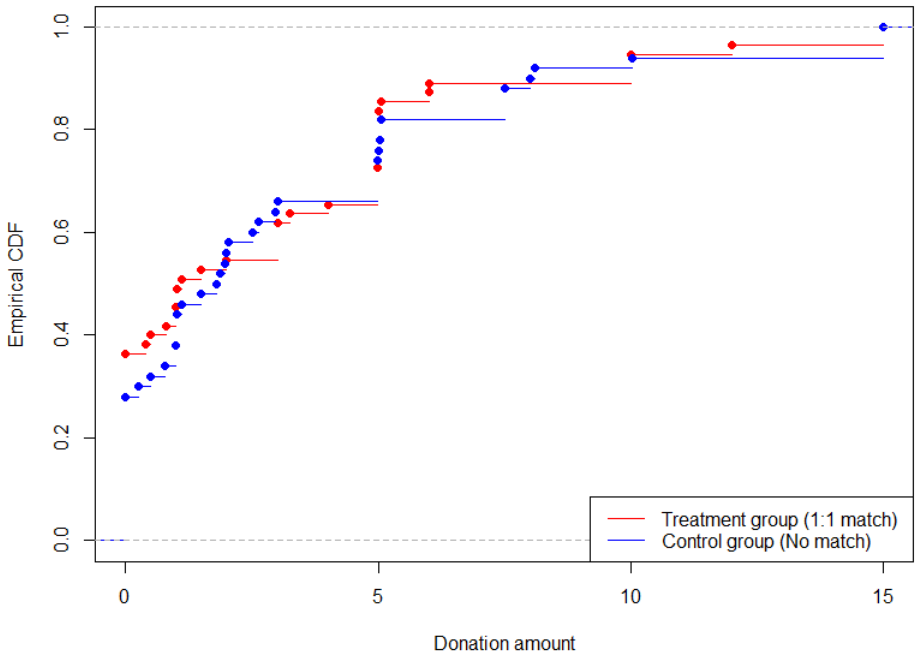
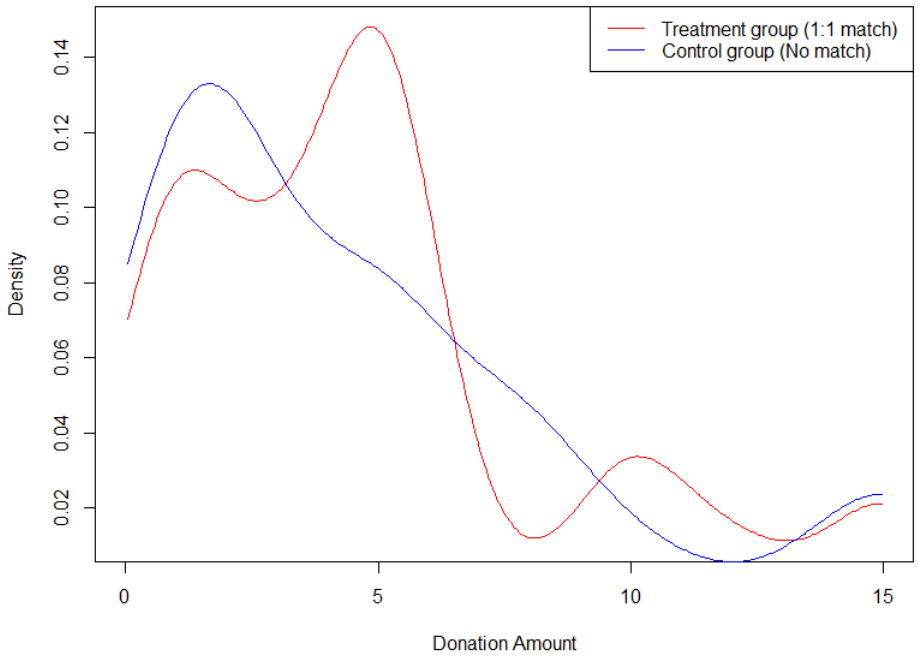


Table 3.9: Results of the donation amount by treatment

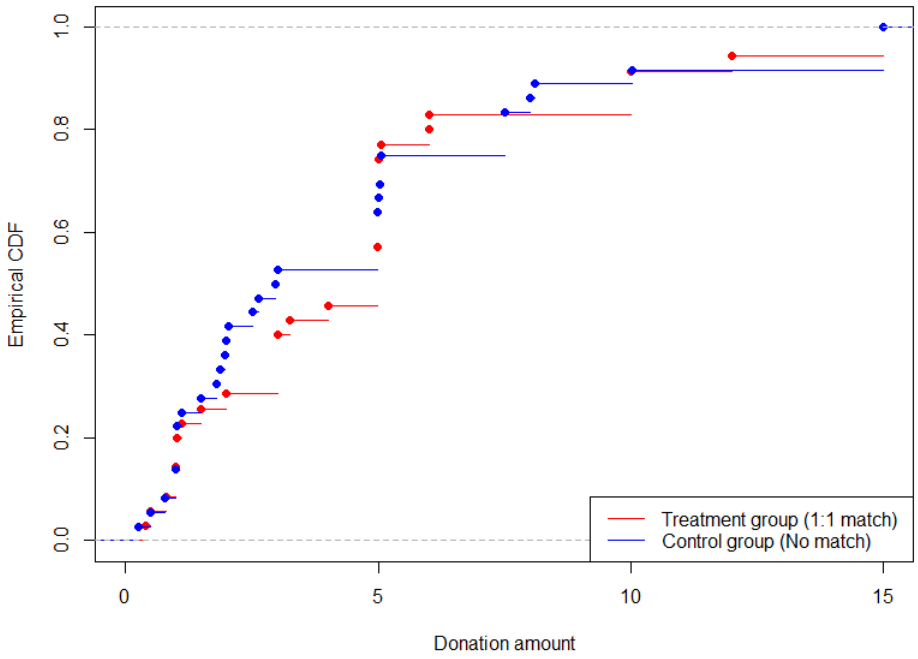
	N	Mean	T-test for equal means	Mann-Whitney test	Kolmogorov-Smirnov test
Treatment group (1:1 match)	53	3.066 (3.850)	$p = 0.790$ (H_1: unequal mean) $p = 0.395$	$p = 0.641$ (H_1: unequal mean) $p = 0.320$	$p = 0.973$ (H_1: unequal distribution)
Control group (No match)	52	3.273 (4.037)	(H_1: control group has a higher mean)	(H_1: control group has a higher mean)	$p = 0.693$ (H_1: the distribution in the control group stochastically dominates that in the treatment group)

Note: Standard errors are shown in the parentheses.

Figure 3.15: Kernel density and empirical CDF plots of donation amount conditional on giving by treatment



(a) Kernel density



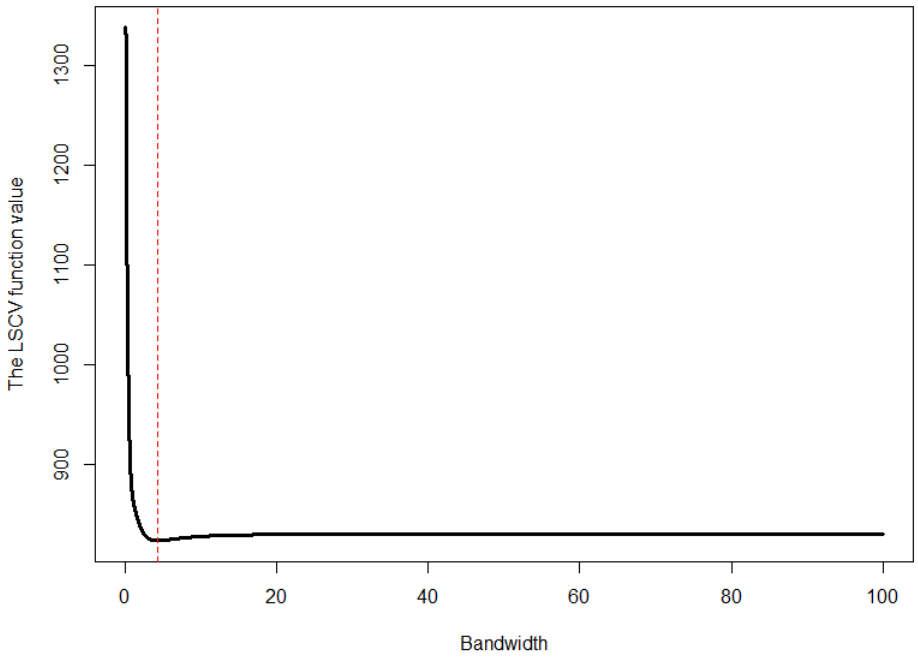
(b) Empirical CDF

Table 3.10: Results of the donation amount conditional on giving by treatment

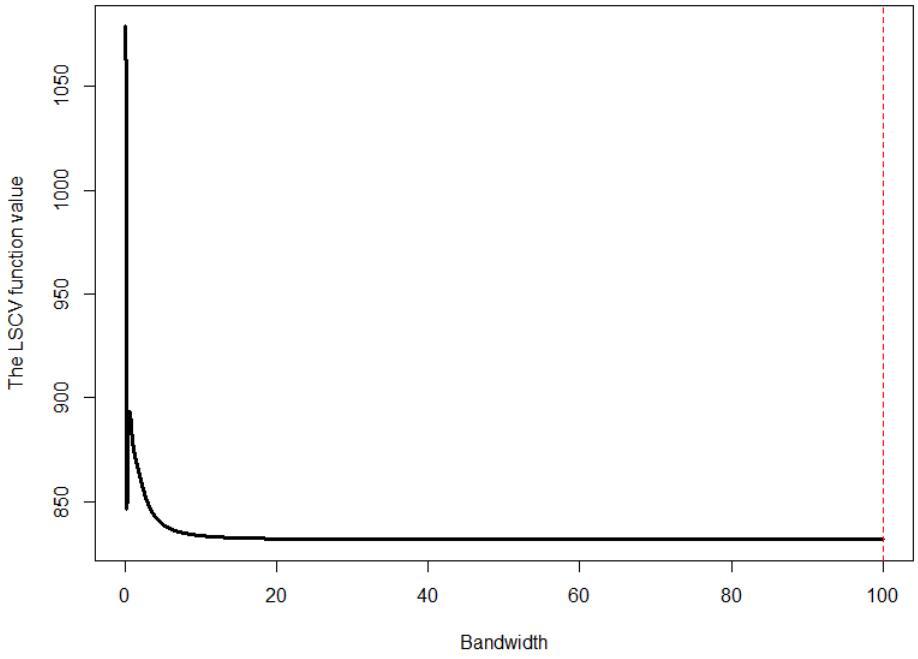
	N	Mean	T-test for equal means	Mann-Whitney test	Kolmogorov-Smirnov test
Treatment group (1:1 match)	35	4.818 (3.854)	$p = 0.773$ (H_1: unequal mean)	$p = 0.624$ (H_1: unequal mean)	$p = 0.389$ (H_1: unequal distribution)
Control group (No match)	36	4.545 (4.107)	$p = 0.387$ (H_1: treatment group has a higher mean)	$p = 0.312$ (H_1: treatment group has a higher mean)	$p = 0.196$ (H_1: the distribution in the treatment group stochastically dominates that in the control group)

Note: Standard errors are shown in the parentheses.

Figure 3.16: LSCV bandwidth selection for the LCLS regression (donation amount including zeros on advertised show-up fee



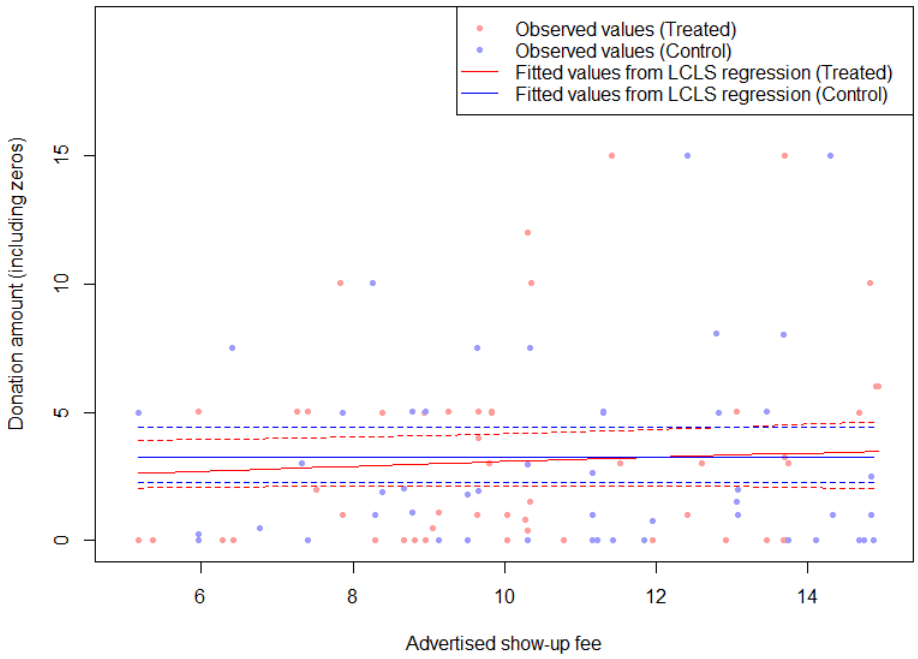
(a) LSCV bandwidth selection for the treatment group



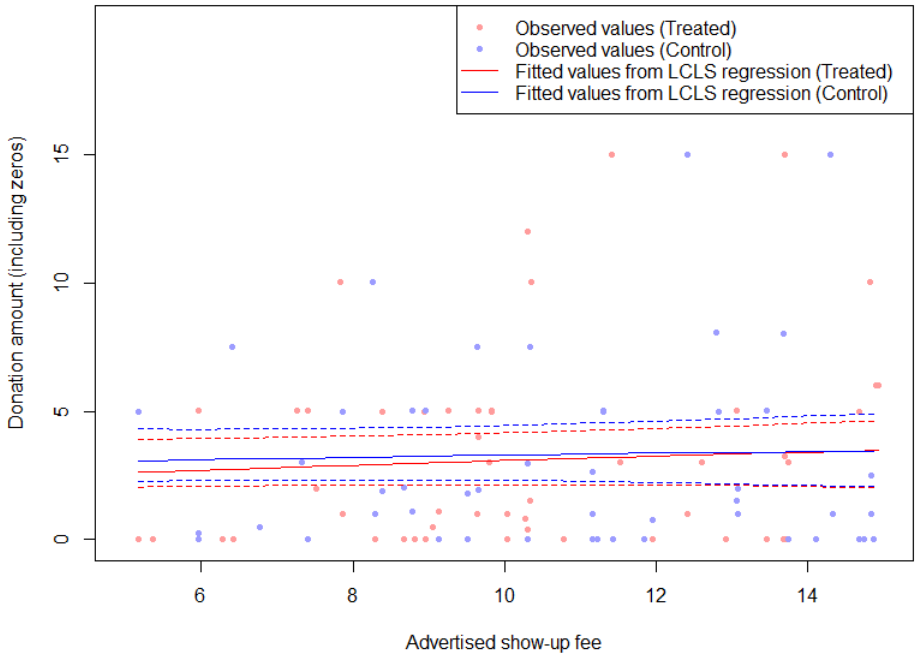
(b) LSCV bandwidth selection for the control group

Note: The LSCV bandwidth for the treatment group is 4.27 based on the LCLS regression for the treatment group, while the LSCV bandwidth for the control group is 100 based on the LCLS regression for the control group. The LSCV bandwidth for the control group might go infinity if we keep increasing the upper bound of the search range.

Figure 3.17: Scatter plots with the fitted values from LCLS regressions (donation amount including zeros on advertised show-up fee) using LSCV bandwidths

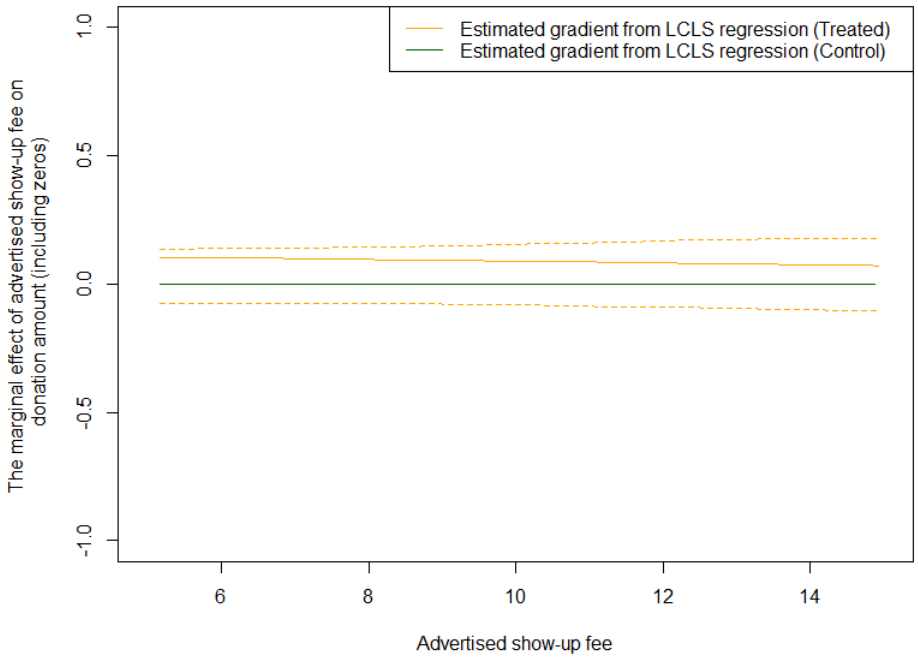


(a) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 4.27$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)

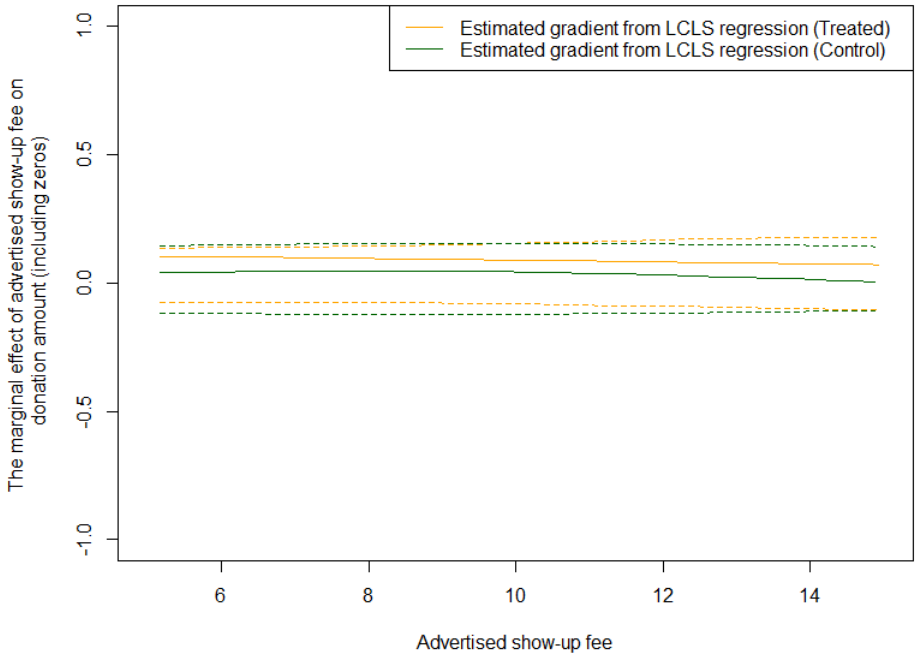


(b) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 4.27$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 4.27$)

Figure 3.18: Gradient plots from LCLS regressions (donation amount including zeros on advertised show-up fee) using LSCV bandwidths



(a) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 4.27$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)



(b) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 4.27$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 4.27$)

Table 3.11: Results on the donation amount including zeros from the linear regression model, separately for the treatment and control groups

	Dependent variable:			
	Donation amount (including zeros)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.289		0.080
		(0.185)		(0.214)
Constant	3.067***	0.100	3.273***	2.409
	(0.519)	(1.968)	(0.571)	(2.385)
Observations	55	55	50	50
Log likelihood	−152.682	−151.446	−141.218	−141.145
AIC	307.364	306.892	284.435	286.290
BIC	311.379	312.914	288.259	292.0262

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.12: Results on the donation amount (including zeros) from the linear regression model using the whole sample

	Dependent variable:			
	Donation amount (including zeros)			
	(1)	(2)	(3)	(4)
Treatment		−0.206	−3.172	−2.309
		(0.770)	(2.103)	(3.074)
Treatment × advertised show-up fee			0.289	0.209
			(0.191)	(0.281)
Advertised show-up fee				0.080
				(0.206)
Constant	3.165***	3.273***	3.273***	2.409
	(0.383)	(0.557)	(0.554)	(2.302)
Observations	105	105	105	105
Log likelihood	−292.993	−292.956	−291.788	−291.711
AIC	587.985	589.913	589.577	591.422
BIC	593.293	597.874	600.193	604.691

Note: *p<0.1; **p<0.05; ***p<0.01

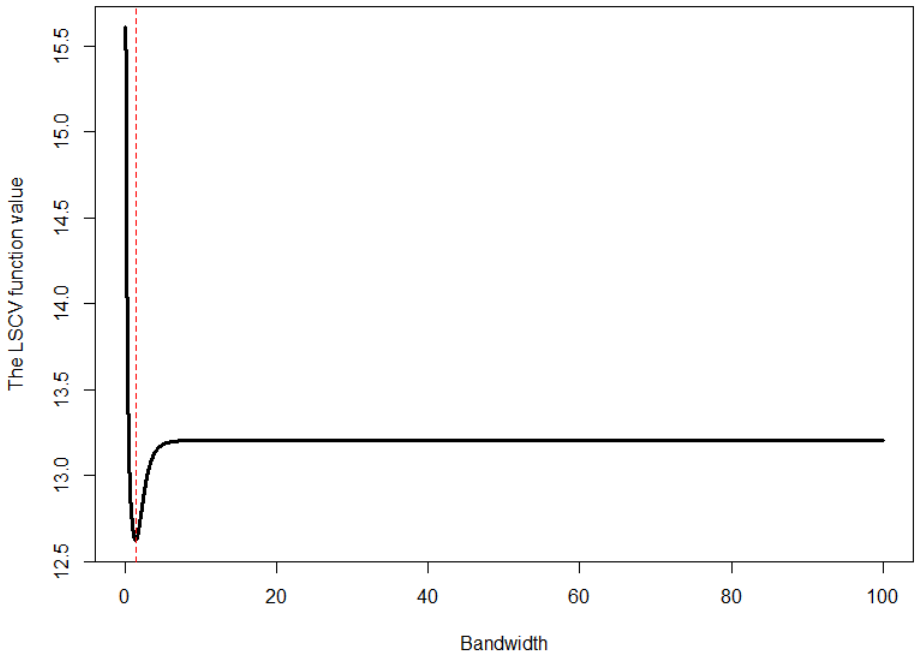
Table 3.13: Results on the donation amount (including zeros) from the linear regression model with a log-level function, separately for the treatment and control groups

	Dependent variable:			
	ln (Donation amount (including zeros))			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.126		−0.040
		(0.090)		(0.093)
Constant	−0.066	−1.356	0.137	0.568
	(0.251)	(0.957)	(0.248)	(1.038)
Observations	55	55	50	50
Log likelihood	−112.801	−111.809	−99.626	−99.531
AIC	227.603	227.617	201.252	203.061
BIC	231.617	233.639	205.076	208.797
Note:				*p<0.1; **p<0.05; ***p<0.01

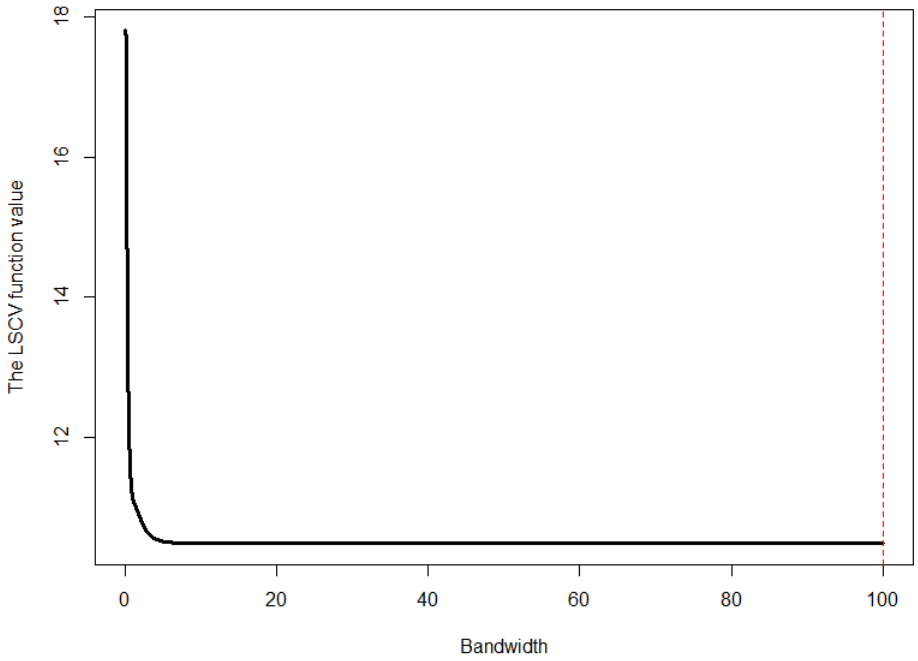
Table 3.14: Results on the donation amount (including zeros) from the linear regression model with a log-level function using the whole sample

	Dependent variable:			
	ln (Donation amount (including zeros))			
	(1)	(2)	(3)	(4)
Treatment		−0.203	−1.493	−1.924
		(0.354)	(0.969)	(1.417)
Treatment × advertised show-up fee			0.126	0.165
			(0.088)	(0.130)
Advertised show-up fee				−0.040
				(0.095)
Constant	0.030	0.137	0.137	0.568
	(0.176)	(0.257)	(0.255)	(1.061)
Observations	105	105	105	105
Log likelihood	−211.688	−211.522	−210.481	−210.390
AIC	425.377	427.044	426.962	428.780
BIC	430.685	435.006	437.578	442.050
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.19: LSCV bandwidth selection for the LCLS regression (binary decision to give on advertised show-up fee)



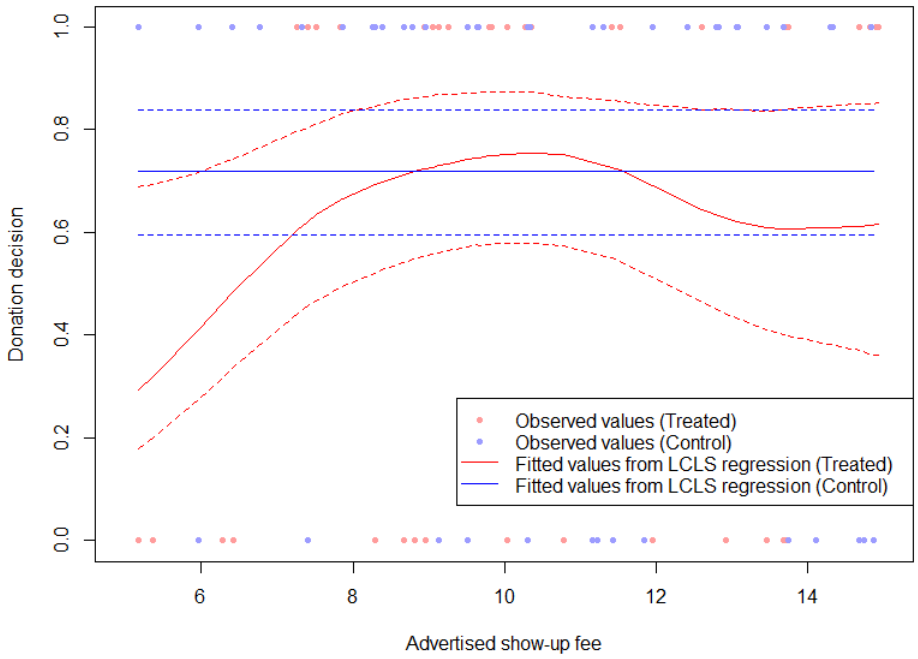
(a) LSCV bandwidth selection for the treatment group



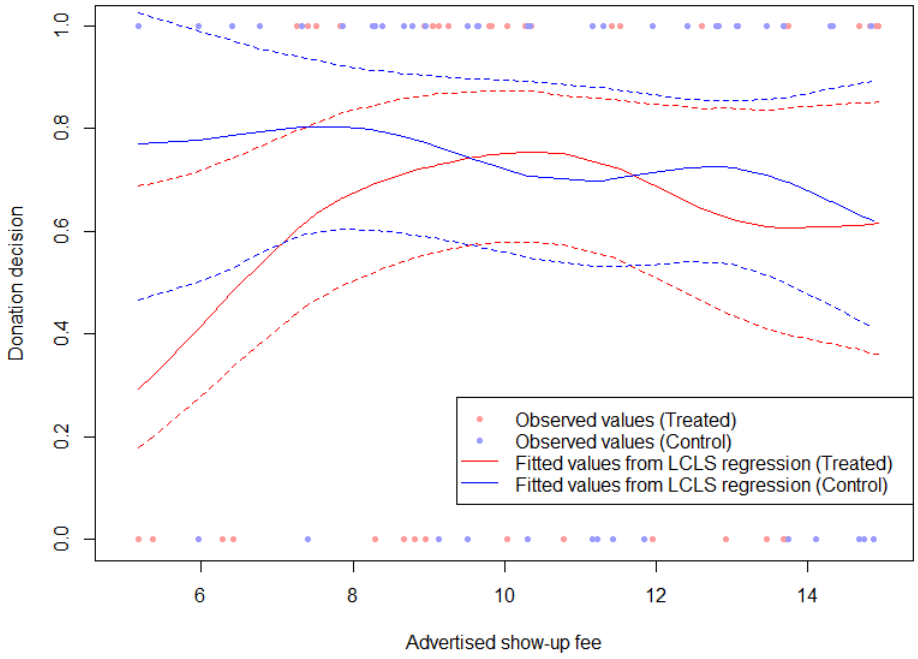
(b) LSCV bandwidth selection for the control group

Note: The LSCV bandwidth for the treatment group is 1.40 based on the LCLS regression for the treatment group, while the LSCV bandwidth for the control group is 100 based on the LCLS regression for the control group. The LSCV bandwidth for the control group might go infinity if we keep increasing the upper bound of the search range.

Figure 3.20: Scatter plots with the fitted values from LCLS regressions (binary decision to give on advertised show-up fee) using LSCV bandwidths

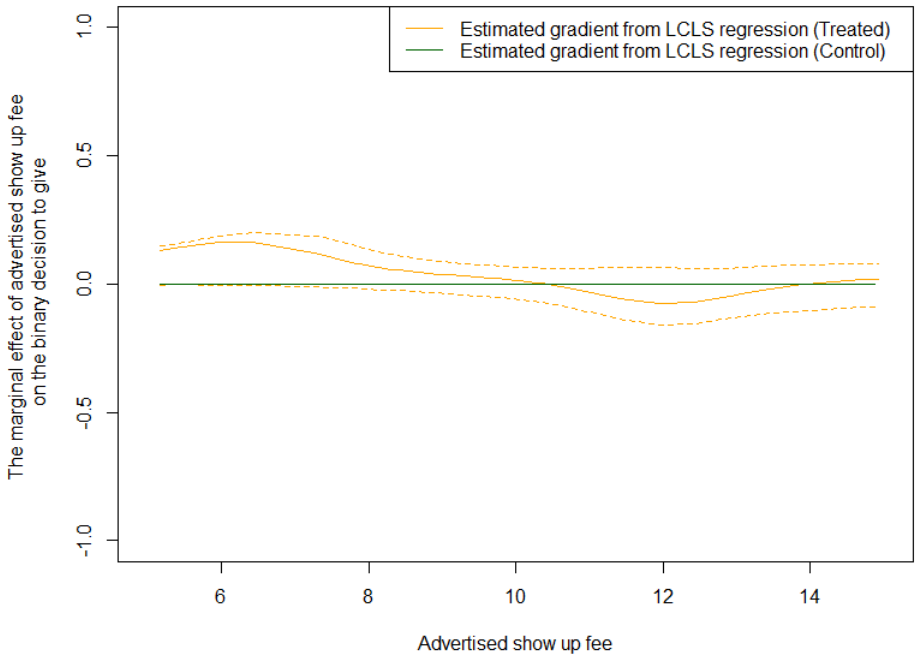


(a) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.40$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)

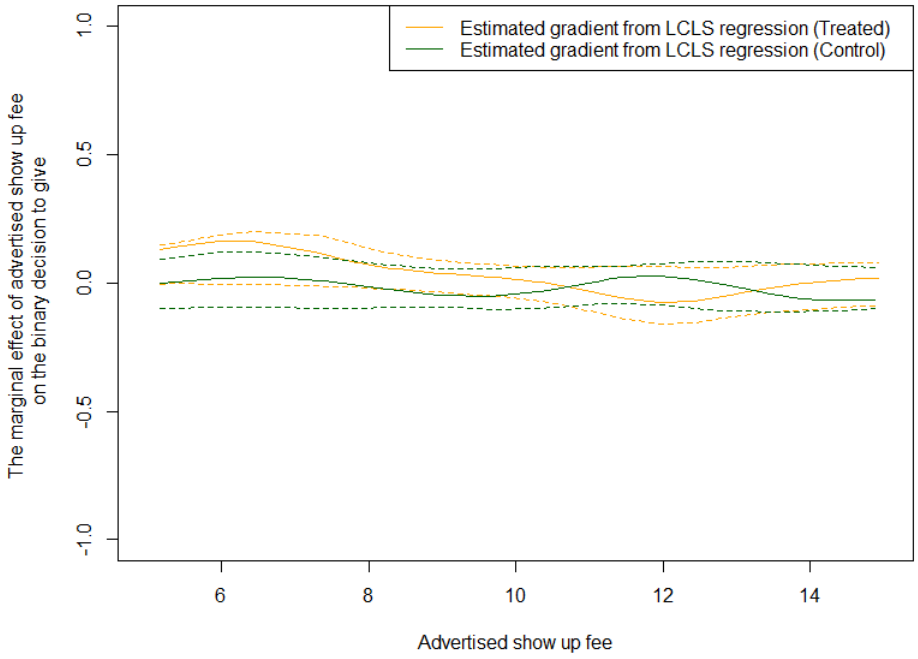


(b) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.40$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.40$)

Figure 3.21: Gradient plots from LCLS regressions (binary decision to give on advertised show-up fee) using LSCV bandwidths



(a) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.40$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)



(b) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.40$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.40$)

Table 3.15: Results on the binary decision to give from the linear regression model, separately for the treatment and control groups

	Dependent variable:			
	Donation amount (including zeros)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.025		−0.024
		(0.024)		(0.024)
Constant	0.636***	0.384	0.720***	0.976***
	(0.065)	(0.251)	(0.064)	(0.266)
Observations	55	55	50	50
Log likelihood	−38.793	−38.236	−31.910	−31.400
AIC	79.586	80.471	65.820	66.799
BIC	83.601	86.493	69.644	72.535
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.16: Results on the binary decision to give from the linear regression model using the whole sample

	Dependent variable:			
	Donation amount (including zeros)			
	(1)	(2)	(3)	(4)
Treatment		−0.084	−0.336	−0.593
		(0.092)	(0.252)	(0.368)
Treatment × advertised show-up fee			0.025	0.048
			(0.023)	(0.034)
Advertised show-up fee				−0.024
				(0.025)
Constant	0.676***	0.720***	0.720***	0.976***
	(0.046)	(0.067)	(0.066)	(0.275)
Observations	105	105	105	105
Log likelihood	−70.247	−69.827	−69.236	−68.760
AIC	142.495	143.655	144.472	145.519
BIC	147.803	151.617	155.088	158.789
Note:				*p<0.1; **p<0.05; ***p<0.01

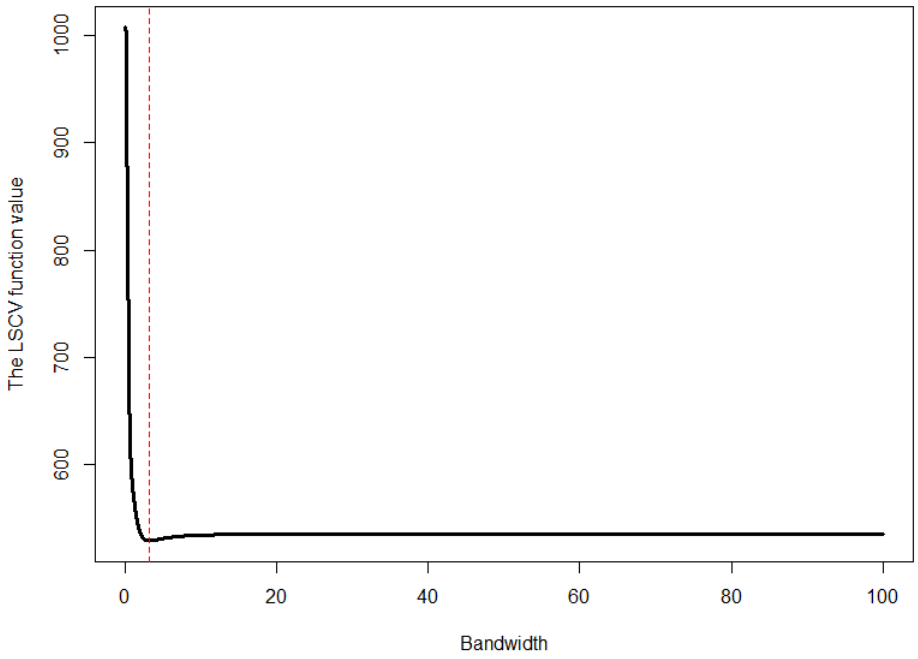
Table 3.17: Results on the binary decision to give from the Probit model, separately for the treatment and control groups

	Dependent variable:			
	Donation amount (including zeros)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.064		−0.071
		(0.063)		(0.072)
Constant	0.349***	−0.309	0.583***	1.367*
	(0.173)	(0.665)	(0.189)	(0.816)
Observations	55	55	50	50
Log likelihood	−36.051	−35.513	−29.648	−29.139
AIC	74.103	75.026	61.295	62.278
BIC	76.110	79.041	63.207	66.102
Note:				*p<0.1; **p<0.05; ***p<0.01

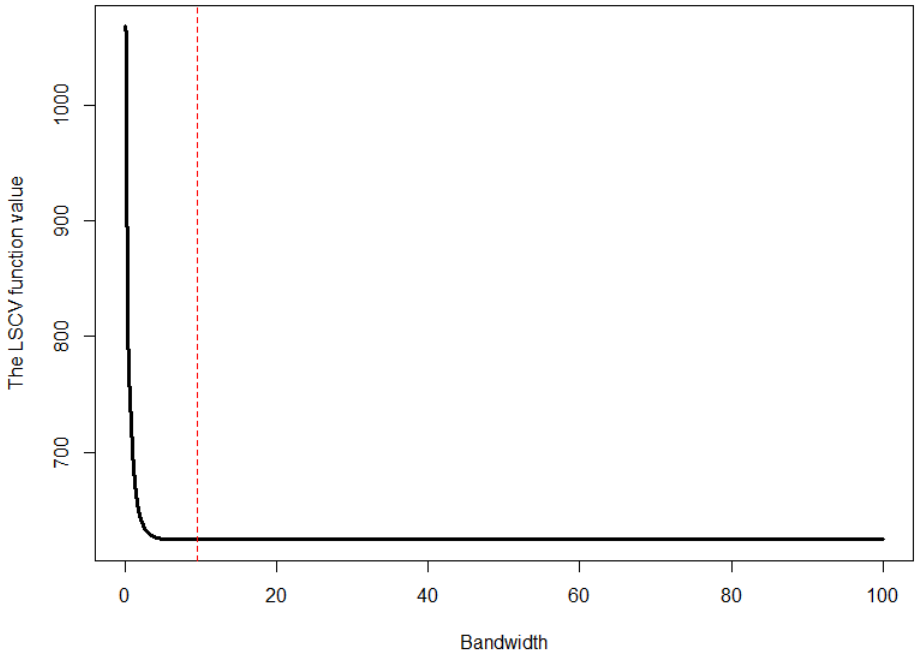
Table 3.18: Results on the binary decision to give from the Probit model using the whole sample

	Dependent variable:			
	Donation amount (including zeros)			
	(1)	(2)	(3)	(4)
Treatment		−0.234	−0.892	−1.676
		(0.256)	(0.692)	(1.053)
Treatment × advertised show-up fee			0.064	0.136
			(0.063)	(0.096)
Advertised show-up fee				−0.071
				(0.072)
Constant	0.457***	0.583***	0.583***	1.367*
	(0.127)	(0.189)	(0.189)	(0.816)
Observations	105	105	105	105
Log likelihood	−66.119	−65.699	−65.161	−64.652
AIC	134.239	135.398	136.321	137.304
BIC	136.893	140.706	144.283	147.919
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.22: LSCV bandwidth selection for the LCLS regression (donation amount conditional on giving on advertised show-up fee



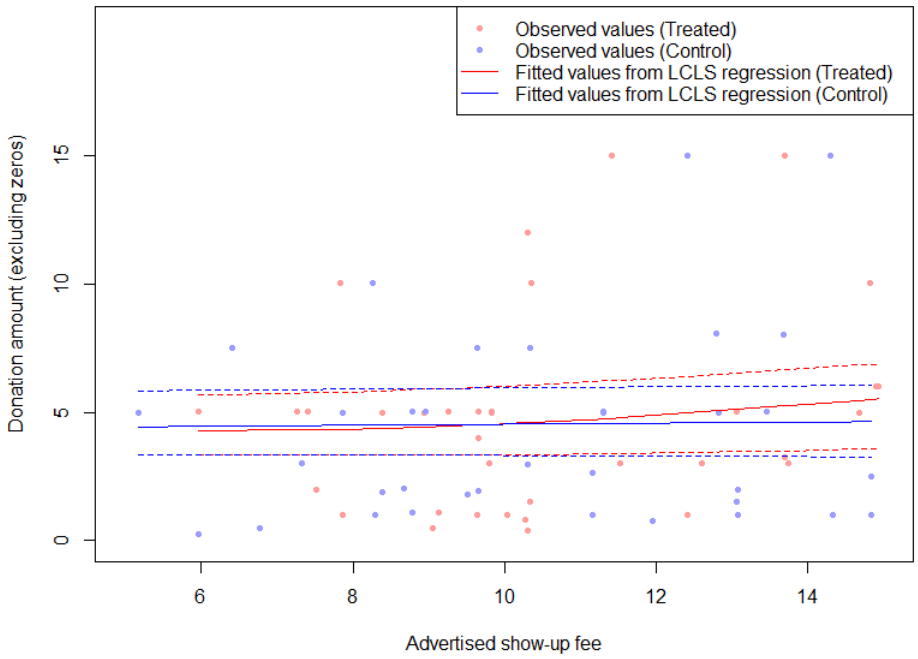
(a) LSCV bandwidth selection for the treatment group



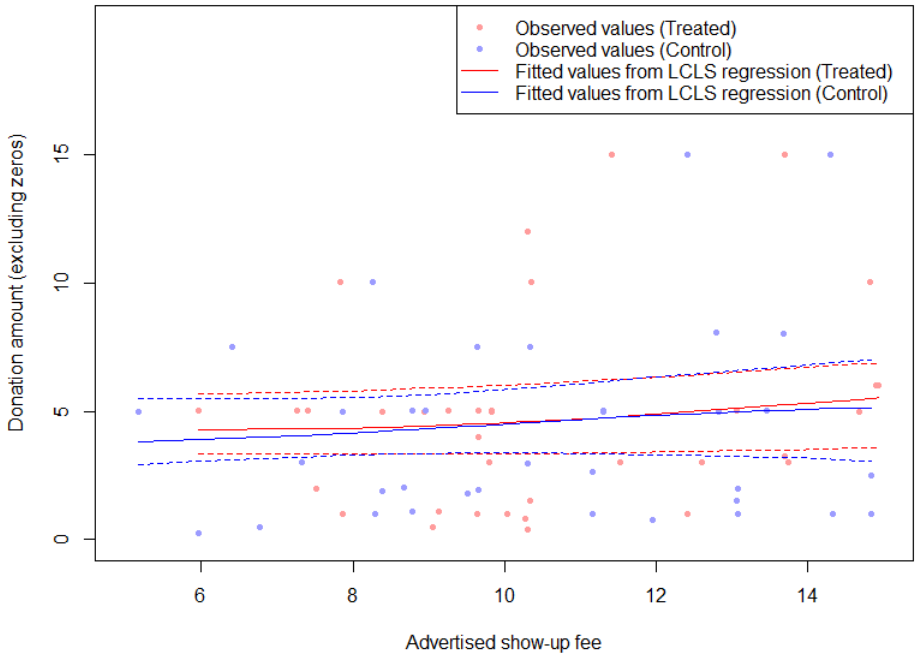
(b) LSCV bandwidth selection for the control group

Note: The LSCV bandwidth for the treatment group is 3.19 based on the LCLS regression for the treatment group, while the LSCV bandwidth for the control group is 9.61 (greater than 3×2.69 and 2.69 is the standard deviation of the advertised show up fee in the control group) based on the LCLS regression for the control group.

Figure 3.23: Scatter plots with the fitted values from LCLS regressions (donation amount conditional on giving on advertised show-up fee) using LSCV bandwidths

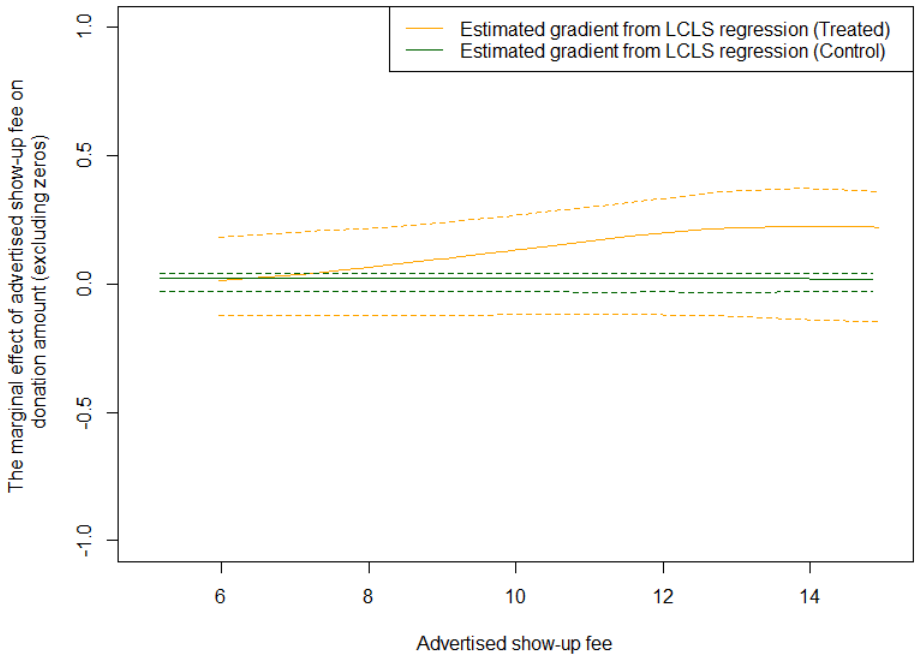


(a) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 3.19$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 9.61$)

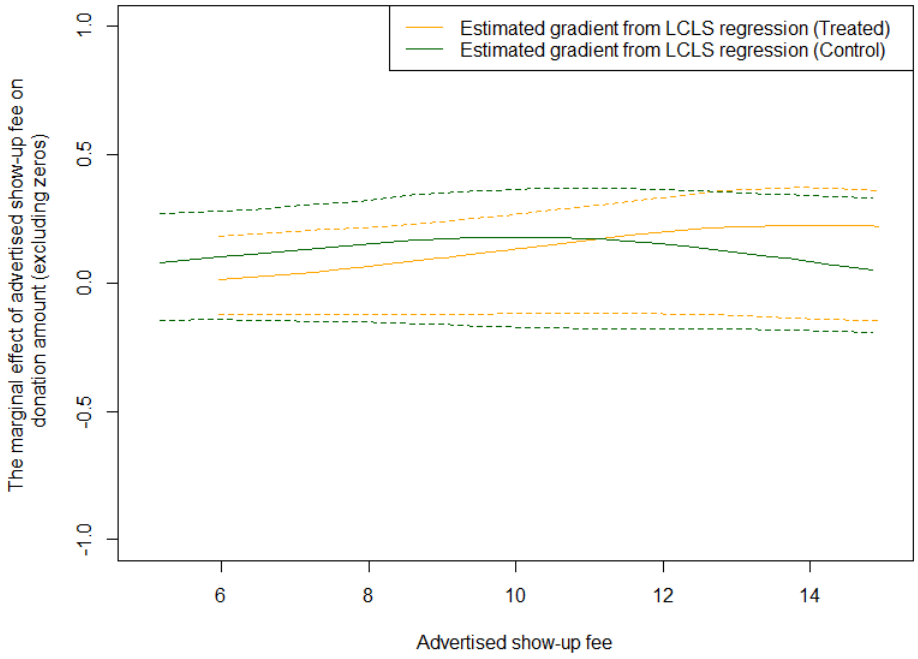


(b) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 3.19$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 3.19$)

Figure 3.24: Gradient plots from LCLS regressions (donation amount conditional on giving on advertised show-up fee) using LSCV bandwidths



(a) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 3.19$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 9.61$)



(b) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 3.19$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 3.19$)

Table 3.19: Results on the donation amount conditional on giving from the linear regression model, separately for the treatment and control groups

	Dependent variable:			
	Donation amount conditional on giving			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.354		0.269
		(0.268)		(0.258)
Constant	4.819***	1.070	4.545***	1.696
	(0.651)	(2.904)	(0.684)	(2.815)
Observations	35	35	36	36
Log likelihood	−97.371	−96.465	−102.431	− 101.864
AIC	196.741	196.930	206.862	207.727
BIC	199.852	201.596	210.029	212.478
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.20: Results on the donation amount conditional on giving from the linear regression model using the whole sample

	Dependent variable:			
	Donation amount conditional on giving			
	(1)	(2)	(3)	(4)
Treatment		0.274	−3.475	−0.626
		(0.946)	(3.093)	(4.062)
Treatment × advertised show-up fee			0.354	0.085
			(0.279)	(0.373)
Advertised show-up fee				0.269
				(0.249)
Constant	4.680***	4.545***	4.545***	1.696*
	(0.470)	(0.664)	(0.661)	(2.719)
Observations	71	71	71	71
Log likelihood	−198.917	−8.874	−198.039	−197.426
AIC	399.835	401.749	402.078	402.852
BIC	404.360	408.537	411.129	414.166
Note:				*p<0.1; **p<0.05; ***p<0.01

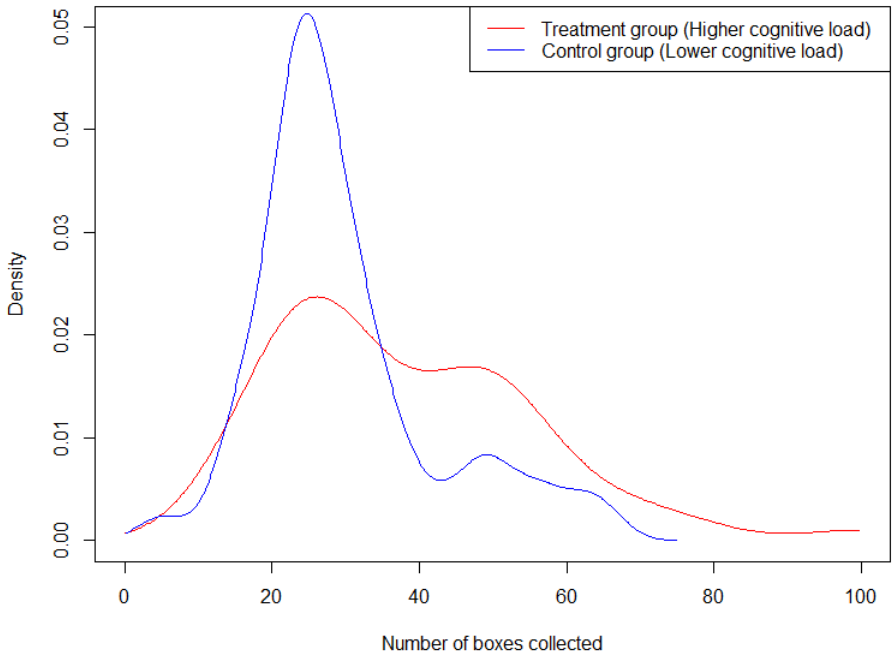
Table 3.21: Results on the donation amount conditional on giving from the linear regression model with a log-level function, separately for the treatment and control groups

	Dependent variable:			
	ln (Donation amount conditional on giving)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.081		0.058
		(0.066)		(0.064)
Constant	1.212***	0.350	1.085***	0.471
	(0.160)	(0.716)	(0.168)	(0.694)
Observations	35	35	36	36
Log likelihood	−48.269	−47.479	−51.881	−51.445
AIC	98.537	98.959	105.761	106.890
BIC	101.648	103.625	108.928	111.640
Note:				*p<0.1; **p<0.05; ***p<0.01

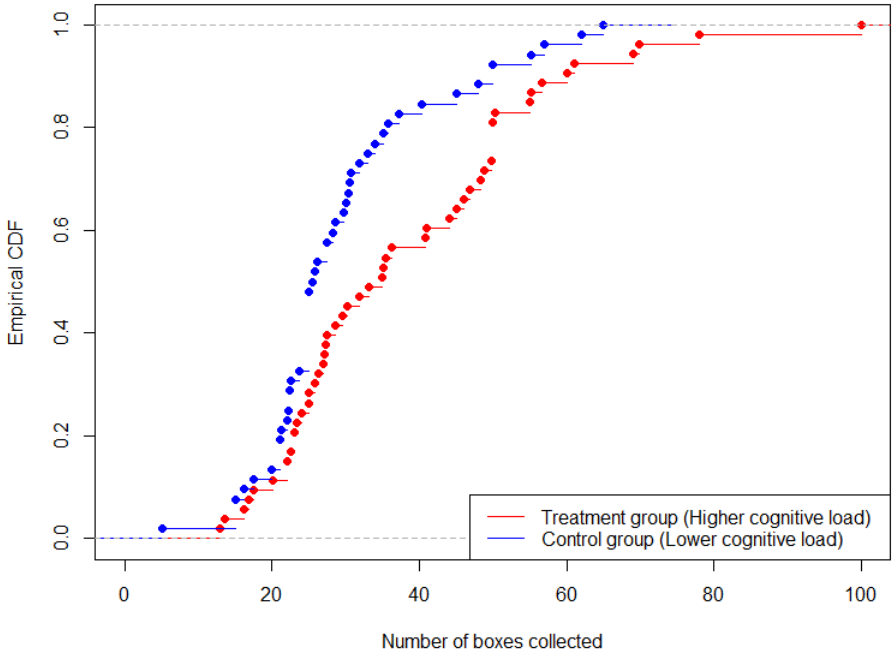
Table 3.22: Results on the donation amount conditional on giving from the linear regression model with a log-level function using the whole sample

	Dependent variable:			
	ln (Donation amount conditional on giving)			
	(1)	(2)	(3)	(4)
Treatment		0.127	−0.735	−0.121
		(0.232)	(0.761)	(1.002)
Treatment × advertised show-up fee			0.081	0.023
			(0.069)	(0.092)
Advertised show-up fee				0.058
				(0.061)
Constant	1.148***	1.085***	1.085***	0.471
	(0.116)	(0.163)	(0.163)	(0.670)
Observations	71	71	71	71
Log likelihood	−99.372	−99.219	−98.489	−98.019
AIC	200.744	202.439	202.978	04.039
BIC	205.269	209.227	12.029	215.352
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 3.25: Kernel density and empirical CDF plots of the p-beauty contest under cognitive load task outcome by treatment



(a) Kernel density



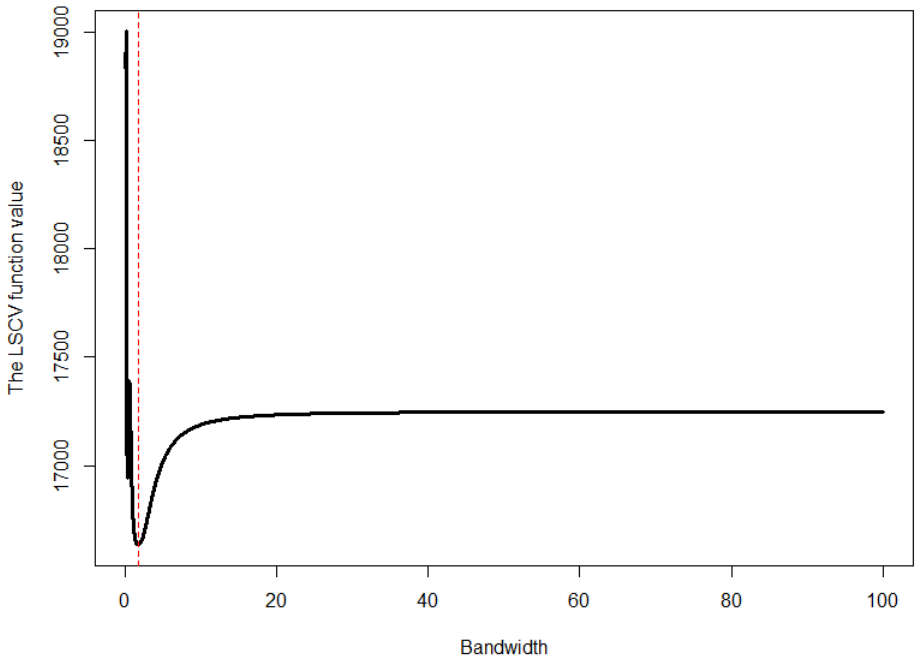
(b) Empirical CDF

Table 3.23: Results of the p-beauty contest outcome by treatment

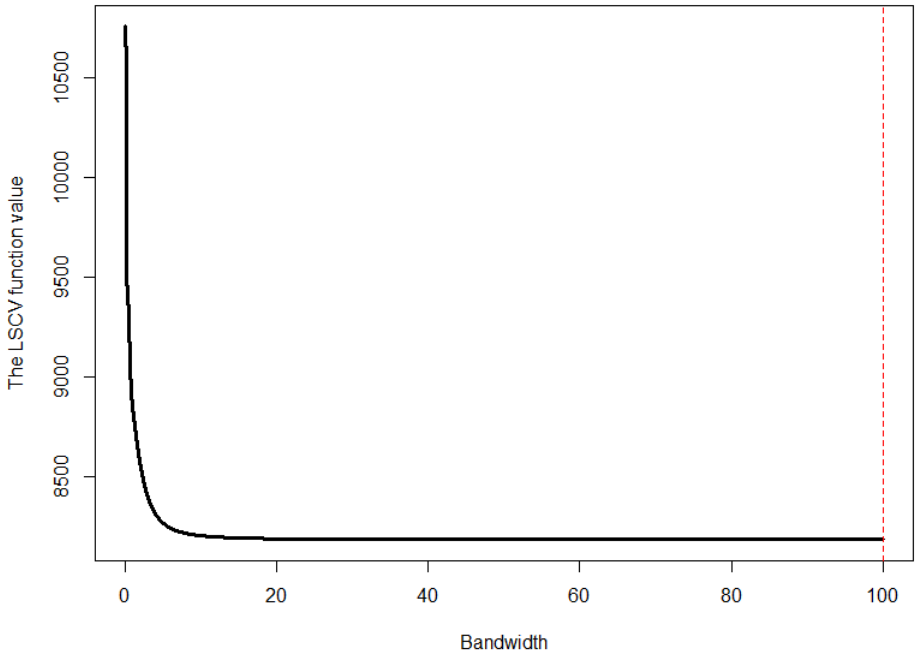
	N	Mean	T-test for equal means	Mann-Whitney test	Kolmogorov-Smirnov test
Treatment group (Higher cognitive load)	53	38.270 (17.869)	$p = 0.005$ (H_1: unequal mean)	$p = 0.014$ (H_1: unequal mean)	$p = 0.032$ (H_1: unequal distribution)
Control group (Lower cognitive load)	52	29.679 (12.423)	$p = 0.003$ (H_1: treatment group has a higher mean)	$p = 0.007$ (H_1: treatment group has a higher mean)	$p = 0.016$ (H_1: the distribution in the treatment group stochastically dominates that in the control group)

Note: Standard errors are shown in the parentheses.

Figure 3.26: LSCV bandwidth selection for the LCLS regression (guess in the p-beauty contest under cognitive load on advertised show-up fee



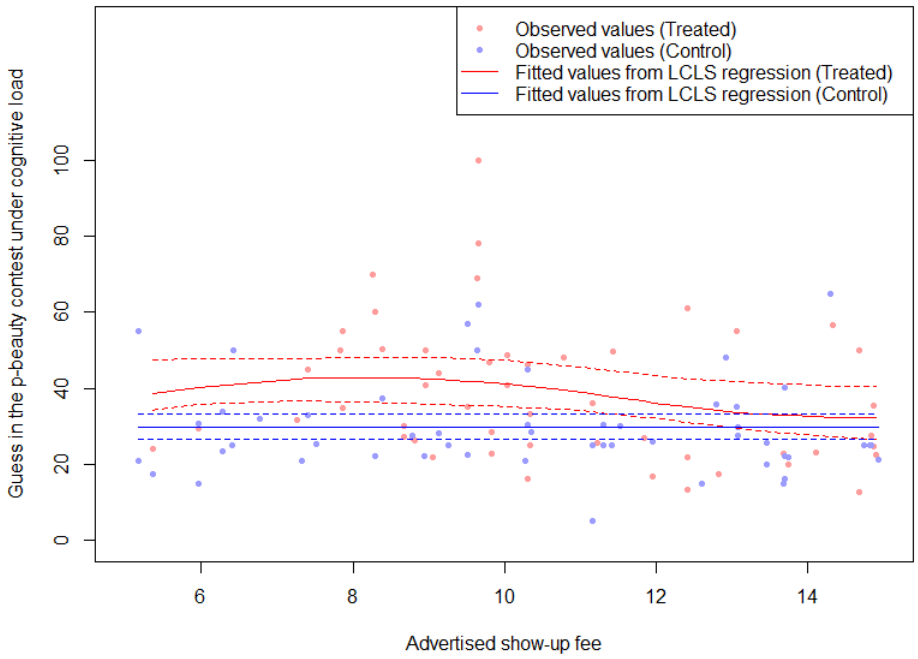
(a) LSCV bandwidth selection for the treatment group



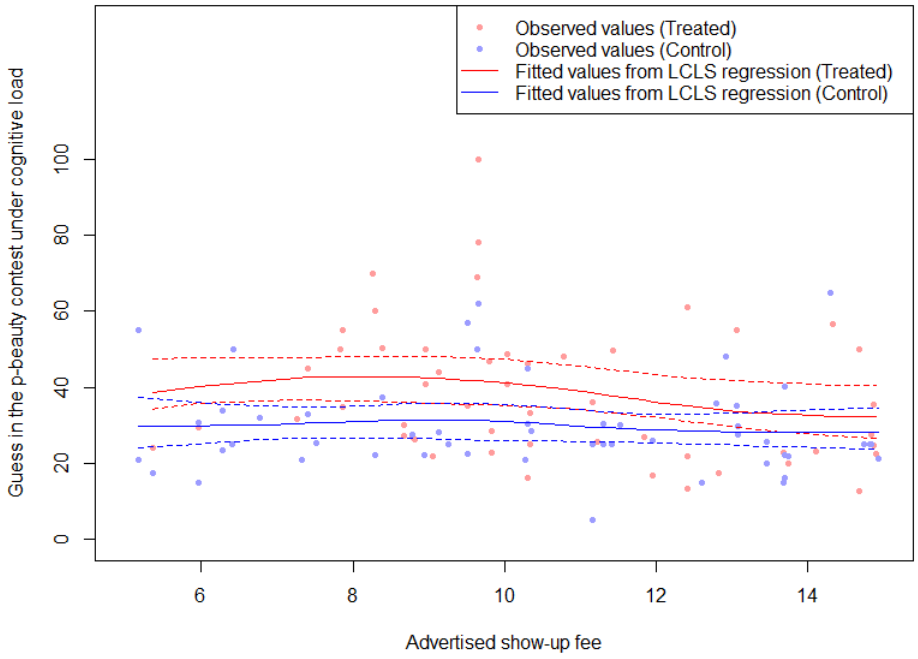
(b) LSCV bandwidth selection for the control group

Note: The LSCV bandwidth for the treatment group is 1.75 based on the LCLS regression for the treatment group, while the LSCV bandwidth for the control group is 100 based on the LCLS regression for the control group. The LSCV bandwidth for the control group might go infinity if we keep increasing the upper bound of the search range.

Figure 3.27: Scatter plots with the fitted values from LCLS regressions (guess in the p-beauty contest under cognitive load on advertised show-up fee) using LSCV bandwidths

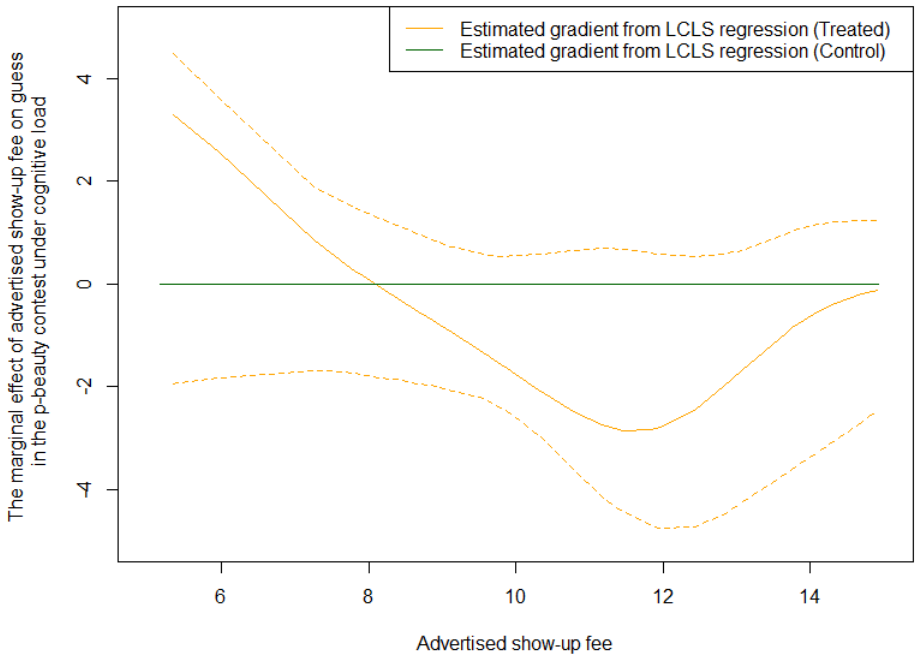


(a) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.75$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)

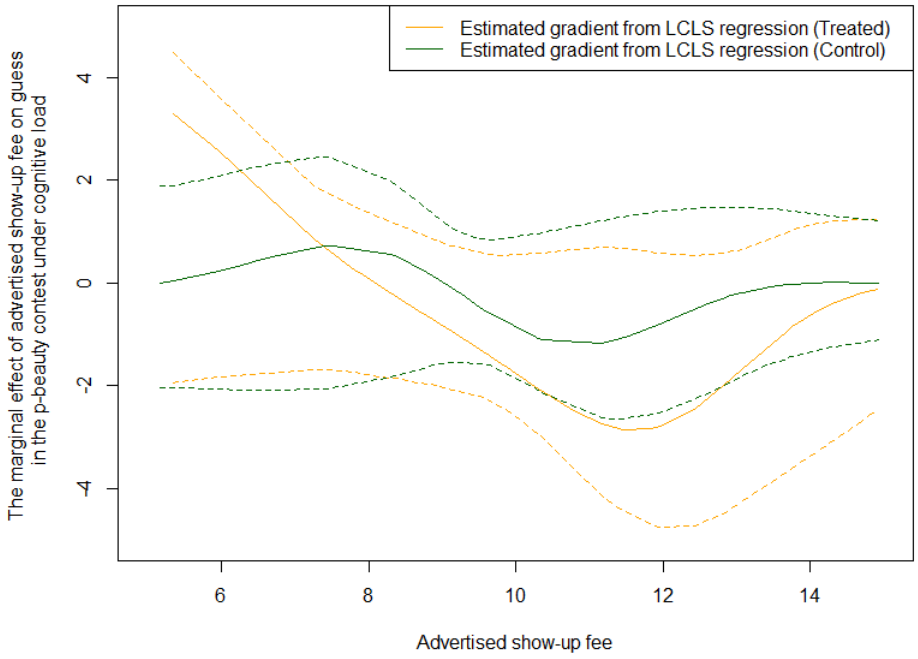


(b) Scatter plots with the fitted values where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.75$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.75$)

Figure 3.28: Gradient plots from LCLS regressions (guess in the p-beauty contest under cognitive load on advertised show-up fee) using LSCV bandwidths



(a) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.75$) and bandwidth for the control group = LSCV bandwidth for the control group ($h = 100$)



(b) Gradient plots from LCLS regressions where bandwidth for the treatment group = LSCV bandwidth for the treatment group ($h = 1.75$) and bandwidth for the control group = LSCV bandwidth for the treatment group ($h = 1.75$)

Table 3.24: Results on the guess in the p-beauty contest under cognitive load from the linear regression model, separately for the treatment and control groups

	Dependent variable:			
	Guess in the p-beauty contest			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		−1.493		−0.321
		(0.965)		(0.585)
Constant	38.270***	54.210***	29.679***	33.014***
	(2.455)	(10.587)	(1.723)	(6.325)
Observations	53	53	52	52
Log likelihood	−228.502	−227.288	−205.295	−205.139
AIC	459.005	458.575	412.591	414.2797
BIC	462.945	464.486	416.493	420.133
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.25: Results on the guess in the p-beauty contest under cognitive load from the linear regression model using the whole sample

	Dependent variable:			
	Guess in the p-beauty contest			
	(1)	(2)	(3)	(4)
Treatment		8.591***	24.531**	21.195*
		(3.009)	(9.398)	(12.019)
Treatment × advertised show-up fee			−1.493*	−1.172
			(0.835)	(1.102)
Advertised show-up fee				−0.321
				(0.716)
Constant	34.015***	29.679***	29.679***	33.014***
	(1.555)	(2.138)	(2.115)	(7.743)
Observations	105	105	105	105
Log likelihood	−440.188	−436.189	−434.568	−434.464
AIC	882.376	876.378	875.137	876.928
BIC	887.684	884.340	885.752	890.198
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.26: Results on the guess in the p-beauty contest under cognitive load from the linear regression model with a log-level function, separately for the treatment and control groups

	Dependent variable:			
	ln (Guess in the p-beauty contest)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		−0.047*		−0.012
		(0.025)		(0.020)
Constant	3.542***	4.045***	3.306***	3.430***
	(0.063)	(0.270)	(0.059)	(0.218)
Observations	53	53	52	52
Log likelihood	−34.750	−32.917	−30.154	−29.974
AIC	71.499	69.835	62.307	63.948
BIC	75.440	75.745	66.210	69.8023
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.27: Results on the guess in the p-beauty contest under cognitive load from the linear regression model with a log-level function using the whole sample

	Dependent variable:			
	ln (Guess in the p-beauty contest)			
	(1)	(2)	(3)	(4)
Treatment		0.235***	0.738**	0.615*
		(0.087)	(0.271)	(0.346)
Treatment × advertised show-up fee			−0.047*	−0.035
			(0.024)	(0.032)
Advertised show-up fee				−0.012
				(0.021)
Constant	3.425***	3.306***	3.306***	3.430***
	(1.555)	(2.138)	(2.115)	(7.743)
Observations	105	105	105	105
Log likelihood	−67.655	−64.055	−62.116	−61.945
AIC	137.310	132.111	130.232	131.889
BIC	142.618	140.072	140.848	145.159
Note:				*p<0.1; **p<0.05; ***p<0.01

REFERENCES

Ahn, H., Powell, J.L., 1993. Semiparametric estimation of censored selection models with a nonparametric selection mechanism. Journal of Econometrics 58, 3–29.

Al-Ubaydli, O., List, J.A., 2013. On the generalizability of experimental results in economics: With a response to Camerer. Technical Report. National Bureau of Economic Research.

Andreoni, J., 1989. Giving with impure altruism: Applications to charity and Ricardian equivalence. Journal of Political Economy 97, 1447–1458.

Andreoni, J., 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. Economic Journal 100, 464–477.

Behaghel, L., Blau, D.M., 2012. Framing social security reform: Behavioral responses to changes in the full retirement age. American Economic Journal: Economic Policy 4, 41–67.

Behaghel, L., Crépon, B., Gurgand, M., Le Barbanchon, T., 2015. Please call again: Correcting nonresponse bias in treatment effect models. Review of Economics and Statistics 97, 1070–1080.

Behrman, J.R., Parker, S.W., Todd, P.E., 2009. Schooling impacts of conditional cash transfers on young children: Evidence from mexico. Economic development and cultural change 57, 439–477.

Blanco, G., Flores, C.A., Flores-Lagunes, A., 2013. Bounds on average and quantile treatment effects of job corps training on wages. Journal of Human Resources 48, 659–701.

Blumenthal, M., Kalambokidis, L., Turk, A., 2012. Subsidizing charitable contributions with a match instead of a deduction: What happens to donations and compliance? National Tax Journal 65, 91.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: methods and applications. Cambridge university press.

Carpenter, J., Graham, M., Wolf, J., 2013. Cognitive ability and strategic sophistication. Games and Economic Behavior 80, 115–130.

Crosetto, P., Filippin, A., 2013. The "bomb" risk elicitation task. Journal of Risk and Uncertainty 47, 31–65.

Czibor, E., Jimenez-Gomez, D., List, J.A., 2019. The dozen things experimental economists should do (more of). Southern Economic Journal 86, 371–432.

Das, M., Newey, W.K., Vella, F., 2003. Nonparametric estimation of sample selection models. The Review of Economic Studies 70, 33–58.

De Mel, S., McKenzie, D., Woodruff, C., 2008. Returns to capital in microenterprises: evidence from a field experiment. The quarterly journal of Economics 123, 1329–1372.

Deck, C., Jahedi, S., 2015. The effect of cognitive load on economic decision making: A survey and new experiments. European Economic Review 78, 97–119.

Eckel, C.C., Grossman, P.J., 1996. Altruism in anonymous dictator games. Games and economic behavior 16, 181–191.

Eckel, C.C., Grossman, P.J., 2003. Rebate versus matching: does how we subsidize charitable contributions matter? Journal of Public Economics 87, 681–701.

Eckel, C.C., Grossman, P.J., 2006. Subsidizing charitable giving with rebates or matching: Further laboratory evidence. Southern Economic Journal , 794–807.

Eckel, C.C., Grossman, P.J., 2008. Subsidizing charitable contributions: a natural field experiment comparing matching and rebate subsidies. Experimental Economics 11, 234–252.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K., Group, O.H.S., 2012. The oregon health insurance experiment: evidence from the first year. The Quarterly journal of economics 127, 1057–1106.

Foundation, G.U., 2020. Giving USA 2020: The annual report on philanthropy for the year 2019.

Frey, B.S., Meier, S., 2004. Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. American Economic Review 94, 1717–1722.

Gee, L.K., Schreck, M.J., 2018. Do beliefs about peers matter for donation matching? experiments in the field and laboratory. Games and Economic Behavior 107, 282–297.

Glewwe, P., Ilias, N., Kremer, M., 2010. Teacher incentives. American Economic Journal: Applied Economics 2, 205–27.

Harrison, G.W., Lau, M.I., Rutström, E.E., 2009. Risk attitudes, randomization to treatment, and self-selection into experiments. Journal of Economic Behavior & Organization 70, 498–507.

Harrison, G.W., Lau, M.I., Yoo, H.I., 2020. Risk attitudes, sample selection, and attrition in a longitudinal field experiment. Review of Economics and Statistics 102, 552–568.

Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, in: Annals of economic and social measurement, volume 5, number 4. NBER, pp. 475–492.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica: Journal of the econometric society , 153–161.

Henderson, D.J., Parmeter, C.F., 2015. Applied nonparametric econometrics. Cambridge University Press.

Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. American economic review 92, 1644–1655.

Horowitz, J.L., Manski, C.F., 1995. Identification and robustness with contaminated and corrupted data. Econometrica: Journal of the Econometric Society , 281–302.

Horowitz, J.L., Manski, C.F., 1998. Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. Journal of Econometrics 84, 37–58.

Horowitz, J.L., Manski, C.F., 2000. Nonparametric analysis of randomized experiments with missing covariate and outcome data. Journal of the American statistical Association 95, 77–84.

Huck, S., Rasul, I., 2011. Matched fundraising: Evidence from a natural field experiment. Journal of Public Economics 95, 351–362.

Huck, S., Rasul, I., Shephard, A., 2015. Comparing charitable fundraising schemes: Evidence from a natural field experiment and a structural model. American Economic Journal: Economic Policy 7, 326–69.

Karlan, D., List, J.A., 2007. Does price matter in charitable giving? evidence from a large-scale natural field experiment. American Economic Review 97, 1774–1793.

Karlan, D., List, J.A., Shafir, E., 2011. Small matches and charitable giving: Evidence from a natural field experiment. Journal of Public Economics 95, 344–350.

Lee, D.S., 2002. Trimming for bounds on treatment effects with missing outcomes.

Lee, D.S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. The Review of Economic Studies 76, 1071–1102.

List, J.A., Lucking-Reiley, D., 2002. The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. Journal of Political Economy 110, 215–233.

Nagel, R., 1995. Unraveling in guessing games: An experimental study. The American Economic Review 85, 1313–1326.

Newey, W.K., 2009. Two-step series estimation of sample selection models. The Econometrics Journal 12, S217–S229.

Newey, W.K., McFadden, D.L., 1994. Large sample estimation and hypothesis testing, in: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics. New Holland. volume 4, pp. 2113–2245.

Rondeau, D., List, J.A., 2008. Matching and challenge gifts to charity: evidence from laboratory and natural field experiments. Experimental Economics 11, 253–267.

Schulz, J., Sunde, U., Thiemann, P., Thöni, C., 2019. Selection into experiments: Evidence from a population of students .

Slonim, R., Wang, C., Garbarino, E., Merrett, D., 2013. Opting-in: Participation bias in economic experiments. Journal of Economic Behavior & Organization 90, 43–70.

Zhang, J.L., Rubin, D.B., 2003. Estimation of causal effects via principal stratification when some outcomes are truncated by "death". Journal of Educational and Behavioral Statistics 28, 353–368.

APPENDIX A

A.1 Proof of Proposition 1 in Chapter 2

Proof: It's sufficient to prove consistency for the estimator of the lower bound $\widehat{\Delta}_1^{lb} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \min\left\{1, \frac{\widehat{\gamma}}{1-\widehat{p}}\right\}$ since a similar argument will follow for the estimator of the upper bound ($\widehat{\Delta}_1^{ub} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i} - \max\left\{0, \frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}\right\}$).

The first term of $\widehat{\Delta}_1^{lb}$ is an OLS estimator of $E[Y_i|R_i = 1, T_i = 1]$ which is a consistent estimator of $E[Y_i|R_i = 1, T_i = 1]$ with
$\sqrt{n}\Big(\frac{\sum Y_i R_i T_i}{\sum R_i T_i} - E[Y_i|R_i = 1, T_i = 1]\Big) \xrightarrow{d} N\Big(0, \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_i T_i]}\Big)$.

For the second term of $\widehat{\Delta}_1^{lb}$, we assume the econometrician know which minimum of the estimator $\widehat{\overline{y}}_{CR} = \min\left\{1, \frac{\widehat{\gamma}}{1-\widehat{p}}\right\}$ is attained.

(1) When $\widehat{\overline{y}}_{CR} = 1$, $\widehat{\Delta}_1^{lb} = \frac{\sum Y_i R_i T_i}{\sum R_i T_i}$ and $\Delta_1^{lb} = E[Y_i|R_i = 1, T_i = 1]$. We have $\widehat{\Delta}_1^{lb} \xrightarrow{p} \Delta_1^{lb}$ and $\sqrt{n}\Big(\widehat{\Delta}_1^{lb} - \Delta_1^{lb}\Big) \xrightarrow{d} N\Big(0, \frac{Var[Y_i|R_i=1,T_i=1]}{E[R_i T_i]}\Big)$.

(2) When $\widehat{\overline{y}}_{CR} = \frac{\widehat{\gamma}}{1-\widehat{p}}$, we know:

(i) $\widehat{\gamma} = \frac{\sum Y_i R_i (1-T_i)}{\sum R_i (1-T_i)}$ is an OLS estimator of $\gamma = E[Y_i|R_i = 1, T_i = 0]$. Thus $\widehat{\gamma}$ is a consistent estimator of γ, i.e., $\widehat{\gamma} \xrightarrow{p} \gamma$ with $\sqrt{n}\Big(\widehat{\gamma} - \gamma\Big)$
$\xrightarrow{d} N\Big(0, \frac{Var[Y_i|R_i=1,T_i=0]}{E[R_i(1-T_i)]}\Big)$.

(ii) $\widehat{p} = 1 - \frac{\widehat{p}_a}{\widehat{p}_b}$ which implies $\frac{1}{1-\widehat{p}} = \frac{\widehat{p}_b}{\widehat{p}_a}$ and $\frac{\widehat{\gamma}}{1-\widehat{p}} = \frac{\widehat{\gamma}\widehat{p}_b}{\widehat{p}_a}$, where $\widehat{p}_a = \frac{\sum R_i T_i}{\sum T_i}$ and $\widehat{p}_b = \frac{\sum R_i(1-T_i)}{\sum(1-T_i)}$. The WLLN shows that $\widehat{p}_a \xrightarrow{p} p_a$ and $\widehat{p}_b \xrightarrow{p} p_b$, where $p_a = Pr(R_i = 1|T_i = 1)$ and $p_b = Pr(R_i = 1|T_i = 0)$.

(iii) Given $p_a > 0$, the Continuous Mapping Theorem (CMT) allows us to show that $\frac{\widehat{\gamma}\widehat{p}_b}{\widehat{p}_a} \xrightarrow{p} \frac{\gamma p_b}{p_a}$, or equivalently, $\frac{\widehat{\gamma}}{1-\widehat{p}} \xrightarrow{p} \frac{\gamma}{1-p}$. Therefore we have $\widehat{\Delta}_1^{lb} \xrightarrow{p} \Delta_1^{lb}$ when $\widehat{\overline{y}}_{CR} = \frac{\widehat{\gamma}}{1-\widehat{p}}$.

To prove the asymptotic normality of $\widehat{\Delta}_1^{lb}$ and $\widehat{\Delta}_1^{ub}$, it's sufficient to show the asymptotic normality of $\frac{\widehat{\gamma}}{1-\widehat{p}}$ and $\frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}$.

We follow by showing the asymptotic normality of $\frac{\widehat{\gamma}}{1-\widehat{p}}$. First, we know $\sqrt{n}\Big(\widehat{\gamma}-\gamma\Big) \xrightarrow{d} N\Big(0, V^{\gamma}\Big)$,where $V^{\gamma} = \frac{Var[Y_i|R_i=1,T_i=0]}{E[R_i(1-T_i)]}$ and we have $\sqrt{n}\Big(\widehat{p}_a - p_a\Big) \xrightarrow{d} N\Big(0, \frac{p_a(1-p_a)}{E[T_i]}\Big)$ and $\sqrt{n}\Big(\widehat{p}_b - p_b\Big) \xrightarrow{d} N\Big(0, \frac{p_b(1-p_b)}{E[1-T_i]}\Big)$ by CLT.

Note that $\frac{\gamma}{1-p} = g(\gamma, p)$ is continuously differentiable in a neighborhood of (γ, p) with $G = G(\gamma, p) = (\frac{\partial g}{\partial \gamma}, \frac{\partial g}{\partial p})' = (\frac{1}{1-p}, \frac{\gamma}{(1-p)^2})$ and $p = 1 - \frac{p_a}{p_b} = f(p_a, p_b)$ is continuously differentiable in a neighborhood of (p_a, p_b) with $F = F(p_a, p_b) = (\frac{\partial f}{\partial p_a}, \frac{\partial f}{\partial p_b})' = (-\frac{1}{p_b}, \frac{p_a}{p_b^2})'$.

Applying the delta method, we have:

$$\sqrt{n}\Big(\frac{\widehat{\gamma}}{1-\widehat{p}} - \frac{\gamma}{1-p}\Big) \xrightarrow{d} N\Big(0, V^{\overline{y}_{CR}}\Big),$$

where

$$V^{\overline{y}_{CR}} = G' \begin{pmatrix} V^{\gamma} & 0 \\ 0 & V^{p} \end{pmatrix} G = \frac{V^{\gamma}}{(1-p)^2} + \frac{\gamma^2}{(1-p)^4} V^{p}$$

and

$$\begin{aligned} V^{p} &= F' \begin{pmatrix} V^{p_a} & 0 \\ 0 & V^{p_b} \end{pmatrix} F = \frac{p_a(1-p_a)}{p_b^2 E[T_i]} + \frac{p_a^2 p_b(1-p_b)}{p_b^4 E[1-T_i]} \\ &= \frac{p_a^2}{p_b^2}\Big[\frac{1/p_a - 1}{E[T_i]} + \frac{1/p_b - 1}{E[1-T_i]}\Big] = (1-p)^2\Big[\frac{1/p_a - 1}{E[T_i]} + \frac{1/p_b - 1}{E[1-T_i]}\Big] \\ &= (1-p)^2\Big[\frac{1/Pr(R_i=1|T_i=1) - 1}{E[T_i]} + \frac{1/Pr(R_i=1|T_i=0) - 1}{E[1-T_i]}\Big]. \end{aligned}$$

We finish the proof by showing the asymptotic property of $\frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}$.

Note that $\frac{\gamma-p}{1-p} = h(\gamma, p)$ is continuously differentiable in a neighborhood of (γ, p) with $H = H(\gamma, p) = (\frac{\partial h}{\partial \gamma}, \frac{\partial h}{\partial p})' = (\frac{1}{1-p}, \frac{\gamma-1}{(1-p)^2})$ and $p = 1 - \frac{p_a}{p_b} = f(p_a, p_b)$ is continuously differentiable in a neighborhood of (p_a, p_b) with $F = F(p_a, p_b) = (\frac{\partial f}{\partial p_a}, \frac{\partial f}{\partial p_b})' = (-\frac{1}{p_b}, \frac{p_a}{p_b^2})'$.

Applying the delta method, we have:

$$\sqrt{n}\Big(\frac{\widehat{\gamma}-\widehat{p}}{1-\widehat{p}}-\frac{\gamma-p}{1-p}\Big)\xrightarrow{d}N\Big(0,V^{\underline{y}_{CR}}\Big),$$

where

$$V^{\underline{y}_{CR}}=H'\begin{pmatrix}V^{\gamma} & 0\\ 0 & V^{p}\end{pmatrix}H=\frac{V^{\gamma}}{(1-p)^2}+\frac{(1-\gamma)^2}{(1-p)^4}V^{p}$$

and

$$\begin{aligned}V^{p}&=F'\begin{pmatrix}V^{p_a} & 0\\ 0 & V^{p_b}\end{pmatrix}F\\ &=(1-p)^2\Big[\frac{1/Pr(R_i=1|T_i=1)-1}{E[T_i]}+\frac{1/Pr(R_i=1|T_i=0)-1}{E[1-T_i]}\Big].\end{aligned}$$

Q.E.D.

A.2 Addressing missing data in Chapter 2

In the second part of Appendix A, we address missing data in Chapter 2. Note that the wave 1 e-mail opening information is missing for both groups. For the 1:1 unlimited group, the wave 5 e-mail receiving and opening information is missing. In order to use the bounding estimation approaches, we assume the proportion of individuals that open e-mails at wave 5 are the same between groups, conditional on whether they opened any e-mail or opened only one/two/three e-mails between waves 2 - 4. In Tables A.1 and A.2, we extrapolate the wave 5 e-mail open rates for the cold list and warm list potential donors, based on the wave 2 - 4 e-mail opening information in the 1:1 unlimited group and the wave 2 - 5 e-mail opening information in the 2:1 unlimited group.

In order to apply the bounding estimation approach in Behaghel et al. (2015), we also need to identify how many potential donors donated and how much they donated in each wave. However, as shown in Table A.3, the wave 4 and wave 5 donation information was (unfortunately) merged together. Tables A.4 and A.5 show the frequencies of donation dates by waves for the 2:1 unlimited group and the 1:1 unlimited group before splitting the wave 4 and wave 5 donation information. Notice that the five waves of e-mails are sent subsequently on 6/26/2013, 6/28/2013, 7/1/2013, 7/3/2013 and 7/5/2013. For each wave, we know that the donation dates should not be earlier than these e-mail sending dates.

To separate the wave 4 and wave 5 donation information, we assume the portion of individuals who donated on or after 7/5/2013 at wave 5 among those who donated on or after 7/5/2013 during wave 4 and wave 5 should be approximately the same for both groups. For the 2:1 unlimited group, among those 49 individuals who donated on or after 7/5/2019 during wave 4 and wave 5, 43 people (87.8%) donated on or after 7/5/2019 at wave 5. For the 1:1 unlimited group, only 7 individuals donated on or after 7/5/2019 during wave 4 and wave 5. Note that all were warm list donors. Based on the aforementioned assumption, there should be $7 \times 87.8\% \approx 6$ individuals who donated on or after 7/5/2013 at wave 5 for the 1:1 unlimited group.

There are 7 possible ways to split the donation information during wave 4 and wave 5 for those who donated on or after 7/5/2019. Table A.6 shows the first 5 cases where one of the five individuals who donated on 7/5/2013 during wave 4 and wave 5 are assigned to wave 5. Table A.7 shows the case where the ones who donated on 7/7/2013 are assigned to wave 5. Table A.8 shows the case where the ones who donated on 7/11/2013 are assigned to wave 5.

For each case, we run 10 simulations on the Lee (2009) and Behaghel et al. (2015) bounds by randomly assigning the wave 5 e-mail opening information (open or not) for the potential donors in the 1:1 unlimited group based on the extrapolated wave 5 e-mail open rates. Both simulations on estimated average treatment effect on the probability to give and the donation amount are run separately for the warm list and cold list potential donors in order to check the robustness of the results. The results of the simulations are not significantly different between each case and are available by request.

Table A.1: Extrapolated wave 2 - 5 e-mail opening information in the 1:1 unlimited group (Cold list)

	Consider waves 2 - 4 only		Consider waves 2 - 5	
	Type	Open rate	Type	Open rate
(1)	Never open any e-mail	84.67%	Open wave 5 only	2.16%
(2)	Open wave 2 only	3.21%	Open waves 2 & 5	0.93%
(3)	Open wave 3 only	1.56%	Open waves 3 & 5	0.22%
(4)	Open wave 4 only	3.43%	Open waves 4 & 5	0.93%
(5)	Open waves 2 & 3	0.94%	Open waves 2, 3 & 5	0.43%
(6)	Open waves 2 & 4	1.78%	Open waves 2, 4 & 5	0.90%
(7)	Open waves 3 & 4	1.05%	Open waves 3, 4 & 5	0.49%
(8)	Open waves 2, 3 & 4	3.37%	Open waves 2, 3, 4 & 5	2.48%
	Sum	100.00%	Sum	8.53%
			Never open any e-mail	82.51%

(a) 2:1 unlimited group

	Consider waves 2 - 4 only		Consider waves 2 - 5	
	Type	Open rate	Type	Open rate
(1)	Never open any e-mail	84.65%	Open wave 5 only	2.16%
(2)	Open wave 2 only	2.57%	Open waves 2 & 5	0.75%
(3)	Open wave 3 only	1.69%	Open waves 3 & 5	0.24%
(4)	Open wave 4 only	3.05%	Open waves 4 & 5	0.83%
(5)	Open waves 2 & 3	1.10%	Open waves 2, 3 & 5	0.50%
(6)	Open waves 2 & 4	2.05%	Open waves 2, 4 & 5	1.03%
(7)	Open waves 3 & 4	1.22%	Open waves 3, 4 & 5	0.57%
(8)	Open waves 2, 3 & 4	3.67%	Open waves 2, 3, 4 & 5	2.69%
	Sum	100.00%	Sum	8.77%
			Never open any e-mail	82.49%

(b) 1:1 unlimited group

Note: Based on the wave 2 - 4 e-mail opening information in the 1:1 unlimited group and the wave 2 - 5 e-mail opening information in the 2:1 unlimited group, we extrapolate the wave 5 e-mail open rates for the cold list potential donors. We assume the proportion of individuals that open e-mails at wave 5 are the same between groups, conditional on whether they opened any e-mail or opened only one/two/three e-mails between waves 2 - 4.

Table A.2: Extrapolated wave 2 - 5 e-mail opening information in the 1:1 unlimited group (Warm list)

	Consider waves 2 - 4 only		Consider waves 2 - 5	
	Type	Open rates	Type	Open rates
(1)	Never open any e-mail	75.62%	Open wave 5 only	3.19%
(2)	Open wave 2 only	4.62%	Open waves 2 & 5	1.36%
(3)	Open wave 3 only	1.90%	Open waves 3 & 5	0.42%
(4)	Open wave 4 only	5.75%	Open waves 4 & 5	1.47%
(5)	Open waves 2 & 3	1.18%	Open waves 2, 3 & 5	0.49%
(6)	Open waves 2 & 4	3.32%	Open waves 2, 4 & 5	1.52%
(7)	Open waves 3 & 4	1.78%	Open waves 3, 4 & 5	0.91%
(8)	Open waves 2, 3 & 4	5.82%	Open waves 2, 3, 4 & 5	4.10%
	Sum	100.00%	Sum	13.46%
			Never open any e-mail	72.43%

(a) 2:1 unlimited group

	Consider waves 2 - 4 only		Consider waves 2 - 5	
	Type	Open rates	Type	Open rates
(1)	Never open any e-mail	75.26%	Open wave 5 only	3.17%
(2)	Open wave 2 only	4.12%	Open waves 2 & 5	1.21%
(3)	Open wave 3 only	1.92%	Open waves 3 & 5	0.42%
(4)	Open wave 4 only	4.77%	Open waves 4 & 5	1.22%
(5)	Open waves 2 & 3	1.41%	Open waves 2, 3 & 5	0.58%
(6)	Open waves 2 & 4	4.42%	Open waves 2, 4 & 5	2.02%
(7)	Open waves 3 & 4	1.62%	Open waves 3, 4 & 5	0.83%
(8)	Open waves 2, 3 & 4	6.47%	Open waves 2, 3, 4 &5	4.55%
	Sum	100.00%	Sum	14.02%
			Never open any e-mail	72.09%

(b) 1:1 unlimited group

Note: Based on the wave 2 - 4 e-mail opening information in the 1:1 unlimited group and the wave 2 - 5 e-mail opening information in the 2:1 unlimited group, we extrapolate the wave 5 e-mail open rates for the warm list potential donors. We assume the proportion of individuals that open e-mails at wave 5 are the same between groups, conditional on whether they opened any e-mail or opened only one/two/three e-mails between waves 2 - 4.

Table A.3: Number of donors in each wave of e-mail

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
(a) 2:1 unlimited					
Number of donors	6	29	27	43	43
Donation rate	0.0134%	0.0649%	0.0605%	0.0963%	0.0963%
Dollars raised - out of pocket	220	680	810	1,069	1,177
(b) 1:1 unlimited					
Number of donors	3	20	15	56	
Donation rate	0.0069%	0.0463%	0.0347%	0.1296%	
Dollars raised - out of pocket	55	605	448	1,492	

Note: For the 1:1 unlimited group, since the wave 5 e-mail receiving and opening information is missing, the wave 4 and wave 5 donation information are somehow merged together and need to be split in the Behaghel et al. (2015) bounding estimation approach.

Table A.4: Frequency table of the donation dates by waves (2:1 unlimited)

Wave 1	Date	6/26/2013	6/27/2013	7/5/2013		
	Frequency	3	1	2		
Wave 2	Date	6/28/2013	6/29/2013	6/30/2013	7/4/2013	
	Frequency	20	4	4	1	
Wave 3	Date	7/1/2013	7/2/2013	7/3/2013	7/4/2013	7/5/2013
	Frequency	14	7	3	2	1
Wave 4	Date	7/3/2013	7/4/2013	7/5/2013		
	Frequency	31	6	6		
Wave 5	Date	7/5/2013	7/6/2013			
	Frequency	40	3			

Table A.5: Frequency table of the donation dates by waves (1:1 unlimited, before splitting wave 4 and wave 5 donation information)

Wave 1	Date	6/26/2013				
	Frequency	3				
Wave 2	Date	6/28/2013	6/29/2013	6/30/2013	7/1/2013	
	Frequency	10	5	3	2	
Wave 3	Date	7/1/2013	7/2/2013	7/4/2013	7/5/2013	
	Frequency	9	3	1	2	
Waves 4 & 5	Date	7/3/2013	7/4/2013	7/5/2013	7/7/2013	7/11/2013
	Frequency	43	6	5	1	1

Table A.6: Frequency table of the donation dates by waves (1:1 unlimited, after splitting donation information, case 1 - 5)

Wave 1	Date	6/26/2013			
	Frequency	3			
Wave 2	Date	6/28/2013	6/29/2013	6/30/2013	7/1/2013
	Frequency	10	5	3	2
Wave 3	Date	7/1/2013	7/2/2013	7/4/2013	7/5/2013
	Frequency	9	3	1	2
Wave 4	Date	7/3/2013	7/4/2013	7/5/2013	
	Frequency	43	6	1	
Wave 5	Date	7/5/2013	7/7/2013	7/11/2013	
	Frequency	4	1	1	

Table A.7: Frequency table of the donation dates by waves (1:1 unlimited, after splitting donation information, case 6)

Wave 1	Date	6/26/2013			
	Frequency	3			
Wave 2	Date	6/28/2013	6/29/2013	6/30/2013	7/1/2013
	Frequency	10	5	3	2
Wave 3	Date	7/1/2013	7/2/2013	7/4/2013	7/5/2013
	Frequency	9	3	1	2
Wave 4	Dates	7/3/2013	7/4/2013	7/7/2013	
	Frequency	43	6	1	
Wave 5	Date	7/5/2013	7/11/2013		
	Frequency	5	1		

Table A.8: Frequency table of the donation dates by waves (1:1 unlimited, after splitting donation information, case 7)

Wave 1	Date	6/26/2013			
	Frequency	3			
Wave 2	Date	6/28/2013	6/29/2013	6/30/2013	7/1/2013
	Frequency	10	5	3	2
Wave 3	Date	7/1/2013	7/2/2013	7/4/2013	7/5/2013
	Frequency	9	3	1	2
Wave 4	Date	7/3/2013	7/4/2013	7/11/2013	
	Frequency	43	6	1	
Wave 5	Date	7/5/2013	7/7/2013		
	Frequency	5	1		

APPENDIX B

B.1 Supplementary tables for Chapter 3

Table B.1: Results on the donation amount (including zeros) from the two-limit Tobit model, separately for the treatment and control groups

	Dependent variable:			
	Donation amount (including zeros)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.485		0.030
		(0.301)		(0.308)
Constant	1.669*	3.314	2.347***	2.025
	(0.858)	(3.269)	(0.843)	(3.418)
Observations	55	55	50	50
Log likelihood	−123.7769	−122.482	−20.195	−120.190
AIC	251.554	250.964	244.390	246.381
BIC	255.568	256.986	248.214	252.117
Note:				*p<0.1; **p<0.05; ***p<0.01

Table B.2: Results on the donation amount (including zeros) from the two-limit Tobit model using the whole sample

	Dependent variable:			
	Donation amount (including zeros)			
	(1)	(2)	(3)	(4)
Treatment		−0.640	−5.662*	−5.339
		(1.171)	(3.341)	(4.710)
Treatment × advertised show-up fee			0.485	0.455
			(0.301)	(0.431)
Advertised show-up fee				0.030
				(0.308)
Constant	2.001***	2.332***	2.348***	2.025
	(0.601)	(0.847)	(0.836)	(3.418)
Observations	105	105	105	105
Log likelihood	−244.132	−243.982	−242.677	−242.672
AIC	492.263	493.964	493.354	495.344
BIC	497.571	501.926	503.970	508.614
Note:				*p<0.1; **p<0.05; ***p<0.01

Table B.3: Results on the donation amount conditional on giving from the Type I Tobit model with right-censoring and a log-level function, separately for the treatment and control groups

	Dependent variable:			
	ln (Donation amount conditional on giving)			
	Treatment group		Control group	
	(1)	(2)	(3)	(4)
Advertised show-up fee		0.090		0.073
		(0.068)		(0.067)
Constant	1.237***	0.281	1.125***	0.358
	(0.167)	(0.733)	(0.180)	(0.729)
Observations	35	35	36	36
Log likelihood	−47.850	−47.850	−52.264	−52.264
AIC	101.448	101.699	109.691	110.528
BIC	104.559	106.366	112.858	115.279
Note:				*p<0.1; **p<0.05; ***p<0.01

Table B.4: Results on the donation amount conditional on giving from the Type I Tobit model with right-censoring and a log-level function using the whole sample

	Dependent variable:			
	ln (Donation amount conditional on giving)			
	(1)	(2)	(3)	(4)
Treatment		0.116	−0.846	−0.090
		(0.245)	(0.795)	(1.038)
Treatment × advertised show-up fee			0.091	0.020
			(0.072)	(0.096)
Advertised show-up fee				0.072
				(0.064)
Constant	1.180***	1.122***	1.122***	0.366
	(0.123)	(0.172)	(0.171)	(0.695)
Observations	71	71	71	71
Log likelihood	−101.803	−101.691	−100.890	−100.266
AIC	207.606	209.383	209.781	210.532
BIC	212.131	216.171	218.832	221.846

Note: *p<0.1; **p<0.05; ***p<0.01

B.2 IRB approval for the field experiment in Chapter 3

THE UNIVERSITY OF ALABAMA | *Office of the Vice President for* Research & Economic Development
Office for Research Compliance

March 1, 2021

Cary Deck, Ph.D.
Professor
Department of Economics Finance & Legal Studies
Culverhouse College of Commerce
The University of Alabama
Box 870224

Re: IRB # EX-18-CM-076-R2-A: "Study of Allocation of Economic Resources"

Dear Dr. Deck:

The University of Alabama Institutional Review Board has reviewed the revision to your previously approved exempt protocol. The board has determined that the change does not affect the exempt status of your protocol.

Please remember that your protocol will expire on July 27, 2021.

Should you need to submit any further correspondence regarding this proposal, please include the assigned IRB application number. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants.

Good luck with your research.

Sincerely,

[redacted]

Carpantato T. Myles, MSM, CIM, CIP
Director & Research Compliance Officer

Jessup Building | Box 870127 | Tuscaloosa, AL 35487-0127
205-348-8461 | Fax 205-348-7189 | Toll Free 1-877-820-3066