Essays on Charitable Fundraising, Free Riding, and Public Good Provision

by

Ran Tao

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2017 by the
Graduate Supervisory Committee:

Daniel Silverman, Chair
Kelly Bishop
Nicolai Kuminoff

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

This dissertation consists of three essays on public good provision.

The first chapter develops a model of charitys choice of fundraising method under two dimensions of asymmetric information, quality and purpose. The main implication is a separating equilibrium where higher-quality charities choose to distinguish themselves by using a traditional fundraising method, while lower-quality ones exploit a low-stakes, take-it- or leave-it, "checkout method. An empirical application reinforced that charities of lower quality are more likely to adopt the checkout method. Despite this, consumers still choose to give in the equilibrium, due to the small requested amount of checkout donations, which disincentivizes serious thinking. Although exploited by lower-quality charities, the checkout method, along with purpose uncertainty, has the potential to alleviate the free-riding problem associated with public good provision and is, therefore, welfare improving.

The second chapter studies why corporations donate to charities and how their donations affect social welfare. I propose that firms make donations out of an image reason. In a model where two firms compete with each other, charitable donation could attract consumers and also signal firm overall social responsibility. I show that there exists an equilibrium where the high responsibility firm overdonates, resulting in a donation level closer to the socially optimal one. This leads to higher consumer welfare due to higher private good consumption as well as higher public good consumption when overdonation is prominent. Overall social welfare is enhanced. Empirical results support social image as an incentive for firms to donate.

The third chapter examines people's marginal willingness to pay for a change in local public good provision. We use a fixed effects hedonic model with MSA level data to study the effect of crime on local housing price. We explore the 1990s crime drop and use abortion data in 1970s and 1980s as an instrumental variable based on

Donohue and Levitt (2001). One result we find is that a decrease in murder of 100 cases per 10,000 people increases housing price by 70%. We further translate this result into a value of a statistical case of homicide, which is around 0.4 million in 1999 dollars.

ACKNOWLEDGMENTS

I am indebted to my committee members: Daniel Silverman, Kelly Bishop, and Nicolai Kuminoff. They showed me what it is like to look into the world with passion and curiosity and transform that passion and curiosity into research. Without their support and guidance, this dissertation would not have been possible. I also want to extend my thanks to ASU professors for their helpful comments and insights. And special thanks to Hong Zhao for all his support and joint work on the third chapter.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

FUNDRAISING UNDER TWO-DIMENSIONAL ASYMMETRIC INFORMATION: THE CASE OF CHECKOUT DONATIONS

1.1 Introduction

Charitable organizations exhibit substantial heterogeneity in many aspects, and the economic activity of charitable donation involves information asymmetry. One source of information asymmetry is the quality of charitable organizations. Even though the vast majority of charitable organizations are sincere in their efforts to support public purposes, the quality of charities could vary. For example, if a charity spends a large proportion of the funds raised on throwing expensive galas or other activities that do not directly contribute to the actual charitable programs of the charity, then the charity might not be considered as an efficient, or high-quality one. This level of efficiency or quality is not immediately apparent to a would-be donor, and requires time and effort to verify.

Another source of heterogeneity comes from the purpose of charities. Charities plead for numerous different causes. Some major purpose categories include education, environment, health, etc., and can be further divided into finer subcategories. Some charities have names that are fully aligned with their purpose, e.g., the Breast Cancer Research Foundation. Others, however, do not fully reveal their purposes through their names. For example, the charitable organization, Futures for Children, is dedicated to help children, but only children with American Indian heritage, to attain post-secondary education. Hence, for consumers encountering a charity for the first time, there exists uncertainty over the exact purpose of the charity. Previous literature has mainly focused on the quality dimension of asymmetric information. I argue that the purpose dimension is also important, because consumers might have different preferences for specific purposes. What is more, purpose is not the only aspect of charities that consumers have heterogeneous tastes for. For example, a charity with a primary goal of supporting women's education might support abortion as well, and a charity that mainly provides food to children might try to teach children religion at the same time. These are characteristics that are not immediately or costlessly verifiable, and consumers could potentially hold strong and different views about them.

In this paper, I study charitable fundraising under the two aforementioned dimensions of asymmetric information. Quality is unknown to consumers, but consumers homogeneously prefer charities with higher quality. The second dimension, which I call "purpose and others", encompasses any unobservable characteristics over which consumers have heterogeneous preferences. The second dimension is important for public good analysis, as public good provision suffers the underprovision problem due to free-riding among consumers, and a mismatch between "purpose and others" and a consumer's taste might provide the consumer with additional incentive to free-ride.

When information is asymmetric, consumers are unable to make fully-informed decisions. Yet, consumers are also sophisticated, in the sense that they are able and sometimes willing to do some research to learn about a charity's type before making a donation decision. For instance, by studying the tax return Form 990 or investigating outside ratings, consumers could obtain a good sense of a charity's quality. And by reading the program details and news articles of a charity, consumers can uncover the "purpose and others" of the charity. But in order to obtain the initially hidden information, consumers need to devote some time or resources to the research, hence search is costly. If they do decide to search, they would be able to make a serious donation when the charity type matches their taste, and donate less or simply opt out when the type is non-matching. However, since search is costly, there exists a trade-off between avoiding the search cost and making a more informed decision. This trade-off provides certain charities with the opportunity to exploit fundraising strategies that discourage consumers to search. And the fundraising strategy that is commonly referred to as "checkout donations" count as one of them.

As the name implies, checkout donations often occur at checkouts in supermarkets and department stores, where the cashier would ask the consumer whether he would like to donate a small amount of money to charity. Common questions include: "would you like to donate a dollar to charity" or "would you like to round up your purchase to the nearest dollar amount as a donation to charity?" Since the amount is so small, and the take-it- or leave-it nature requires only a simple answer of "yes" or "no", the decision does not require much serious thinking: in expectation, the combined value of the cause and quality either exceeds the value of the small donation, or it does not. Moreover, also due to the small donation size, the cost of searching exceeds the benefit of acquiring extra information. Hence, charities that do not want consumers to find out about their types, might want to take advantage of the "mindlessness" of consumers by using checkout donations. On the other hand, some other charities might prefer to have consumers know about their type in order to make a more serious, informed donation, and hence adopt more conventional fundraising methods where they, for instance, send letters in the mail to ask for open-ended donations. The conjecture is that the fundraising method a charity uses might serve as a signal to consumers on its hidden quality type.

In order to study this signaling mechanism, I set up a game between a charity with two-dimensional asymmetric information and consumers of heterogeneous taste for "purpose and others." The charity chooses a fundraising strategy between open-ended and checkout, which can also serve as a signal for quality. Consumers, after seeing the signal, play a simultaneous game where they decide to search or not and then how much to donate taking into account their beliefs and tastes. By focusing on a separating Bayesian Nash Equilibrium, I find that a charity of higher quality tends to adopt the conventional open-ended fundraising strategy, while a charity of lower quality utilizes the take-it- or leave-it checkout fundraising strategy. Hence, the fundraising strategy a charity adopts functions as a signaling mechanism through which quality is revealed. The next step I take is to seek empirical evidence to evaluate the model. Using data obtained from both an email survey and Better Business Bureau, I set up the testable hypothesis that charities with lower program percentage levels exhibit higher tendency to adopt checkout donation methods. Logit and Probit regression results indicate that, on average, one percentage point decrease in program percentage is associated with approximately one percentage point increase in the probability of using checkout donations. Hence, the empirical results are consistent with the separating Bayesian Nash Equilibrium predictions.

If it is truly the case that only lower quality charities adopt the checkout strategy and thereby separate themselves from others, then it might seem strange that consumers, being aware of the separation, still choose to donate to these charities, making checkout donation a common phenomenon. Although the model explains that the expected value of donating exceeds that of searching or not donating, there is still a fair question to raise regarding whether authority should step in and prohibit checkout donations all in all, as a means to prevent bad charities from exploiting of the "mindlessness" of consumers. To provide an answer to this question, I conduct a counterfactual analysis in which the checkout method is banned. The welfare result shows that compared to this world without checkout donations, the separating equilibrium in the baseline scenario actually exhibits higher ex-ante welfare. The reason behind this result does not involve costs of search, but rather concerns the second dimension of asymmetric information, where consumers' taste heterogeneity comes into play. When facing a checkout donation request, consumers do not bother to search or reject the small request, taking into account that the charity, although not high-quality, has at least a high chance of matching their taste for "purpose and others." So checkout donation, together with purpose uncertainty, has the potential to alleviate the free-riding problem that is characteristic of public goods. As a result, checkout donation enhances social welfare by improving private provision of public goods and, therefore, should not be banned.

One thing that should be acknowledged is the behavioral aspect of checkout donations. The presence of observers undeniably affects consumers' decision to donate. However, in this paper I only study the other fundamental aspect of checkout donations, and that is the "mindless" aspect. Whether there are more observers present or not, consumers are urged by the cashier, who evidently does not represent the charity, to make a snap decision. This imposed time constraint alone could spur skepticism on the credibility of the charity. Moreover, it is unlikely for the consumers to thoroughly inquire into the detailed purpose or other hidden characteristics of the charity under the time constraint. But my model suggests that exactly due to the "mindless" nature of checkout donations, that is, the small donation size and purpose obscurity, consumers reach the snap decision of donating to the charity, even if the charity might not be a great one. On the other hand, the behavioral aspect, although important, is not unique to checkout donations. For example, DellaVigna *et al.* (2012) find through a field experiment of door-to-door fundraising that a class of consumers avoid social pressure by "opting out" of the campaign if given the chance, or donate small amounts to lesson the discomfort of social pressure when not given the chance to "opt out." So one may expect that the observer effect at checkouts would also induce this class of consumers to donate.

1.2 Related Literature

One problem that is characteristic of voluntary provision of public good is the free-riding problem. The traditional model predicts that the total level of public goods provided is below the socially optimal level, and government provision crowds out private provision completely (Bergstrom *et al.*, 1986, for example). However, as noted in Andreoni (1988), household participation rate in voluntary giving is high, and government grants only partially crowd out private donations. Andreoni further shows that when assumptions in the traditional model are relaxed, then the traditional model "fails to explain either the extensive or intensive nature of giving."

There exists a strand of literature that gives attention to fundraising methods and investigates how fundraising behavior may attract higher levels of giving than the traditional model predicts. Andreoni (1998) studies the effect of "seed money" on total funds raised and concludes that the existence of seed money could pull the society out of the suboptimal zero-contribution equilibrium. Harbaugh (1998) adds a "prestige" component to individual's utility when analyzing how "brackets" (e.g., silver, gold, platinum) works to improve public good provisions. Cornes and Sandler (1984, 1994) study impure public goods and show that when private and public outputs are jointly provided by one single good, free-riding is less of a problem, as compared to pure public good models. Kotchen (2005) follows and extends this line of research on impure public goods by focusing on environmentally friendly goods with the existence of substitutes for both the public and private characteristics. Morgan (2000) models fixed-prize raffles as a fundraising method that mitigates free-riding and suggests that the negative externality of entering a fixed-prize lottery counteracts the positive externality from the public good and hence improves public good provision. And numerous experimental studies shed light on various fundraising methods such as rebates and matching grants (Eckel and Grossman, 2003; Karlan and List, 2007; Huck *et al.*, 2015), sequential giving (Potters *et al.*, 2007; Bracha *et al.*, 2011), gift exchange (Falk, 2007; Eckel *et al.*, 2016), etc. While the different fundraising methods could induce giving, some of them could also increase the overhead cost, and hence lower the perceived quality of a charity. Rose-Ackerman (1982) points out that donors might not like charities with high overheads and models charitable fundraising under asymmetric information. However, only more recently have studies focused on the potential for donors to acquire hidden information on charity quality.

Vesterlund (2003) studies sequential fundraising where a charity has the option to announce the initial contribution. In her model, there is a lead giver who moves first, where he decides whether to purchase information on quality, and a second donor who follows. The quality of the charity is either high or low. If it is low, then the donors do not receive any utility from the public good built. The charity chooses whether to announce or not to announce the initial contribution of the lead donor, and the lead donor can then decide whether to purchase information on quality or not. If the charity chooses to make the announcement, then the lead donor has the opportunity to make a second donation, together with the second donor. If there is no announcement of the initial contribution, then the two donors donate simultaneously. Vesterlund focuses on perfect Bayesian equilibria where the charity chooses to announce initial contributions. Under information cost that falls into a certain range, the lead giver would purchase information when and only when announcement is made. Then a lead giver who knows the charity is of higher quality would make a substantially larger donation than if the quality is public information, in order to send a signal to the second donor. Hence, in this particular equilibrium, the announcement strategy itself does not reveal the quality type of the charity, but since the lead donor chooses to purchase information and overdonate, the quality type is revealed and the free-riding problem is mitigated. Andreoni (2006) considers a similar setting where a lead giver has the option to purchase information from the charity and donates an amount that might signal the quality to the second-round givers. One difference is that a bad charity could have either zero value to consumers or a low but positive value. And the consequence is that there will be a wealthy, voluntary lead giver who gives an exceptionally large initial gift to a good charity in order to convince the followers that the quality is indeed high. Following the theoretical papers, Potters *et al.* (2007) obtain experimental evidence supporting the theoretical predictions of leadership in sequential giving.

This paper belongs to the pool of studies on fundraising methods. The model shares some features in Vesterlund (2003) such as unknown quality of charity and the ability to costly search for it. Yet, this paper also adds some new features to the existing literature. First, it examines a distinct fundraising strategy, namely checkout donations, which has not yet been formally studied. Second, it introduces an extra dimension of asymmetric information, namely "purpose and others" of charities, on top of unknown quality, and this purpose uncertainty assists the checkout fundraising method in alleviating the free-riding problem. As for the unknown quality, the low type does not offer zero utility as in previous literature, but rather a strictly positive utility for the public good it builds. Another small distinction is that, unlike aforementioned studies, the search ability is not restricted to any particular giver, thus making all consumers "sophisticated" players of the game. Last but not the least, the signaling mechanism is different. In Vesterlund (2003) and Andreoni (2006), the lead donor has the opportunity to signal the charity's quality by donating different amounts. In this paper, I consider the possibility of the charity itself to directly signal its quality to the consumers by using differentiating fundraising strategies. This means that although the consumers could potentially search for quality, they do not have to solely rely on the search mechanism to reveal the hidden information.

In addition to the theoretical studies mentioned above, there are a few experimental studies that are close to the checkout fundraising method. Andreoni *et al.* (forthcoming) set up a field experiment with the Salvation Army, which collects cash donations at store entrances during the Christmas season. Although the charity does not ask for take-it- or leave-it amounts, the donations collected are typically small, and are, therefore, similar to checkout donations. The authors find in a set of experiments that explicit verbal asking increases giving significantly. They also find that consumers who are uncomfortable with refusing to give seek to avoid the explicit asking but not the silent solicitation. These results do not undermine this paper's results, as checkout fundraising does not necessarily require any verbal asking. Since the cashier does not work for the charity, it is very common for the cashier to skip the verbal ask and simply let the consumer decide to donate a dollar or not on the checkout screen. Moreover, checkout fundraising does not have to take place in physical stores but could also work at online checkouts. This is another reason why this paper focuses on the signaling aspect of checkout fundraising and not on the behavioral aspect.

The experimental study by Charness and Cheung (2013) is more directly related to the checkout fundraising method. The authors conduct a field experiment in a restaurant, where they set up a donation jar at the cashier with a suggested donation amount of $0.5, $1, or $2, or without a suggested amount. The result is that the treatments with suggested amount of $0.5 and $1 yield significantly higher average daily donations than the $2 and no suggestion cases. Although the experiment considers suggested amounts instead of take-it- or leave-it amounts, it has a similar flavor to checkout donations. And the experiment results are compatible with my model's predictions, which will be discussed in a later section.

1.3 Model

1.3.1 Preliminaries

Charities are heterogeneous in quality and "purpose and others." To simplify notation, I will use "purpose" to denote the second dimension. A charity can be of two quality types, $Q \in \{high, low\}$. [1] The purpose of a charity can either be Purpose A or Purpose B, i.e., $P \in \{A, B\}$. The joint distribution of the charity's type is uniform, and is common knowledge, but the realized type is privately known by the charity. The charity's strategy set includes: to ask for an open-ended donation that is optimal for the consumer, and to ask for a take-it- or leave-it donation in a small amount, i.e., less than a certain upper bound. I call the open-ended fundraising method "open" and the take-it- or leave-it method "epsilon." The upper bound of epsilon is denoted as $\bar{\varepsilon}$. It is natural to assume that $\bar{\varepsilon}$ should be small, hence "mindless." This is because, in the real world, it is rare for a charity to ask for a donation bigger than a dollar without offering the option to change the amount.

There are two consumers, A and B. Their utility takes the form:

$$U_i(x_i, G) = x_i + \alpha_i log(G)\,,$$

where x_i is Consumer i's private good consumption, G is the total level of public good or donations raised, that is, $G = \sum g_i$, and α_i is Consumer i's taste, or more

[1] The so-called "low" type is not necessarily a bad charity per se, it is just not as lean as the high type. For example, CharityWatch, a nonprofit charity watchdog organization, considers a charity to be efficient when its program percentage is 75% or higher (*Charity Rating Guide and Watchdog Report*, Volume 62. December 2012. CharityWatch.). But a charity with a program percentage lower than that cutoff is by no means a bad charity for sure. The public good such a charity builds still benefits consumers, but consumers might judge them as a lower-quality charity, due to reports from watchdog organizations such as CharityWatch.

precisely, perceived value of the public good. This perceived value is determined by both the quality and purpose of the charity that provides the public good. Both consumers prefer a high-type charity to a low-type one, so α_i is higher if the charity has $Q = high$. But consumers value the two purposes differently. Consumer A cares more about Purpose A, and consumer B finds Purpose B more important. And this is commonly known. So α_i is higher if the charity's purpose is more aligned with Consumer i's interest. To be more specific, I let α_i take the form:

$$\alpha_i = \begin{cases} 1, & Q = high,\ P = i \\ q, & Q = low,\ P = i \\ p, & Q = high,\ P \neq i \\ qp, & Q = low,\ P \neq i \end{cases}$$

where P, $i \in \{A, B\}$, and q, $p \in (0, 1)$. q being strictly larger than zero means that even if the charity type is low, consumer still gains some, although possibly slender, utility from the public good it builds. And p being strictly less than one indicates that consumers feel strictly better after giving to one cause than the other, even though they feel good about giving to either.

Consumer i chooses x_i and g_i to maximize utility subject to budget constraint

$$x_i + g_i = w\,.$$

Consumers can conduct some search to find out Q and P of the charity simultaneously at a fixed cost, $K > 0$. The cost enters utility linearly, if search is undertaken. Before donating to maximize their utility, consumers decide whether to search or not.

The structure of the game is the following. In the first stage, nature selects the charity's type to be one of $highA$, $highB$, $lowA$, $lowB$, with equal probability for simplicity. Knowing its type, the charity chooses its fundraising strategy to be open or epsilon. If the charity chooses epsilon, it also decides a numerical amount, ε, which does not exceed the upper bound, $\bar{\varepsilon}$. After observing the fundraising strategy, Consumer A and Consumer B update their prior belief about charity's type and play a simultaneous game. In the simultaneous game, A chooses to search or not search and then g_i to maximize expected utility, and B does the same thing at the same time. The total level of public good, G, is simply the total donation from A and B, i.e., $g_A + g_B = G$. Then each consumer's total realized payoff is the utility from private good consumption plus public good consumption adjusted by taste and minus search cost as needed. And the charity's payoff is the total level of G raised.

1.3.2 Main Result: A Separating Pure-Strategy BNE

Existence of the Equilibrium

The main focus is on a separating pure-strategy Bayesian Nash Equilibrium. The equilibrium is separating in the sense that a high-type charity would adopt the open fundraising method, while a low type chooses to use epsilon. Both consumers, when facing the same information, choose the same action, either search or no search. The separation feature makes the equilibrium particularly interesting. It indicates that the fundraising strategy the charity adopts is no longer just means to raise proceeds, but can also work as a signaling mechanism to inform the consumers about charity's hidden type.

The upper bound of the epsilon amount is set to be q, the taste parameter for a low type, and remains the same throughout all equilibrium discussions. This assumption should not be too restrictive, because q is also the optimal amount a purpose-matched consumer would donate given he knows that he faces a low type and that the non-matched consumer does not give.[2] As argued earlier, checkout donations are supposed

[2]When a consumer is approached by a seemingly untrustworthy charity (or simply a person)

to be small, since it is hard for a charity to ask for a take-it- or leave-it amount higher than consumer's willingness to give. Hence, an upper bound higher than q seems to be unfair, while an upper bound moderately smaller than q would not affect the equilibrium results.

Proposition 1. For any q small enough and p large enough, there exists an interval of K that support a separating pure-strategy Bayesian Nash Equilibrium, in which:

(a) A high-type charity uses strategy open and a low type uses strategy epsilon with $\varepsilon = \frac{q+qp}{2} log(2)$;

(b) When consumers observe open, both update their belief to "high for sure," and both search. Then the consumer who cares less about the purpose fully free-rides on the one who cares more;

(c) When consumers observe epsilon, both update their belief to "low for sure," take on no search, and donate ε.

Proof: See Appendix.

Proof Intuition:

To prove such an equilibrium exists, I need to show that once the equilibrium is reached, there is no deviation on both the consumers' and the charity's sides. First of all, given the charity adopts open, and both consumers know consequently that the charity is of type high, both consumers would want to search and find out about the purpose and then optimize utility, provided the search cost K is small enough. On the other hand, when the search cost K is large enough, given the charity asks for ε, both consumers know that the quality is low, and hence they would rather just give

and asked for money, it seems generous enough to donate an amount assuming the charity is not a good one and no one else is donating. And if the charity asks for a large take-it- or leave-it amount that exceeds the consumer's highest willingness to pay, and the consumer counter-offers a smaller amount, such as q in this case, it is unlikely that the charity would decline.

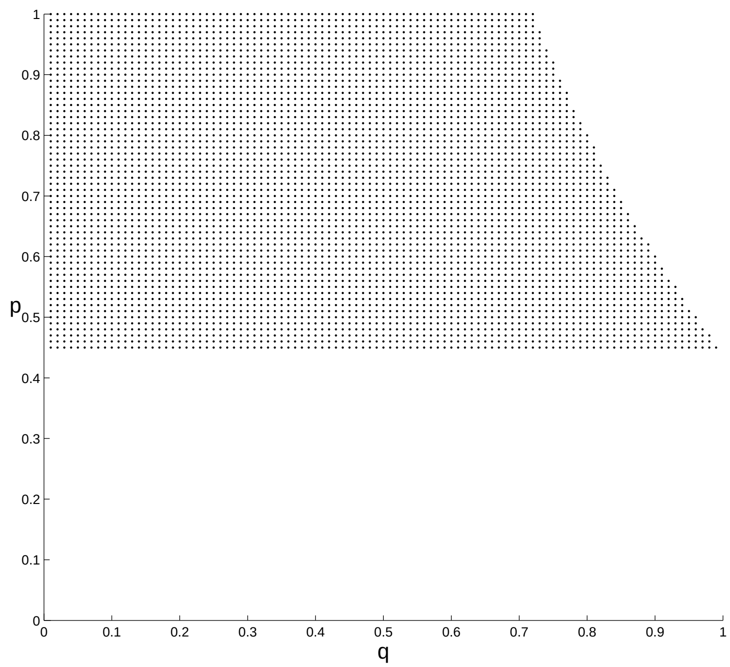
ε without knowing the true purpose, as the ε amount is too small to be worth the trouble of finding the purpose out. On the charity's side, when q is small enough, i.e., the difference between high and low types is significant, the high type would not want to ask for epsilon donations, because it would be taken as a low type, hence the payoff would be lower. And when p is large enough, i.e., the non-matching purpose still gives the consumer high enough value from donating, then both consumers at the epsilon side have high incentives to donate even though the quality is low. Consequently, for a low type, it would not deviate to open and pretend to be high, because the consumers would then search and know its true type, which would lead to lower payoff for the low-type charity.

Figure 1.1 shows in the $q - p$ space the region where the separating equilibrium can be supported. The bottom region fails to support the equilibrium because the consumers do not care enough for a public good that does not match their interest, so they are less willing to contribute to checkout donations. As a result, low type can no longer benefit much from utilizing the checkout strategy and might as well deviate to the open-ended strategy. In the upper-right corner, where q is too large, the difference between a high and a low charity is small. Then a high charity might also want to use the checkout strategy, as it would not mind being regarded as a low type and collecting non-trivial checkout donations from both consumers.

A Numerical Illustration

To provide a better sense of magnitude for the search cost, K, I calculate the interval of K that supports the separating equilibrium for $q = 0.2$ and $p = 0.8$. The choice of these numbers has no special meaning and is just for illustration purposes. When consumers' perceived values of certain types of public goods are set to be the above, the search cost that supports the separating equilibrium ranges from 0.0069 to 0.165.

Figure 1.1: Parameter Values Supporting the Separating Equilibrium
Notes: The figure shows in the $q - p$ space the region (shaded) where the separating equilibrium in Proposition 1 can be supported. The horizontal axis, q, is the parameter reflecting how much utility from public good consumption is discounted due to the lower quality of the charity. The vertical axis, p, is the parameter reflecting how much utility from public good consumption is discounted due to an unmatched purpose as compared to a matched one.



Any search cost within this range is not too big for the consumers to incur when facing a high charity, and is also not too trivial for the consumers to "mindlessly" donate the small take-it- or leave-it amount. Also, this range of search costs is well below the donation given to a high charity by the consumer who cares more about the purpose, which equals to 1.

1.3.3 Other Types of Equilibria

There exist various types of equilibria along with different off-path beliefs to support them. In this subsection, I briefly discuss the existence (or nonexistence) of other types of equilibria. The first type to consider is the reverse separating equilibrium where a high type adopts an epsilon strategy, while the low type chooses open. It turns out that this type of separation cannot be supported for any values of search cost. Suppose such a separating equilibrium exists, where high adopts epsilon and low chooses open. Then, consumers know it is a high type when seeing epsilon, and low when seeing open. If both consumers search at epsilon and donate regardless of purpose, then there would be profitable deviation to no search. Suppose the high type asks for an epsilon amount that is the highest amount a consumer would give without searching and given the other consumer gives, then its payoff must be larger than the upper bound, q (as the low type receives payoff higher than q in the baseline separating equilibrium). Since the low type cannot receive a payoff higher than q at the open side, the low type would want to deviate to epsilon and pretend to be high, as there is no search at epsilon. Suppose the high type asks for q, then one case is that consumers do not search and both donate, then again, the low type would want to deviate. The other case is that consumers search and only the purpose-matching consumer donates. If consumers do not search at the open side, then the low type, receiving less than q would again benefit from deviating to epsilon. If consumers do search at the open side, then the high type would want to deviate to open to receive the higher, optimal amount from the purpose-matching consumer. To sum up, there is always profitable deviation on the charity's side. Hence, a separating equilibrium where high type chooses epsilon and low, open does not exist.

One way to "pool" is to choose open no matter what type the charity is. Such equilibrium could exist if the search cost is so small that the consumers choose to search even when facing a small epsilon request. However, assuming the same range of search costs and parameter values that support the baseline equilibrium, it can be shown that this type of pooling equilibrium does not exist. When seeing open, the consumers would always search (see proof of Proposition 2). If consumers' off-path belief is that a type that deviates to epsilon is low, then they would choose no search and donate ε, just as in the baseline separating equilibrium. If, on the other extreme, the consumers' off-path belief is that only a high type would deviate to epsilon, then it is clear that consumers' best responses is still to choose no search and donate ε. Since consumers never search off-path, a low type would benefit from deviating to epsilon, breaking the pooling equilibrium.

Another way to pool is for both types to choose epsilon with a same ε amount. Such equilibrium could exist when search cost is so large that the charity could ask for high epsilon amounts without the consumers searching, so even the high type is better off at the epsilon side. This pooling can be broken when the search cost is not that large. Even if search cost is large, when the epsilon amount asked is restricted to be less than $\frac{1}{2}$, that is, half of the optimal amount consumer gives at the open end in the main separating equilibrium, it would always be profitable for the high type to deviate to open. And the low type would not deviate given consumers' off-path belief satisfies the Cho-Kreps criterion.

One more type of equilibrium that could be of potential interest is a semi-separating equilibrium, where the high type chooses open, while the low type randomizes between open and epsilon. For this equilibrium to exist, the low type has to be asking for an epsilon amount of $\frac{q}{2}$, because otherwise the low type would not want to play a mixed strategy. Given this condition is satisfied, the semi-separating equilibrium can be supported by the range of search costs in the main separating equilibrium. At the open side, consumers' belief is in accordance with the Bayes' rule, and they would choose to search and find out the hidden information. At the epsilon side, they believe it is a low type for sure, do not search, and donate the epsilon amount requested. This semi-pooling equilibrium gives a separation result with a similar flavor as in the baseline separating equilibrium, but a less stark one. In the next section, I compare welfare levels between the baseline separating case and a scenario where epsilon strategy does not exist. If instead of the fully separating equilibrium, I use this semi-separating result as the baseline scenario, I can only obtain a weak comparison. So instead, I focus on the fully separating case in the welfare discussion.

1.4 Welfare Discussions

1.4.1 A Question

The intuition behind the baseline separating equilibrium is straightforward. When there is a signaling mechanism, the high-type charity would not risk being misjudged as low quality by asking for small donations. Rather, it would want the consumers to know about its quality and make a serious donation, even if that means losing entirely the consumer who does not care much about its purpose. On the other hand, the low quality charity would not want the consumers to make fully-informed decisions. If the low type pretends to be a high type, then the consumers will search and find out about both its true quality and purpose. Then, the consumer who does not care about its purpose would opt out, while the consumer who cares more would not donate much more than the epsilon amount since he knows the quality is low. Hence, the gain on the intensive margin does not offset the loss on the extensive margin, and the charity would rather adopt the checkout strategy to benefit from consumers not knowing and not caring to know about its purpose.

This equilibrium result clearly implies that the low, and only the low quality charities are taking advantage of the "mindlessness" of consumers. If this is the case, then it might seem strange that consumers would still be willing to give, even when these checkout charities are known to be the bad ones. The model provides a simple answer to this question. When a charity asks for a small checkout donation, even though the consumers know that the charity is not highly efficient, they also do not regard the public goods raised by this charity as negligible. So because the search cost is non-trivial, when facing a small, take-it- or leave-it request, consumers do not bother with the search, and simply donate the small amount in response to the slightly positive value of a lower quality charity. Moreover, they understand that there is a 50% chance that the charity's purpose would match their interest, and for the other 50%, even though the purpose does not match, the perceived value is still high enough. As a result, consumers would rather donate the small amount than privately consuming it simply because the charity might not be of the highest quality.

But even if consumers willingly let bad charities exploit checkout donations, one might still wonder whether the authority should step in and ban the checkout donation method all in all, because in this case, low-quality charities will lose this chance to exploit the "mindlessness" of consumers. In the next subsection, I provide an answer to this question by comparing the baseline separating equilibrium with a situation where the checkout donation method is prohibited. It turns out that consumers are better off when checkout donations exist.

1.4.2 Counterfactual Analysis

In this counterfactual analysis, I take the main separating equilibrium in Section 3 as the baseline scenario, and consider the alternative scenario where the charities are not allowed to use the epsilon strategy. The range of search cost and the upper bound of epsilon are set to be the same as in the baseline model for a fair comparison. If epsilon is removed, and the only strategy the charity could use is open, then consumers are no longer able to infer the quality without searching. And indeed, they will always choose to search and find out the hidden information.

Proposition 2. When there is only the open strategy, there is an equilibrium where consumers always choose to search. And this equilibrium yields lower ex-ante welfare than the separating equilibrium in the baseline model.

Proof: See Appendix.

The result of Proposition 2 grants merit to the usage of checkout donations. This might not seem straightforward, because one might think that the consumers would be better off if they always search to find out about all hidden types and then optimize under full information. However, this is not the case. The baseline separating equilibrium yields higher ex-ante welfare even though purpose is unclear at the epsilon side. The reason is two-fold and does not concern the search cost. (Consumers' expected payoff from the baseline equilibrium outcome is already higher than the expected payoff from the no-epsilon scenario even before considering the extra search cost in the no-epsilon scenario.) First, the existence of the signaling mechanism in the baseline equilibrium effortlessly reveals charity's quality–one of the two dimensions of asymmetric information. When a low type cannot pretend to be high and dilute the overall expected quality, inefficiency due to information asymmetry is largely reduced. Secondly, and more importantly, the remaining dimension of asymmetric information, namely the "purpose and others" of the charity, has the potential to alleviate the free-riding problem that is characteristic of public good provisions. When a charity asks for a small donation, and purpose is still uncertain, the consumer who cares less and would have otherwise chosen to free-ride, has now the incentive to donate. This gain on the extensive margin increases overall public good provision. Certainly, purpose obscurity alone might not improve public good provision. But it works under the checkout donation method: Since the take-it- or leave-it request is so small, consumers do not bother to search and consequently have less incentive to free-ride without having the exact information on purpose and other characteristics they potentially care about.

If it were a normal consumption good that we are considering, this positive effect of purpose uncertainty on welfare would not show. But for a public good, there already exists welfare loss due to the positive externality from public good provision. Purpose uncertainly induces consumers to make uninformed decisions when asked for a checkout donation, but at the same time, it counteracts the positive externality by hindering the consumers from free-riding. In other words, it enhances welfare of the whole by hurting the individual in a Nash game. This logic is somewhat analogous to the one in Morgan (2000) on fixed-prize raffles.

To sum up, when there is two-dimensional information asymmetry, an extra strategy of checkout donations helps improve social welfare by clearing up the inefficient dimension but also ameliorates the free-riding problem by keeping the other dimension intact along with the incentive to donate. And when the checkout donation method is actually welfare-enhancing, there is certainly no reason to prohibit its existence.

1.4.3 Connection to Charness and Cheung (2013)

Charness and Cheung (2013) set up a donation jar at a restaurant cashier and conduct four treatments where either there is a suggested amount of \$0.5, \$1, or \$2, or no suggested amount. The average daily donation for the first two treatments are \$3.568 and \$3.023, respectively, which are significantly higher than that of the last two treatments, which are \$1.581 and \$1.506, respectively. The mechanism in Charness and Cheung is different, in particular, they consider suggested amount instead of take-it- or leave-it amount. But the magnitude is small, similar to checkout donations, and it could be the case that the small suggested amount also serves as a signal to consumers saying that the charity is of low quality. As a result, when seeing \$0.5 or

$1, consumers might think it is a lower quality charity, do not care to search, and donate. When there is no suggested amount, however, consumers might engage in search and find out that the charity is not so lean, hence only the consumers who really care about the charity's purpose donate. As for the case of a $2 suggested amount, one possibility is that it exceeds the amount consumers are willing to donate when expecting the charity to be bad. As a results, they might search and find out about the true type, leading to a lower donation rate. Charness and Cheung did not provide the participation rate under the four treatments. The average donations from the first two treatments, which are roughly twice the magnitude of the last two treatments, could have resulted from an increase on the intensive margin. But it could also very well be due to a gain on the extensive margin, i.e., lesser free-riding. This is why the results by Charness and Cheung might be explainable using the reasonings of this paper and imply that fundraising strategies that have a similar flavor as the checkout strategy might have a similar effect of alleviating the free-riding problem and enhancing welfare.

1.5 Empirical Evidence

The main separating equilibrium provides theoretical basis for the following testable hypothesis.

Hypothesis: A higher-quality charity, i.e., a charity that spends a higher percentage of total expenses on charitable programs, has less tendency to adopt a checkout donation strategy.

In this section, I empirically test the above hypothesis using logit and probit estimations.

1.5.1 Data and Variables

The main data source I use is the website of BBB Wise Giving Alliance, a nonprofit organization that helps donors make better-informed decisions. For over seven hundred charities, BBB collects and reports important figures from their tax return Form 990 and evaluate charities according to BBB's own standards. My key independent variable of interest and additional control variables come from BBB's database.

The dependent variable is a dummy variable that indicates whether a charity uses a small-size donation strategy. I collected the data for this variable through an email survey of the charities listed on the BBB website. The main question asked is: "Does your organization raise funds through collaborating with grocery/department stores at checkouts or through other special methods to collect donations in small amounts, e.g., around $1?" I do not restrict to donations occurring at store checkouts because I want to focus on small-size donations with checkouts being a typical example. This way, I also obtain more affirmative observations. One problem is that only a fraction of the surveyed charities actually responded to the email. This potentially creates a non-response bias, which I will discuss in more details in a later subsection.

The key explanatory variable is the percentage of total expenses spent on charitable programs, or, more shortly, program percentage. This number is commonly used by charity rating institutes as an important indicator of a charity's quality. A higher program percentage usually implies higher efficiency. Since the hypothesis I intend to test is that a better charity has less tendency to adopt a checkout strategy, I expect to see a negative coefficient for this key variable of interest.

BBB also provides their own charity rating measures. They have listed in total twenty standards mainly concerning a charity's governance, effectiveness, finances, etc. and reported for each charity whether those standards are met. From this piece of data, I obtain the variable, number of BBB standards not met by a charity. Since BBB's rating measure is not as straightforward and commonly used as program percentage, I only regard it as supplementary information on a charity's quality.

A third variable, reserve, defined as net asset divided by total income, might also provide some additional information on quality. The inclusion of this variable is motivated by reports provided by the charity rating institute, CharityWatch. According to CharityWatch, a good charity should not have asset reserves of more than three years, since the dollars raised should be spent to carry out their missions, not to stock up on assets.[3]

In terms of other control variables, I first include charity's age. Age might have an effect on the tendency to adopt checkout donations. For instance, a long-existing charity might have a larger member base, and hence relies more on traditional fundraising methods, while a young charity might be more eager to try newer methods like checkout ones.

I also include the natural log of total income in the year where the charity is reviewed by BBB. Total income tells about the size and scale of a charity. It could be the case that a larger charity has more resources to build relationships with grocery stores to facilitate checkout donations. But a larger charity also tends to have more resources to hold other types of fundraising events (e.g., galas), so it could also be the case that a larger charity tends to forgo checkout opportunities.

Some additional charity characteristics are also included in the regression. Number of paid staffs complements the income variable in controlling for the charity's size, but more on the physical side. The size of the board might affect the charity's behavior, since a larger board means more monitoring. And top compensation presents a tradeoff between CEO's incentive and charity's program percentage.

[3] *Charity Rating Guide and Watchdog Report*, Volume 62. December 2012. CharityWatch.

Lastly, I create a set of dummy variables indicating which purpose categories a charity belongs to. There are seven large categories such as education, health, and international relief. The category division is originally based on summary statistics found in CharityWatch reports. I could have used the 20 categories listed by BBB, but due to a small sample size, I decide that a coarser division should be better for identification.

1.5.2 Summary Statistics

Table 1.1 reports some summary statistics. The overall sample size is 185. The subsample of charities with checkout strategies is 30, and the subsample without is 155. The first two columns give the variable means and medians of the subsample with checkout strategies, and the next two columns the ones without. The last two columns report p-values for the t- and z-test for the hypothesis that the two groups have the same means or medians, respectively. For the key variable, program percentage, the group with checkout strategies has both lower mean and lower median than the group without. And to reduce effects of possible outliers, I winsorise the variable at the 5% level, which is also done in the later regressions. The t- and z-tests show that the 3.21 difference in mean and the difference of 1 in median are both significant at the 10% level after the winsorization.

So without controlling for any other factors, it seems that on an average sense, charities with checkout strategies have lower program percentage, hence efficiency, than the ones without. Another variable that displays significant difference in mean and median for the two groups is log income. The group with checkout strategies are, on average, larger in scale than the one without. However, since there are no controls or regression models used, these results are inconclusive.

Table 1.1: Summary Statistics

Notes: The table presents the summary statistics of charity characteristics based on 185 nation-wide charities with data on whether checkout donation is implemented. Data on checkout donation is obtained through an email survey. All other data are obtained from BBB, which provides charity information based on Form 990. *Program percentage* is the the percentage of total expenses spent on charitable programs. *BBB standards not met* is the number of BBB standards that a charity does not meet. *Reserve* is net asset divided by total income. *Age* is a charity's age. *Log(income)* is the natural logarithm of a charity's total annual income. *Compensation %* is the top executive total compensation divided by total income. Board size is the size of charity board. *# paid staff* is the number of charity's paid staff. Columns (1) to (4) report the sample mean and median for two subgroups based on whether checkout donation is implemented, and columns (5) and (6) list the p-values for the test on equality of sample mean and median using t-test and Wilcoxon z-test, respectively. See the text for further details on the definition of variables. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	A: Mindless (N=30)		B: No Mindless (N=155)		Tests on Sample Difference	
	Mean	Median	Mean	Median	Mean	Median
	(1)	(2)	(3)	(4)	(5)	(6)
Program %	80.60	82.00	83.81	83.00	0.048**	0.065 *
BBB standards not met	0.66	0.00	1.00	0.00	0.387	0.097
Reserve	1.00	0.66	0.95	0.67	0.824	0.623
Age	29.13	27.50	31.58	23.00	0.557	0.780
Log(income)	16.94	16.55	15.72	15.60	0.000***	0.000***
Compensation %	0.02	0.01	0.03	0.02	0.000***	0.072*
Board size	18.66	16.00	15.16	12.00	0.131	0.124
# Paid staff	260.93	51.00	140.14	21.50	0.477	0.041**

1.5.3 Model and Results

To estimate the effect of program percentage on "probability of checkout," I run a probit regression (and a logit regression as a robustness check) with the following model:

$$Pr(Checkout\,Donation) = F(X\beta)$$

where X is a vector of variables including program percentage and other controls, and F is the standard normal distribution for the probit (or logistic distribution for the logit) model. The estimation results for three specifications are reported in Table 1.2.

Table 1.2: Charity Quality and Checkout Donation

Notes: The table presents the coefficients from probit and logit regressions of the use of checkout donation on charity quality. The dependent variable is equal to 1 if a charity uses checkout donation and 0 otherwise. *Program percentage* is the the percentage of total expenses spent on charitable programs. Seven category dummies for charity purpose are created. The last row reports the sample average marginal effect of *program percentage*. The numbers in the parenthesis are standard errors. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Probit			Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
Program %	-0.0283*	-0.0542***	-0.0610***	-0.0516*	-0.1000***	-0.1069***
	(0.0147)	(0.0183)	(0.0203)	(0.0269)	(0.0334)	(0.0366)
BBB standards not met		-0.0640	-0.0616		-0.1013	-0.1074
		(0.0817)	(0.0926)		(0.1466)	(0.1623)
Reserve		-0.0238	0.0477		-0.1021	0.0146
		(0.1202)	(0.1341)		(0.2271)	(0.2563)
Age		-0.0165**	-0.0144*		-0.0310**	-0.0258*
		(0.0074)	(0.0081)		(0.0144)	(0.0150)
Log(income)		0.2758*	0.2965**		0.5029*	0.5439**
		(0.1431)	(0.1489)		(0.2632)	(0.2746)
Compensation %		-11.3234	-13.2894		-20.1049	-23.7835
		(9.1942)	(10.0334)		(17.0454)	(18.4936)
Board size		0.0099	0.0050		0.0164	0.0060
		(0.0136)	(0.0155)		(0.0234)	(0.0282)
# Paid staff		-0.0003	-0.0004		-0.0004	-0.0006
		(0.0002)	(0.0003)		(0.0004)	(0.0005)
Category dummies			✓			✓
Observations	181	181	181	181	181	181
F-statistics	3.81	28.80	45.44	3.83	31.97	45.25
Pseudo R^2	0.024	0.181	0.285	0.024	0.181	0.284
Average marginal effect of program	-0.0067	-0.0108	-0.0105	-0.0068	-0.0112	-0.0104

In the first specification, I include only the key variable, program percentage. The probit coefficient in column one is -0.028 and is significant at the 10% level. The average marginal effect is -0.0067, which means that on average, a 1 percentage point decrease in program percentage is associated with a 0.67 percentage point increase in the probability of using checkout.

In the second column, I include all the control variables except for the category dummies. The variable program percentage is still negative and significant, but now at the 1% level. The average marginal effect also becomes larger at -0.0108, which means that a 1 percentage point decrease in program percentage is associated with a 1 percentage point increase in probability of checkout. Two other variables significant are age and log income. Age has a coefficient of -0.017, is significant at the 1% level, and has an average marginal effect of -0.0033. As mentioned before, a possible explanation is that older charities have more affiliate members so it is easier to perform more traditional fundraising acts such as membership appeals, while younger charities do not have specific targets and thus use checkout methods to target everybody. The other significant–but only at the 10% level–variable is log income, with a positive coefficient of 0.276. So it seems that larger charities have more tendency to use checkout methods. One thing that should be noted is that, a charity's size does not tell much about its quality–a charity with high income could have a low program expense percentage and a high asset reserve, hence low efficiency. Indeed, the correlation between log income and program percentage is around 0.11 and insignificant at the 10% level for both the Pearson and Spearman methods. Both reserve and BBB standards unmet have negative signs, which is consistent with the conjecture that higher quality charity does less "checkout", but both are too insignificant to draw any conclusion from.

In the third column, I further include the set of category dummies. The inclusion of dummies does not change the results much. The coefficient for program percentage is now-0.061, and still highly significant, with an average marginal effect of -0.0105. Age has become slightly less significant, but still significant at the 10% level. And log income is now significant at the 5% level. Reserve now has a positive coefficient, but is still highly insignificant. The results for the rest of the variables are similar to the second specification. The only category dummy that is significant is education, which has a marginal effect of 0.14. This can be interpreted as charities that have education as their main purpose are 14% more likely to adopt checkout methods. This is not surprising, in the sense that it is common to encounter charities concerning kids and their education in places such as grocery stores.

The results from the logit regressions are highly similar to the ones from probit. For the key variable, program percentage, the average marginal effect is also around -0.007 when no controls are added, and around -0.01 when controls are added. In any case, program percentage is negative and significant. And on average, a 1 percentage point decrease in program percentage is associated with approximately 1 percentage point increase in probability of checkout.

To provide a richer illustration of program percentage's effect than the average marginal effect, I also calculate the "predicted probabilities of checkout" of an "average" charity with program percentage valued at different sample deciles. The charity is "average" in the sense that all its non-dummy variables (except for program percentage) takes value at sample means, and in terms of the category dummies, it falls into either education or health category, since they are the two largest categories and education is the only category with positive and significant coefficient. The predicted results for both the Probit and Logit models are reported in Table 1.3. As depicted in the table, for both categories in both models, as program percentage goes from a higher decile to a lower one, the probability of checkout rises in a monotonic manner. For instance, for an average charity in the education category, its predicted probability of using checkout is 0.043 if program percentage is in the highest decile, as compared to 0.494 for the lowest decile. This is to say, this average charity with a program percentage of 72 is over ten times more likely to adopt a checkout strategy than an otherwise same charity with a program percentage of 100. The results for an average charity in the health category are similar. The probability of checkout increases as program percentage decreases, with 0.021 for the highest program percentage decile and 0.369 for the lowest. The same patterns appear in the Logit model as well as when using the second specification in Table 1.2 without controlling for category dummies.

Table 1.3: Predicted Probability of Implementing Checkout Donation

Notes: The table presents the predicted probabilities of a charity using checkout donation based on probit and logit regressions with full explanatory variables from Table 1.2. Probabilities are evaluated at deciles of program percentage and sample means of all other non-dummy explanatory variables. Columns (1) and (3) provide the predicted probabilities for a charity belonging to the education category, and columns (2) and (4) the health category. The last row reports the differences between the lowest and highest deciles. Column (5) gives the market share of charities in education or health categories in each decile in terms of number of charities.

	Predicted Probability				Market Share of Education and Health
	Probit		Logit		
	(1)	(2)	(3)	(4)	(5)
1^{st}: program%=72	0.494	0.369	0.513	0.389	0.81
2^{nd}: program%=76	0.398	0.282	0.407	0.293	0.42
3^{rd}: program%=79	0.329	0.224	0.333	0.231	0.25
4^{th}: program%=82	0.266	0.173	0.266	0.179	0.47
5^{th}: program%=83	0.247	0.158	0.245	0.164	0.79
6^{th}: program%=85	0.210	0.130	0.208	0.137	0.71
7^{th}: program%=88	0.161	0.095	0.160	0.103	0.52
8^{th}: program%=91	0.120	0.068	0.121	0.077	0.65
9^{th}: program%=95	0.078	0.041	0.083	0.052	0.56
10^{th}: program%=100	0.043	0.021	0.050	0.031	0.67
Difference between 1^{st} and 10^{th}	0.451	0.348	0.463	0.358	

The results presented above provide empirical support for the prediction from the separating equilibrium in the model section that it is the lower-quality charities that have a higher tendency to use checkout donations.

1.5.4 Non-Response Bias

Since the dependent variable is collected from an email survey, the data could potentially suffer from a non-response bias. From the 761 emails sent, only 185 got replied. In order to make sure that the response group is representative of the entire surveyed group, I perform a set of tests to show that the characteristics of the response group are not significantly different from those of the entire surveyed group. Table 1.4 gives the test results.

Table 1.4: Non-Response Bias

Notes: The table presents p-values of a set of hypothesis tests on non-response bias. For all non-dummy variables, column (1) shows p-values from testing the hypothesis that the response group and the entire group have the same variable means, using t-test with unequal variance. Column (2) shows p-values from testing on the equality of variable medians of the two groups. For dummy variables, the test is on the equality of proportion and the p-values are reported in column (3). Column (4) reports the p-values from testing the hypothesis of zero Pearson's correlation between a charity's characteristics and whether it responded. More variables are considered here than those used in regression specifications. *Log(assets)* and *Log(net assets)* are the logarithm of a charity's total assets and net assets. *Fundraising %* and *Administrative %* are the percentage of fundraising expenses and administrative expenses on total expenses. Charities are divided into four region groups based on their incorporated states, and region dummies are created accordingly. *Leverage* is the ratio between a charity's debts and total assets. See the text for details on the definition of other variables. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Mean (1)	Median (2)	Proportion (3)	Correlation (4)
Program %	0.626	0.935		0.531
BBB standards not met	0.389	0.563		0.281
Reserve	0.631	0.870		0.631
Age	0.101	0.279		0.049**
Log(income)	0.878	0.623		0.848
Compensation %	0.028**	0.174		0.205
Board size	0.358	0.334		0.406
# Paid staff	0.209	0.907		0.359
Log(assets)	0.934	1.000		0.918
Log(net assets)	0.932	0.656		0.910
Fundraising %	0.671	0.449		0.570
Administrative %	0.228	0.289		0.141
Leverage	0.720	1.000		0.621
Category dummy 1			0.256	0.134
Category dummy 2			0.323	0.203
Category dummy 3			0.891	0.859
Category dummy 4			0.665	0.576
Category dummy 5			0.947	0.931
Category dummy 6			0.623	0.527
Category dummy 7			0.764	0.700
Region dummy 1			0.874	0.840
Region dummy 2			0.607	0.511
Region dummy 3			0.813	0.762
Region dummy 4			0.341	0.217

Column 1 shows the p-values from testing the hypothesis that the response group and the entire group have the same (non-dummy) variable means allowing for unequal variance. The only significant one is top compensation. But since the test is performed on thirteen variables and only one difference is significantly non-zero at the 95% confidence level, this is close to the expected number of significant differences if the two groups are sampled from the same distribution. Therefore, there is no clear sign of a systematic bias due to nonresponses. The second column gives the p-values from testing the equality of variable medians of the two groups. None of the differences are significant. For the dummy variables, a test on the equality of proportion is performed and none of the p-values in column 3 indicates any significant difference. The last column reports the p-values for testing the hypothesis of zero-correlation between a charity's characteristics and whether it responses, and the only significant variable is age.

The above test results suggest that the response group is overall not significantly different from the entire surveyed group. Hence, the inference made from the response group should not suffer from any systematic non-response bias.

1.6 Conclusion

The meaning for checkout donations to exist is two-fold. First, checkout donations provide the inferior charities with a platform to exploit the "purpose and others" dimension of asymmetric information, i.e., the "mindlessness" of consumers, to achieve a higher public good provision. Since the take-it- or leave-it amount is so small, consumers do not bother to search. And since consumers do not acquire information regarding purpose and others of the charity, they simply donate the asked amount. So in a sense, checkout donation and purpose obscurity hurt the consumers by not letting them make fully informed decisions. Yet exactly through this "trick," checkout donations achieve a higher public good provision, which, in turn, benefits the consumers. Second, checkout donations provide charities with the opportunity to signal their quality. If there is no such signal to separate high and low quality charities, the overall quality is diluted by the existence of low quality charities and consumers become even more reluctant to give. So the extra strategy of checkout donations clears up this inefficiency. Then a reasonable question to ask would be, is this signaling feature unique to the checkout donation strategy? The answer is likely to be no. In the real world, we observe similar fundraising strategies with slight differences. For instance, a frequently observed strategy similar to the take-it- or leave-it one is to also collect donations at checkouts but ask for open-ended donations with small suggested amounts to select from, e.g., "$1, $2, $5 or other." On the contrary, letters arriving in the mail might ask for "$100, $200, $500, or other." It is not "open vs. epsilon" anymore, but it still has a similar flavor–one could then think of the suggested amount as a signal of quality. Furthermore, one might argue that search is not really an option in the grocery store due to time constraint. If this is the case, then the locational choice itself might serve as a signal of quality, since inferior charities are more likely to think of exploiting this inconvenience of shoppers. Hence, this study on checkout donations might serve as a primitive study of a whole class of fundraising strategies that operate as signaling mechanisms for quality. And even if this class of fundraising strategies is adopted by lower quality charities only, we might not want to hastily judge such strategies with a negative frame of mind, because they might very well come with negative externalities that hurt consumers individually, while benefiting them collectively by combating the positive externality of public goods.

Chapter 2

A WELFARE IMPLICATION OF CORPORATE GIVING: THEORY AND EVIDENCE

2.1 Introduction

It is frequently observed that corporations make donations to charities. This phenomena might seem contradictory to the universal assumption that firms are profit-driven. It is especially puzzling, as charities build public goods, and publilc goods should only benefit consumers, not corporations. Since it is hard to imagine that corporations are giving money away on a purely altruistic ground, there has to be underlying channels of chritable giving through which profits are generated. In this paper, I propose two intertwined channels linked to firm image to explain the donation behavior of profit-driven firms.

Corporate donation is often covered in news articles, disclosed on company websites, or made observable to consumers through other forms of communication. For example, the Fortune magazine, with help from The Chronicle of Philanthropy, has published the top 20 most generous companies of the Fortune 500 in 2016.[1] Without help from newspapers or magazines, stores could also put up signs telling shoppers how much they have donated. So charitable donations, due to high visibility, could generate an image effect among the consumers. I call this channel a direct image effect. When this effect exists, firms would want to use charitable donations to enhance company image and attract consumers. However, charitable giving is not the only action deemed socially responsible. Corporate social responsibility (CSR) in-

[1]Preston, Caroline. "The 20 Most Generous Companies of the Fortune 500." Fortunes, 22, 06, 2016, http://fortune.com/2016/06/22/fortune-500-most-charitable-companies/.

corporates a wide variety of activities that "go beyond firm's legal and contractual obligations", as described by Bénabou and Tirole (2010). Therefore, environmental friendliness, product safety, employee relations, and so on all factor into an overall level of responsibility. Not all of these CSR activities count towards public goods. For example, a firm that treats its employees exceptionally nicely is regarded as socially responsible, but this does not directly benefit all consumers in the economy. But all things equal, consumers might still like such a firm better than other ones, simply because it is "nicer." As for environmental friendliness, which is usually considered a public good, there exist past events indicating that consumers care. For example, Barrage *et al.* (2014) find out that following the British Petroleum oil spill in 2010, consumers punish BP by switching to its competitors. Yet, a firm's overall responsibility might not be fully observable to the consumers. A great deal of social responsibility is predetermined by a firm's production technology, (e.g. whether the technology is clean or not, safe or not, etc.) which is most likely unobservable. But consumers like firms that are "nicer." As a result, it is sensible for firms to reveal its virtue to consumers using an observable signal. And even if a firm's CSR activities are observable, it could be hard to get a precise measure on them. On the other hand, charitable donations are easy to measure in monetary terms. Hence, charitable donation, due to its visibility and measurability, could be a good signal for overall CSR. In this way, charitable donation affects firm image both directly and through the revelation of firm's overall responsibility.

To incorporate both the direct image and the signal motives into a competitive environment, I adopt a simple Hotelling type of model (Hotelling, 1929), where two firms compete with each other using both price and donation. All else equal, a lower price attracts more consumers, but a higher donation does so as well due to an enhanced social image. Both the price and the direct image effects are reflected in consumer's utility. So there is not only an optimal price, but also an optimal donation level that yields the firm highest profits. I also assume that firms are of different levels of responsibility, which is a predetermined trait of the firms. For example, oil companies are commonly regarded as environmentally unfriendly. But an oil company that uses efficient drilling rigs and adopts better safety measures could be more responsible than its competitors. When such responsibleness is hidden from the consumers, charitable donation works as a costly signal, and I show that there exists an equilibrium where a high type overdonates in order to separate itself from the low type.

The model explains why it could be optimal for profit-driven firms to give money to charities, especially when there is asymmetric information. But the information asymmetry leads to overdonation of the high type, which is a source of inefficiency for the firm, as compared to a world with no information asymmetry. However, this information failure might not yield a bad consequence for the overall economy. This is because when public goods are involved, there already exists a source of inefficiency due to free-riding among consumers. Classic models such as Bergstrom *et al.* (1986) predict that voluntary provision of public goods is always socially suboptimal. Here, the information asymmetry leads to a high level of firm donation, which could enhance social welfare. First of all, the overdonation of the high type leads to a higher company image enjoyed by the consumers. Second, the higher overall firm donation level crowds out voluntary giving, but also allows the consumers to allocate spared income to private good consumption, while still being able to enjoy a high level of public goods due to the firms' contribution. And last but not the least, if firm donation is very high, consumers could become constrained to free-ride on the firms as much as they would like to, so free-riding is mitigated on a certain level. Overall, when the image and signaling mechanisms exist, firms might earn lower profits due to charitable givings, but the consumers and society as a whole benefit from better public good provisions.

To investigate whether there exists an image effect of charitable donations, I use donation data collected from database NOZAsearch to test whether a firm that is "closer" to consumers make more donations. I construct a "closeness" variable using the Bureau of Economic Analysis input-output data that tells how much of a firm's costumer base consists of common consumers. The "closer" to consumers a firm is, the more it should care about its image, and hence the more it should donate. I find that firm donation is positively affected by the firm's "closeness" to consumers, and this effect is especially significant when only local donations are considered, as local donations are more readily observed by consumers. In an additional test, I use a fractional multinomial logit model to test whether there is any relation between a firm's industry and the cause category of its charitable donations. I obtain the results that firms donate to causes that are related to its line of business. For example, food firms donate more to the food cause, and medical firms donate more to the health cause. One explanation for this result is again the image story.

2.2 Related Literature

Although this paper concerns mainly charitable donations, it can fit into the broader literature of corporate social responsibility (CSR). One main theme of the CSR literature focuses on why firms behave as good "corporate citizens" (term adopted from Bénabou and Tirole (2010)) and many studies in this literature are closely related to this paper. For example, Kotchen and Moon (2012) finds that corporations use CSR activities to offset previous irresponsible behavior. Although they do not explicitly mention image, their result clearly imply that firms care about their company image and do not want consumers to see them as socially irresponsible. Siegel and Vitaliano (2007) divide products into "search" and "experience" goods, where "search" goods have attributes fully observable to consumers before the purchase, while "experience" goods need to be consumed first for their attributes to be revealed. The authors find that companies that sell "experience" goods are more likely to be socially responsible than the ones that sell "search" goods. This result implies that CSR activities should have some signaling values when product attributes are unobservable. Servaes and Tamayo (2013) show that advertising intensity is positively related to CSR news coverage. By interacting advertising intensity and CSR intensity, they further show that CSR has a more positive effect on firm performance when advertising intensity is higher. Their interpretation is that advertising increases consumers' awareness of the firms, leading to a stronger effect of CSR on firm value. The relationship between CSR and firm value is another main theme of the CSR literature. Margolis *et al.* (2007) conduct a meta-analysis of such empirical studies, whose results are mixed, and the average relation is positive but small.

I should note that the standard dataset used in the CSR literature is the Kinder, Lydenberg, Domini Research & Analytics (KLD) Social Ratings Database, which is a panel dataset on social performance of public firms. This paper, however, does not use the KLD dataset. This is mainly because KLD encompasses a broad range of CSR activities including things such as corporate governance and employee relations, and is not pertinent to the donation topic of my study. In this study, I only focus on donations, because it is a very direct and clear form of contribution to public goods. Hence, I hand-collect charitable donation data from the NOZAsearch database. Unlike KLD variables that are all dummy variables, the donation data I collect has specific values or ranges. Therefore, by focusing on donation data only, I forgo data of some CSR activities that are less relevant to public goods, but arrive at a more precise measure of donation than the KLD data.

Unlike the above mentioned studies, this paper does not only focus on empirical findings, but also builds a model to explain corporate donations and subsequently analyze the effect on public good provisions. The provision of public good has been extensively studied. The classic model predicts that the level of public goods provided is below the socially optimal level due to free-riding among consumers, and government provision crowds out private provision completely (Bergstrom *et al.*, 1986, for example). Besley and Ghatak (2007) model CSR in a Bertrand competitive environment and reach their main conclusion that CSR does not change the total level of public good provisions, either. Bagnoli and Watts (2003), on the other hand, reach an inconclusive result and state that CSR might increase public good provision under certain conditions.

The paper is also related to some papers in both the economic and finance literature regarding how firms might adopt costly means to signal hidden characteristics. A well-known example is the study by Milgrom and Roberts (1986) on advertising intensity. The authors point out that expensive advertisements are not just used by firms to introduce products to consumers, but rather a costly signal to convey the message that the product is of good quality. Miller and Rock (1985) argue that a firm's dividend is not just money paid to the shareholder but also works as a signal to let the stock market know of the firm's good financial performance. The signaling story in this paper has a similar flavor, and builds on the classic Spence (1973) job market signaling model.

2.3 Model

2.3.1 Setup

There are two firms in the market of a private good. Although they belong to the same market, the products they sell are not the same in the eyes' of the consumers, because the consumers have slightly differentiated taste for the two products. One example of this situation is the market of soda drinks, where Coca Cola and Pepsi are almost the same product, but consumers might prefer one over the other to varying extents. This type of differentiated taste is captured by a standard Hotelling (1929) model. If a consumer is "located" closer to firm 1 than to firm 2, then he has a personal taste that makes him more inclined to purchase the good from firm 1, when all else is equal. I assume that the length of the market is l, and use x to denote a consumer's "distance" to firm 1. However, the distance of the consumer is not the only thing that determines whether he purchases from firm 1 or 2: the two firms also have different overall responsibility. I let r_1 denote the predetermined level of CSR (e.g., clean technology) of firm 1, and assume that r_1 is higher than r_2 of firm 2. The overall responsibility is reflected in consumer's utility of purchasing the good. The consumer's purchasing decision is also affected by the prices the firms charge, as in the standard Hotelling model.

Apart from the private good, there is also a public good in this economy. Both the firms and the consumers could contribute to the public good by making charitable donations. Firms' donations, d_1 and d_2, are measured as effective donations, that is, how much public good is built. (Effective donation, the amount of public good raised, might be lower than the actual monetory donation made by the firm.) Consumers get to enjoy the public good, but not the firms. Hence, the action of donating is costly for the firms and does not benefit firms directly. However, it might help firms attract consumers through a direct image effect. So all else equal, making a larger donation means capturing a larger share of the market. On top of being an image enhancer, I further consider the case where donation could function as a signal when there is asymmetric information. I assume that the responsibility of firms are known among the firms, but not the consumers. That is, consumers only know that firms could be either high type or low type with equal probability, but do not know the actual types of the firms. In this case, the firms, especially the good type, might want to signal their types through public good donations. I assume that before setting the product price, the donation decision is made and observed by all agents.

The profit function of a type j firm can be written as

$$\pi_j(d_j, r_j) = b(d_1, d_2, r_1, r_2) - k(d_j, r_j)$$

where b is the benefit, and k is the cost function, with $b_{r_j} > 0$, $b_{d_j} > 0$, $k_{d_j} > 0$ and $k_{d_j d_j} > 0$. The cost function also satisfies the single-crossing condition $k_{r_j d_j} < 0$. That is, the firm with the higher responsibility, r_1, has a lower marginal cost of donating. This assumption is especially realistic if there exists positive assortative matching between firms and charities, that is, higher responsible firm donates to higher quality charities. Then firms of higher responsiblity spends less to have a certain level of effective donations made. There also exists anecdotal events that show that good charities might refuse donations from companies considered irresponsible, such as tobacco companies. Also, the cost of donating could also include non-pecuniary costs such as transaction cost. If firms with higher social responsibility have better established partnerships with charities, then they are likely to incur lower transaction costs when arranging donations. The benefit $b(r, d)$ is derived as in the Hotelling model. The benefit is calculated as price times the share of the market. But since the pricing decision is made after the donation decision is made, price is just a function of d, and does not show in the b function. The share of the market a firm gets is determined by the consumers' decision of which firm to purchase the private good from.

The consumers observe the effective donations and product prices of the two firms. They then decide which one of the two firms to purchase the good from. For a consumer at distance x, his utility from the purchase is $c = -x - p_1 + \alpha r_1 + \mathbf{1}(d_1 \geq d_2) \cdot \beta \cdot (d_1 - d_2)$, if purchased from firm 1; and is $c = -(l - x) - p_2 + \alpha r_2 + \mathbf{1}(d_2 \geq d_1) \cdot \beta \cdot (d_2 - d_1)$, if purchased from firm 2. Since the good itself is the same whether purchased from firm 1 or 2, the utility gained directly from the good itself is simply set to be 0. The distance is subtracted from utility as in the standard Hotelling model, and p_1 and p_2 are prices charged by the two firms, respectively, and are thus also subtracted. α is a parameter between 0 and 1 that represents consumers' general awareness about corporate social responsibility. Consumers know that the high type is associated with r_1 and low type with r_2, but when there is information asymmetry, they do not know whether a firm is high or low. In this case, consumers rely on the signals, d_1 and d_2 to make the purchasing decision.

Apart from serving as a potential signal, d_1 and d_2 also enter consumers' utility directly as an image component. $\mathbf{1}(d_1 \geq d_2)$ is an indicator, so if firm 1's donation exceeds that of firm 2, then consumers will receive a positive direct image effect from firm 1, which is the difference between the two donation levels multiplied by a parameter, β, also between 0 and 1; but there is no direct image effect for firm 2, the one with the lower donation level. The reason I use the difference between donation levels instead of the actual levels to characterize the direct image effect is that for certain industries, firms are expected or even required to give, especially locally, to charities, so the level of giving should not fully capture the direct image effect on consumers, but rather, a firm needs to "go the extra mile" to convince the consumers that it is socially responsible. Therefore, I assume that only the firm with the higher donation level attracts consumers with a direct image component in the utility, and the image is characterized by the exceeded amount in donation.[2]

The consumers decide the firm to purchase from to maximize utility of the private good consumption. In addition to the purchasing decision, consumers could also make voluntary donations to build the public good. Consumers' objective is to maximize utility. But since the purchasing decision and donation decision do not affect each other, the purchasing decision could be viewed as the first stage, and the discussion of the donation decision is delayed to the welfare section.[3]

The sequence of events is the following. Knowing their own and each other's type, firms first decide simultaneously how much donations to make. Then, after donations are observed by all agents, firms decide simultaneously how much price to charge for the good they sell. After observing firms' decisions, consumers choose which of the two firms to purchase the good from. (They also decide their own donations to the public good, which will be discussed in the welfare section.) In the following subsections, I solve the model backwards for both the symmetric and asymmetric information cases.

2.3.2 Solving the Model, Symmetric Information

Before focusing on separating equilibria under asymmetric information, I first solve the simple model assuming that firm type is observed by the consumers. Since consumers only differ in x, the distance to firm 1, I solve the model using the standard

[2] The results in Propositions 1 and 2 do not change if I replace $\mathbf{1}(d_i \geq d_j) \cdot \beta \cdot (d_i - d_j)$ with $\beta \cdot (d_i - d_j)$ in consumer's utility function. It only changes the solution in social planner's problem and makes the welfare in social planner's problem even higher compared to the welfare under symmetric and asymmetric information.

[3] I use log-linear form for consumer's full utility, that is $u_i = c_i + w - g_i + \gamma log(G + d_1 + d_2)$. Then, the solution for the c_i part and the solution of g_i are separate from each other. Moreover, the purchasing decision is determined by firms' donation decisions, and firms' donation decisions are not affected by consumers' voluntary donations. So the purchasing decision should not change whether the full utility including public good consumption is considered.

technique in solving the basic Hotelling model. I start by finding the indifferent consumer, denoted as x^*, who is indifferent between purchasing from firm 1 or firm 2, that is

$$-x^* - p_1 + \alpha r_1 + \beta(d_1 - d_2) = -(l - x^*) - p_2 + \alpha r_2 \, ,$$

then the indifferent consumer is located at

$$x^* = \frac{1}{2}l - \frac{1}{2}(p_1 - p_2) + \frac{\alpha}{2}(r_1 - r_2) + \frac{\beta}{2}(d_1 - d_2) \, .$$

Since all consumers closer to firm 1 than the indifferent consumer will purchase the product at firm 1, and all consumers farther than the indifferent consumer go to firm 2, the profits of the two firms can be computed as follows:

$$\pi_1 = x^*(p_1 - k) - k(d_1, r_1)$$

$$\pi_2 = (l - x^*)(p_2 - k) - k(d_2, r_2)$$

Since the firms set prices after the donation decisions are made and observed, the optimal prices can be written as functions of the d's. Then after substituting and simplifications, the profits of the firms can be rewritten as functions of d's only:

$$\pi_1 = 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1 - d_2)]^2 - k(d_1, r_1)$$
$$\pi_2 = 2[\frac{l}{2} + \frac{\alpha}{6}(r_2 - r_1) + \frac{\beta}{6}(d_2 - d_1)]^2 - k(d_2, r_2)$$

Claim 0. Let (d_1^{SI}, d_2^{SI}) be the symmetric information solutions. Then we have $d_1^{SI} > d_2^{SI}$.

Proof: See Appendix.

It is intuitive that the firm with a lower cost to donate donates more. It is also intuitive that the more consumers care about the direct image from donations, the more firms are willing to donate in order to attract consumers, as the following result states.

Proposition 0. The sum of d_1^{SI} and d_2^{SI} is increasing in β.

Proof: See Appendix.[4]

2.3.3 A "Separating" Equilibrium, Asymmetric Information

The focus of the paper is more on the scenario when the firm types are not observable by the consumers. In this case, firms use public good donations as a signal for their hidden type. I will focus on separating equilibrium where firm 1 donates more than firm 2 to signal its high type because firm 1's cost of donating is lower. The separating equilibrium is not the standard one in signaling games. Instead of one firm of two possible types, there are actually two firms of two types. But the spirit of "signaling" remains, in the sense that when consumers observe the equilibrium levels of donations, they fully infer the type of the firms. The game includes two firms competing with each other to attract consumers using donation and price. Since price is set after donations are made and observed, the best response price is just a function of the donation decisions. So in a separating equilibrium, consumers only need to observe the donation levels to know the firm type and equilibrium price and make the decision of which firm to purchase from accordingly. On the firms' side, the firms fully anticipate consumers' reaction to the donation decisions and maximize their profits accordingly. Since the existence of equilibria hinges upon the belief consumers hold when observing the donation levels, I will first pin down one particular belief and discuss separating equilibria based on that belief.

Consumer Belief

Suppose that firm 1's equilibrium action is to donate $\widetilde{d}_1$, and firm 2's equilibrium action is $\widetilde{d}_2$, where $\widetilde{d}_1 > \widetilde{d}_2$. Let the consumers' belief be the following: if they

[4]For Claim 0 and Proposition 0, only interior solutions are considered. When β is small, that is, the direct image effect is small, firms might not have enough incentive to donate to public goods.

observe a donation level greater than or equal to $\widetilde{d}_1$, they believe it is a high type; if they observe a donation level strictly lower than $\widetilde{d}_1$, they believe it is a low type. The construction of this belief is adopted from Spence (1973). Given this particular belief, I can state some properties of separating equilibria in the following two claims. Both claims aim to compare the donation levels in any separating equilibrium under asymmetric information with donation levels in the symmetric information solution mentioned above.

Claim 1. If $(\widetilde{d}_1, \widetilde{d}_2)$ is a separating equilibrium outcome under the above belief, then it has to be the case that $\widetilde{d}_1 \geq d_1^{SI}$.

Proof: See Appendix.

Under certain parameter levels, it could be the case that the (d_1^{SI}, d_2^{SI}) solution in the symmetric information case can be supported as a separating equilibrium under the asymmetric case. This requires that the responsibility component in consumers' utility does not differ too much for the two types of firms, that is, $\alpha\Delta r$ is small. I will provide a specific numerical example later in the paper to provide an idea of what is "small". However, I will focus on the case where $\alpha\Delta r$ is large, that is, firm's type matters a lot to consumers, as the separating equilibrium under asymmetric information is then different from the symmetric information solution. So for the results stated later in the paper, I will impose the assumption that $\alpha\Delta r$ is large enough so that under asymmetric information, firm 1 donates strictly more than under symmetric information. The intuition behind this case is that when consumers care about firm's responsibility, the low type has an incentive to pretend that it is a high type. Then, the real high type, firm 1, would have to overdonate as compared to the symmetric case, in order to separate itself from the lower type and ensure that the lower type would not want to mimic the high donation.

Claim 2. A separating equilibrium with $\widetilde{d}_1 > d_1^{SI}$ has the following properties:

(1) $\widetilde{d}_1 + \widetilde{d}_2 > d_1^{SI} + d_2^{SI}$

(2) $\widetilde{d}_1 - \widetilde{d}_2 > d_1^{SI} - d_2^{SI}$

Proof: See Appendix.

Claim 2 states that for any separating equilibrium with firm 1 overdonating, both the sum of and the difference between the two firms' donations exceed those in the symmetric information case. The higher sum is contributed solely by firm 1's over-donation, as firm 2's donation actually decreases as compared to the symmetric case. But as shown in the proof, since firm 2's best response with respect to firm 1's donation has a negative but flat slope, the sum of the two firms' donations is higher. And since firm 1 increases donation while firm 2 decreases donation, the difference between the two firms' levels is also higher under asymmetric information as compared to the symmetric case. The intuition behind firm 1's increase in donation is explained in the previous claim. For the decrease in firm 2's donation, the intuition is the following. As firm 1 increases its donation, it also gets a higher share of the market than in the symmetric case. A higher donation leads to a higher price, but since firm 2 is left with a lower share of the market, its marginal benefit of donating at the symmetric equilibrium level is lower than the marginal cost. So it would move to a lower donation level so that marginal cost and marginal benefit are equal again.

Based on the above results, I further the analysis by showing the existence of one particular separating equilibrium, due to the reason that I would like to use one equilibrium as the reference point for asymmetric information case when comparing social welfare with the symmetric information case. As mentioned before, in order for the separation to hold, firm 1 needs to make enough donation such that firm 2 has no incentive to mimic firm 1's high level. The construction of the consumers' belief makes sure that firm 1 would not want to deviate downwards from its equilibrium level, because it would be mistaken as a low type. The higher firm 1's equilibrium level, the less incentive there is for firm 2 to mimic firm 1. The particular equilibrium I state below has firm 1 donating at a level where firm 2 has exactly zero incentive to pretend to be high type.

Proposition 1. When firm's type q is unknown to consumers, there exists a "separating equilibrium" with (d_1^{AI}, d_2^{AI}) as the equilibrium donation levels, where

(1) d_1^{AI} is such that firm 2 has exactly no incentive to deviate to d_1^{AI} from d_2^{AI};

(2) Given d_1^{AI}, d_2^{AI} solves firm 2's FOC: $\frac{\beta}{3}[\frac{l}{2}+\frac{\alpha}{2}(r_2-r_1)+\frac{\beta}{6}(d_2-d_1^{AI})]-k_d(d_2,q_2)=0$;

(3) The equilibrium prices are best responses to the donation levels;

(4) The consumer belief is that if $d \geq d_1^{AI}$, it is a firm of type q_1; if $d < d_1^{AI}$, it is of type q_2. And consumers choose to buy good from the firm that yields higher utility.

Proof: See Appendix.

The proof of the proposition is detailed in the Appendix. Here I provide some brief intuition to why there is no incentive to deviate for both firms. First, firm 1 would not deviate to a higher level of donation than d_1^{AI} because given d_2^{AI}, firm 1's profit goes down as firm 1 increases its donation. But firm 1 would also not want to decrease donation, because with the belief constructed this way, firm 1 would be regarded as a low type if deviated down, and any lower donation level could not bring firm 1 a higher profit than staying at d_1^{AI}. As for firm 2, it would not deviate to d_1^{AI} by construction in condition (1). And it would not deviate to any level higher than d_1^{AI} even though it could appear to be high type because, intuitively, if firm 1 wouldn't benefit from deviating up, firm 2 would certainly not benefit from that, as it is more costly for firm 2 to spend its money on donations than firm 1. Finally, for any donation level lower than d_1^{AI} and other than d_2^{AI}, firm 2 would not benefit from deviating to those because d_2^{AI} solves the condition in (2), which is firm 2's FOC given d_1^{AI}.[5]

Claim 2 provides some properties of separating equilibria in general, so it should apply to the equilibrium (d_1^{AI}, d_2^{AI}) as well: $d_1^{AI} + d_2^{AI} > d_1^{SI} + d_2^{SI}$ and $d_1^{AI} - d_2^{AI} > d_1^{SI} - d_2^{SI}$. That is, both the sum of and difference between the two firms' donations are higher in the equilibrium in Proposition 1, where there is asymmetric information, than when information is symmetric. In the separating equilibrium, information asymmetry results in inefficiency for the firms, because a costly and dissipative signal is needed in order to convey the hidden information. However, when there is public good involved, there is also another type of inefficiency, namely the underprovision of the public good. Due to the tendency to free-ride among consumers, the increase of firm donation would crowd out consumers' voluntary donations. But this also allows consumers to enjoy the spared income. And the increase in the two firms' difference might also improve consumer welfare, since a larger difference in donation is reflected as a stronger image effect in consumers' utility. So consumer welfare should go up. However, the question of concern is whether the increase in consumer welfare offsets the loss in firm profits and makes social welfare higher when there is asymmetric information.

2.4 Welfare Comparison

To calculate welfare, I first specify consumers' full utility, that is, utility from private good consumption as well as public good consumption, along with the decision of donating to public goods. Suppose there are N consumers in the economy, and

[5] Under the same parameter levels, equilibrium donations higher than d_1^{AI} for firm 1 might also support separation, where firm 2 setting its equilibrium donation by again solving its FOC. These equilibria will have firm 2 strictly prefer not to mimic firm 1's level instead of being indifferent as in (d_1^{AI}, d_2^{AI}). I only look at (d_1^{AI}, d_2^{AI}) since the same welfare result would apply to those equilibria.

consumer i's full utility is the following[6]:

$$c_i + w - g_i + \gamma log(G + d_1 + d_2)$$

where c_i is the utility from consuming the product, as specified in the previous section, w is an exogenous income that is the same to all consumers and can be used to purchase a numeraire good that is not produced, g_i is the consumer's voluntary donation to public goods, and $\gamma log(G + d_1 + d_2)$ is the utility from total public good consumption, which includes all consumers' donations, that is G, and firms' donations. [7] As mentioned earlier, the purchasing decision and donation decision do not affect each other . So I regard the c_i part as maximized in a first stage, and consumers play a game where each chooses the voluntary donation, g_i, to maximize his full utility in the second stage. This approach is valid due to both the log-linear form of the utility, and the fact that the purchasing decision is determined by firms' donation decisions, which are not affected by consumers' voluntary donations. If consumers could choose for the firms how much to donate, then the whole model would not be much different from a standard voluntary donation game where consumers decide how much of their endowment to contribute to public good. So it is an important assumption that firms' donation decisions are made upfront by some decision maker independent of the consumers.

Suppose there is a social planner who maximizes social welfare as the sum of all consumers' utility and the firms' profits. That is, the planner can choose d_1, d_2, p_1,

[6]The reason of assuming N consumers instead of a continuum of consumers is a technical one. If there is an infinite number of consumers, then the simple Nash equilibrium concept cannot be applied to solve the donation game between them. So instead, I adopt the common approach in setting up voluntary donation games, that is to assume that there are N consumers in the economy. One way to reconcile this with the previous infinite setting is to argue that even though there is a finite amount of consumers, their demand for a good could be of a different, continuous measure. Another possibility is to alter the previous Hotelling setting into a market where N consumers locate uniformly but discretely on a length of l. The model can still be solved with the same approach, but is simply more straightforward to illustrate with the usual continuously uniform market.

[7]If CSR acitivities of the firms also count as public goods, then G could also include those.

p_2, which firm to purchase from and how much g_i to donate for every consumer in order to maximize:

$$C + W + G + N\gamma log(G + d_1 + d_2) + \Pi$$

where C is the aggregate private consumption, W the aggregate wealth, and Π the aggregate profits of the two firms. The social planner's solution is optimal, because it takes into account the social benefit of the public good, so there is no free-riding among the consumers. However, when consumers make their own donation decisions using the simple Nash equilibrium concept, they do not take into account the positive externality their own donation creates, and hence the public good is suboptimally provided. In the baseline case, where there is no asymmetric information, firms donate d_1^{SI} and d_2^{SI}, respectively. So given this, consumer i chooses g_i to maximize utility. And one constraint is that g_i cannot be strictly negative. As for the asymmetric information case, firms donate the equilibrium amounts d_1^{AI} and d_2^{AI} as in Proposition 1, and consumer i chooses $g_i \geq 0$ to maximize utility $c_i + w - g_i + \gamma log(G + d_1^{AI} + d_2^{AI})$. Similar to the social planner scenario, the social welfare is the sum of all consumers' utility plus firms' profits. This can be denoted as $C^{SI} + W + G^{SI} + N\gamma log(G^{SI} + d_1^{SI} + d_2^{SI}) + \Pi^{SI}$ and $C^{AI} + W + G^{AI} + N\gamma log(G^{AI} + d_1^{AI} + d_2^{AI}) + \Pi^{AI}$ for the symmetric and asymmetric information cases, respectively. Proposition 2 gives a welfare comparison between the two.

Proposition 2. Social welfare is higher under asymmetric information than under symmetric information. That is, $C^{AI} - G^{AI} + N\gamma log(G^{AI} + d_1^{AI} + d_2^{AI}) + \Pi^{AI} > C^{SI} - G^{SI} + N\gamma log(G^{SI} + d_1^{SI} + d_2^{SI}) + \Pi^{SI}$.

Proof: See Appendix.

Information asymmetry is usually a source of inefficiency and lowers social welfare. This result, however, demonstrates that when public good is involved and there already exists inefficiency due to free-riding, asymmetric information might counteract the free-riding problem and eventually lead to higher social welfare. There are several channels through which asymmetric information enhances welfare here. First of all, when firm type is unclear, and consumers value the high type a lot more than the low type, then the low type would like to disguise as a high type, while a high type would not want to be mistaken as a low type. So the high type would go an extra mile to separate itself from the low type by choosing a higher donation level that the low type would not want to adopt. As a consequence, the difference between high type and low type's donations is higher. As mentioned before, this leads to a higher direct image effect in consumers' utility when purchasing from the high type firm and enhances consumer welfare.

In addition to the difference, the sum of the two firms' donations also rises due to firm 1's overdonating. A higher corporate donation certainly crowds out consumers' voluntary donation. But this is not necessarily bad for consumers, since they can allocate the spared income towards consumption of the numeraire good as they free-ride on firms' donations. This alone leads to higher consumer welfare. However, more interesting is the scenario when the crowding out is not one to one. This happens when firms' donations are high enough, consumers decrease their voluntary donations until they have hit the nonnegativity constraint and could not further decrease g_i even though they would like to. That is, consumers might not regard it as individually optimal, because they desire to free-ride even more, but less free-riding means higher social efficiency, and benefits consumers in the end. This binding situation of the nonnegativity constraint allows for higher public good provision, and even higher consumer welfare. Even though firms lose profit due to asymmetric information, the inefficiency on the firm side is offset by the gain in efficiency due to higher public good provision on the consumer side. Hence, social welfare is improved.

In the next part, I use a specific numerical example to provide some straightforward illustration of the above result .

2.4.1 Numerical Example

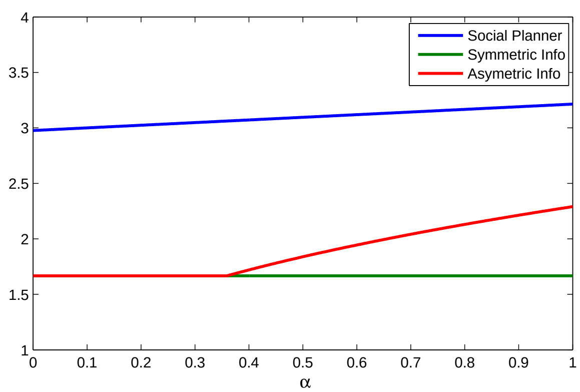
In the numerical example, I set the parameters to be specific values and graph how donations and welfare varies under the cases of social planner, symmetric information, and asymmetric information. I also specify the cost function of firms to be a simple quadratic form

$$k(d_i, r_i) = (K - (r_i - d_i))^2$$

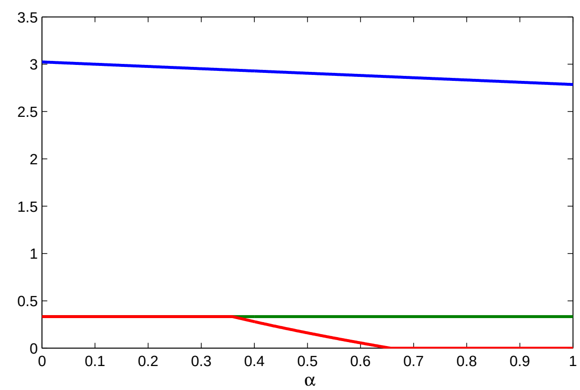
where K is a constant larger than r_1 and r_2 to ensure that $k_{d_i} > 0$ for all $d_i > 0$.The parameter values are as follows: $l = 10$, $\beta = 0.8$, $r_1 = 5$, $r_2 = 4$, $A = 5$, $\gamma = 2$, $N = 3$. I let α to vary between 0 and 1. I vary α because it represents the benefit of a good firm separating itself from a bad firm. Therefore, it is interesting to see how the benefit of signaling affects the welfare difference between symmetric and asymmetric information. Figure 2.1 shows the numerical results. The horizontal axis is α and the vertical axis represents different outcome variables of interest. Figure 2.1a plots the total corporate donations. Social planner's choice of $d_1 + d_2$ is the highest, because the planner takes into account the social benefit of firms' contribution to public goods. Firms, on the other hand, only care about profits and use donation to attract consumers, so the firms' decisions lie below the planner's choice, regardless of symmetric or asymmetric information. But when information is asymmetric and α is large enough, firms donate more in total due to firm 1's overdonation in the separating equilibrium. Exactly due to this increase in firm donation under asymmetric information, consumers' total voluntary donation decreases, as illustrated in Figure 2.1b. Yet more interesting is when firm donations are high enough, consumers' donations go to zero and is bounded by the nonnegativity constraint. That is, consumers would like to free-ride more on firms' donations by donating negative amounts, but are not allowed to. So in a sense, free-riding is mitigated to some extent in the asymmetric case when α is high enough. But the social inefficiency is still evident as the planner's choice for G is still above the individual decisions.

Figure 2.1: Parameter Values Supporting the Separating Equilibrium
Notes: This figure presents the numerical example of how donations and welfare vary under the cases of social planner, symmetric information, and asymmetric information. The x-axis is α, the responsibility parameter in each graph. The y-axis is firm public goods provision in Figure (a), individual public goods contribution in Figure (b), firm profit in Figure (c), and total welfare in Figure (d). Model parameters are as follows: $l = 10$, $\beta = 0.8$, $e_1 = 5$, $e_2 = 4$, $A = 5$, $\alpha = 2$, $N = 3$. See Section 2 and Appendix for the calculation of each outcome variable.

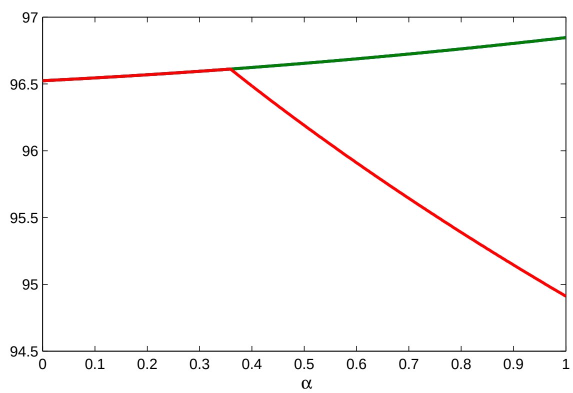
(a) Firm Public Goods Provision ($d_1 + d_2$)



(b) Individual Public Goods Contribution (G)



(c) Firm Profit (Π)



(d) Total Welfare

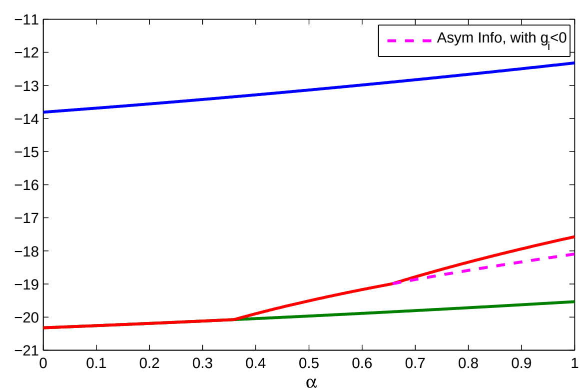


Figure 2.1c plots Π, which is firms' total profit. It shows that firms lose some profits when information is asymmetric. Figure 2.1d corresponds to the result in

Proposition 2. The total social welfare is higher under asymmetric information than symmetric, and the relation is strict as long as α is large enough. And at the point of α where consumers hit the nonnegativity constraint, actual social welfare takes off from the original trend which is denoted in the dash line. With an α this high, firm 1 takes the extra step in donation level to separate from firm 2, leading to a high total level of firm donations. The firm donations are so high that consumers are unable to free-ride as much as they desire because of the nonnegativity constraint. And the mitigated free-riding eventually leads to an even higher social welfare. This mitigation effect is stronger when there is more consumers in the economy.

2.4.2 Model Conclusion

When information is perfect, a profit maximizing firm might find it optimal to donate to public goods if donations affect company image and attract customers. When there is information asymmetry, firms have an additional reason to over-contribute in public goods. But the overdonation is not necessarily a bad thing. The welfare results suggest that in a model where there already exists inefficiency due to externality of public goods, adding a second layer of inefficiency, namely information asymmetry, might not distort the economy further, but could potentially counteract the inefficiency created by the public good provision and eventually improve social welfare.

2.5 Empirical Tests

In this section, I report results of two empirical tests are consistent with the hypothesis that the more consumers care about company image, regardless through direct or indirect signaling channels, the more firms would donate to charities. It is difficult to test directly the effect of firm donations on company image. So instead, I construct a "closeness" variable that tells whether a firm operates in an industry that deals with consumers directly, and test whether a "closer" firm would donate more to public goods. The underlying assumption is that the "closer" a firm is to the consumers, the more noticeable and relevant its donations to the consumers, or, the more image matters to both the consumers and the firm. In the second test, I use a fractional multinomial logit regression aiming to see whether there is any relation between a firm's industry and the charitable cause of the donations it makes.

2.5.1 Data and Variables

The sample firms are the S&P 500 firms in year 2014. I use the 2014 list to search for donations made by the firms from 1998 to 2014. The dataset is an unbalanced panel because some firms do not donate to charities in some years. The donation data is collected from NOZAsearch, a charitable donation database. NOZA's database is comprised of detailed charitable donation information collected from public available internet locations. As of 2015, it contains more than 100 million donation records. When a firm's name is searched, the database will list out the donations made by the firm, with details such as donation amount (or range), donation year, recipient of the donation, cause of the donation, and so on. I aggregate firm donations at year level and create a variable **donation**, which is the total amount of donations in millions made by a firm in a year.[8] NOZAsearch also reports whether donations are made to charitable programs operating on a local or national scale, which the variables **local donation** and **national donation** take into account. Very few donations are made on an international scale.

The "closeness" variable is constructed using the 2007 benchmark input-output data provided by the Bureau of Economic Analysis. It is called the **personal con-**

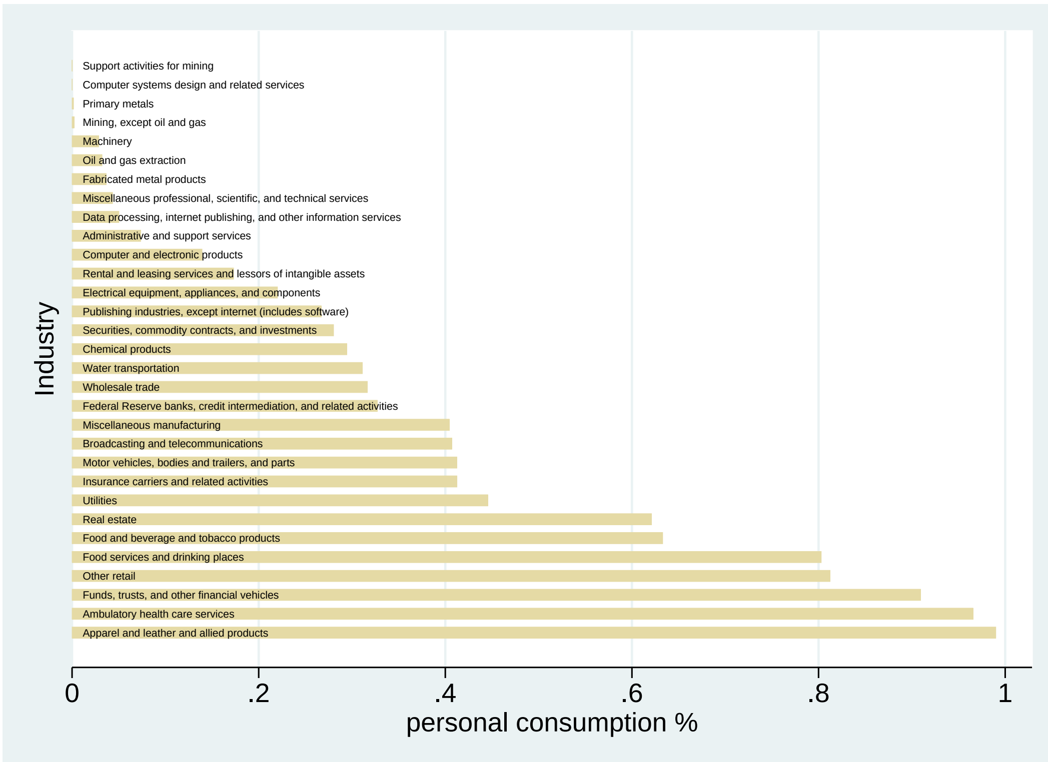
[8]All dollar amounts are stated in 2000 dollars deflated/inflated using CPI.

sumption percentage (PCP), which is the percentage of an industry's total revenue that is attributed to direct consumer purchases. The detailed construction of PCP follows the procedure used by Ahern and Harford (2014) and is detailed in the Appendix. Take the telephone apparatus manufacturing industry for an example. The make table of BEA's input-output data tells that the industry is the primary producer of telephones, and secondary producer of other commodities such as wireless equipments. For each of these commodities, I can compute the market share of the telephone apparatus manufacturing industry in the commodity market from the make table. Meanwhile, for each commodity, the use table lists all the industries that consume the commodity and it can be calculated how much revenue a certain industry contributes to each commodity market. The use table includes an industry called "personal consumption expenditures", which entails the common consumers. So using both the make and use tables, I obtain the share of revenue consumers contribute to a certain industry, such as the telephone apparatus manufacturing industry. I call this the personal consumption percentage and use it as a proxy for "closeness" to consumers. The BEA uses its own industry classification, but also provides conversion instructions into NAICS industry codes. One might be concerned that the industry classification is broad, but the NAICS six-digit code in fact provides a rather refined classification. Using the six-digit classification, I obtain 193 industries for the sample firms, and the PCP variable takes 127 different values. In Figure 2.2 , I use the more aggregated, three-digit NAICS code to show the mean PCP values for some aggregated industries. One can see that for an industry like mining, PCP is close to zero, while the apparel industry has a PCP value close to one.

Table 2.1 reports the summary statistics. The first part include the donation-related variables. On average, a sample firm donates 4 million of 2000 dollars to charities in one year. The average local donation is 1.6 million, and average national

Figure 2.2: Output Purchased by Final Consumers

Notes: This figure presents personal consumption percentage (PCP) in different industries. PCP is calculated based the 2007 benchmark input-output data provided by the Bureau of Economic Analysis. PCP measures the percentage of revenue that consumers contribute to a certain industry. PCP is calculated for each BEA detailed industry (389 industries). For this figure, PCP is aggregated to BEA summary industry (71 industries) based on weighted average of each detailed industry with industry total output as weight. BEA summary industries with number of 2007 Compustat firms below sample median (65 firms) are not reported in this figure. See Section 4 and appendix for detailed PCP construction procedures.



donation is 2.2 million. Statistics of five main cause categories are also reported. Education is the largest category with a mean firm-year donation of 2 million dollars, and the second largest category is health with 1 million dollars. Firm characteristics are reported in the second part of the table. The average sample firm has total assets of 51.6 billion dollars and net income of 1.6 billion dollars. This is consistent with the fact that these are large S&P 500 firms. Finally, regarding the "closeness" measure, the average personal consumption percentage is 0.35.

Table 2.1: Summary Statistics

Notes: The table reports the summary statistics for a sample of S&P 500 firms with available data from 1994 to 2014. Donation is the total amount of a firm's donation in a year collected from the NOZAsearch database. The summary statistics excludes firm-year observations where there is no donation reported from NOZAsearch of that firm year. *Local donation* and *national donation* are the amount of firm's donation to local organizations and national organizations, respectively. *Food donation* refers to donation where recipient's cause is food and agriculture. *Environment donation* refers to donation where recipient's cause is environment or animal-related. *Education donation* refers to donation where recipient's cause is education or youth development. *Science donation* refers to donation where recipient's cause is science and technology. Health donation refers to donation where recipient's cause is health care, medical research, or mental health & crisis intervention. All donation amounts are reported in millions of 2000 U.S. dollars. Firm characteristics are obtained from Compustat. *Total assets* and *net incomes* are in billions of 2000 U.S. dollars. *Return on equity* is the ratio of net incomes to equity. *Return on assets* is the ratio of net incomes to total assets. *Personal consumption %* calculated from BEA input-output matrix measures the percentage of revenue that consumers contribute to the firm's industry.

	mean	p50	sd	p25	p75
Donation ($M)					
Donation	4.08	1.13	9.12	0.32	3.76
Local donation	1.58	0.27	4.46	0.00	1.20
National donation	2.24	0.50	5.79	0.09	1.94
Food donation	0.08	0.00	0.74	0.00	0.00
Environment donation	0.14	0.00	0.94	0.00	0.00
Education donation	2.05	0.45	4.79	0.00	1.94
Science donation	0.13	0.00	0.81	0.00	0.00
Health donation	1.02	0.00	4.45	0.00	0.42
Firm					
Total assets ($B)	51.59	15.28	128.46	6.06	36.16
Net incomes ($B)	1.57	0.68	2.72	0.28	1.61
Personal consumption %	0.35	0.32	0.33	0.03	0.54
Observations	3,734				

2.5.2 Test 1

The regression results of the first test is reported in Table 2.2. The dependent variable is donation, and the key independent variable is PCP. Firm characteristics and industry characteristics are added as additional controls. I also add time fixed effects and state fixed effects to absorb any time or location trend. [9] In the first column, I include all donations made by firm-years, and this gives me a sample size of 3,734. The key variable PCP is significant at 10% level, weakly supporting my "image gain through closeness" story. The second column, however, only includes those

[9] I do not add industry fixed effects because the key variable PCP is at industry level.

donations made on a local scale, and the third only national ones. The differentiation between local and national donations is based on the conjecture that donations are more observable by consumers on a local level, so firms might have higher incentive to donate on a local scale to attract local consumers using a good local image. Indeed, the variable PCP is significant at 5% level when only local donations are considered. The coefficient is 0.535, which can be translated into an increase of 0.176 million dollars of donations, when personal consumption percentage goes up by a standard deviation (0.33). This increase is about 11% of the mean local donations as reported in Table 2.1. However, for national donations, PCP is insignificant as shown in column 3. This result supports the conjecture that a firm responds more to "closeness" with local donations instead of national ones due to observability of local donations by consumers.

In column 4, I include not only firm-years with positive donation amount, but also those firms that donated a positive amount in at least one, but not necessarily all years. As a results, I get a larger sample size. For a firm that donated in some years but not others, there is the concern that the zero-donations are not actually zeros, but rather missing entries in the NOZAsearch database. As a result, I use column 1 with positive donations only as the benchmark regression and add column 4 as a robustness check. The coefficients in the two columns do not appear to be significantly different. And for column 5, I include all the S&P 500 firms, regardless of whether they make donations at all. One can interpret the first column as focusing on the intensive margin of donations given that a firm donates in a given year, while the last column is more about the extensive margin. In the last column, the variable PCP is not significant when the extensive margin is also considered.

The control variable size, which is the natural log of total assets, is significant at 1% level in all five specifications. This is consistent with the intuition that larger

Table 2.2: Closeness and Donation

Notes: The table presents OLS regression of firm donation on industry personal consumption percentage. The sample consists of S&P 500 firms with available data from 1994 to 2014. The dependent variable in columns (1), (4) and (5) is firm's total donation in a year. The dependent variable in columns (2) and (3) is local donation and national donation, respectively. Columns (1) to (3) only include firm-year observations where firm's total donation in that year is not zero. Column (4) further includes firm-years where firm's donation in that year is zero, but the firm has denoted at least once during the entire sample period. Column (5) includes all firm-years. *Personal consumption %* is the percentage of revenue that consumers contribute to the firm's industry. *Size* is the logarithm of firm's total assets. *Net incomes* is the logarithm of firm's net income. *Industry size* is the logarithm of the total assets of the NAICS industry which the firm belongs to. *Industry Herfindahl* is the Herfindahl index based on sales. *Commercial bank* is a dummy variable which equals 1 if the firm is a commercial bank. Year and state fixed effects are included in all columns. Standard errors reported in the parentheses are robust and two-way clustered by industry and year. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Donation (1)	Local Donation (2)	National Donation (3)	Donation (4)	Donation (5)
Personal consumption %	1.9978*	0.5708**	0.9742	1.3789*	0.9893
	(1.0811)	(0.2730)	(0.6655)	(0.8108)	(0.6585)
Size	1.4900***	0.5902***	0.7611***	0.9085***	0.7405***
	(0.3297)	(0.1616)	(0.1559)	(0.2076)	(0.1769)
Net incomes	0.2143*	0.0494*	0.1192*	0.1759**	0.1552**
	(0.1109)	(0.0299)	(0.0693)	(0.0866)	(0.0768)
Industry size	-0.0780	0.0205	-0.0790	0.1053	0.1296
	(0.2335)	(0.1090)	(0.1177)	(0.1328)	(0.1125)
Industry Herfindahl	0.1407	0.4115	-0.0843	0.6679	1.0146
	(1.2470)	(0.4344)	(0.7571)	(0.8310)	(0.6988)
Commercial Bank	-0.2198	0.7270***	-0.8450	-0.0359	0.0834
	(0.8281)	(0.2695)	(0.5363)	(0.5428)	(0.6068)
Time fixed effects	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓
Observations	3,734	3,734	3,734	5,874	7,123
Adjusted R^2	0.2313	0.2278	0.1835	0.2029	0.1750

firms should donate more to charities. The indicator variable commercial bank is significant at 1% when only local donations are used. In all other specifications, commercial bank is insignificant. This confirms the fact that banks are required to support their local communities. The results in the "closeness" regression provides some evidence that firms care about their public image perceived by consumers, as firms that are "closer" to consumers donate more to public goods, especially reflected in local donations that are more visible.

2.5.3 Test 2

In the second test, I use the donation cause information provided by NOZAsearch to test to see whether firms tend to donate to causes that are related to their industries. I divide all donations into six cause categories, food, environment, education, science, health, and others. A fractional multinomial logit model is used to characterize firm's decision of what percentage of donation it puts into each category. The others category is omitted in the regression. Regarding firm characteristics, I consider three industry dummies, food industry, high-tech industry, and medical industry, as these are the industries that have clear connection to some of the cause categories. Since explanatory variables are time invariant, I also construct donation percentage at firm level rather than firm-year level.[10] The sample I arrive at has 405 firms.

Table 2.3 Panel A reports the regression coefficients. For firms in the food industry, the coefficient for the food category is significant at 1% level and has a magnitude larger than the coefficients for all remaining categories. This means that food companies are significantly more likely to donate to the food cause.[11] For high-tech firms, the coefficient for the science category is highly significant and has a magnitude larger than other coefficients. So high-tech firms has a higher tendency to donate to the science cause among all causes. The coefficient for education is also significant at 1%, but the magnitude is not the largest among all coefficients. This can be interpreted as high-tech firms donates more to the education cause than at least the omitted "others" category. And for medical firms, the coefficient for the health category is

[10]Using firm-year donation as dependent variable and clustering standard errors at firm-year level yield the same results.

[11]For multinomial logit models, an increase in explanatory variable x leads to an increase in dependent variable category y_i if $\beta_{x,i}$ is larger than $\beta_{x,j}$ for any $j \neq i$.

highly significant and largest in magnitude, so medical firms has a tendency to donate to the health cause among all causes.[12]

In Panel B of Table 2.3, I report the marginal effects translated from the regression coefficients. The last line in the panel means that an average firm would allocate 2% of its donations to the food cause, 3% to environment, 41% to education, 2% to science, 18% to health, and 34% to other categories. The first marginal effect of 0.0621 means that a food company on average allocates 8.21% of its total donations to the food cause as compared to 2% of an average firm. For high-tech firms, the coefficients in Panel A are significant for both education and science causes. Their marginal effects show that high-tech firms donate 49.25% to education and 4.51% to science, significantly higher than an average firm. And for medical firms, they donate significantly more to the health category–46.94% as compared to 18% of an average firm.

The fractional multinomial logit results suggest that there exists some connection between the firm's line of business and the cause they like to donate to. The underlying reason might be an image one. For example, people who consume a lot of medical products are more likely to be people with health problems. These are also likely to be the people who pay more attention to medical research, health care issues, and donations made for these purposes. So the image effect would also be larger for these consumers. Hence, it makes sense for medical firms to donate more to the health cause in order to attract customers. However, there could also be other reasons causing this donation pattern. For example, a high-tech firm that donates more to education and science categories might see the donations as long-term investments and expect future gains through higher human capital and knowledge spillovers. If

[12]The health cause also include medical researches. For more detailed description of the cause categories, please refer to the notes of Table 2.1.

this is true, then there seems to exist some complementarity between a firm's line of business and donations that might increase firm profits in the long term. This kind of complementarity is not explored in this paper, but could be an interesting direction for future research on corporate giving.

2.6 Conclusion

Charitable giving, due to its visibility and measurability, does not only enhance a firm's social image directly, but also conveys a credible message of the firm's overall level of social responsibility, which might be hidden from or hard to measure for the consumers. Consumers, even though self-interested, would still prefer to purchase products from firms that seem "nicer." In this case, it is profit-maximizing for firms to donate money to charities in order to attract consumers. In a separating equilibrium, a high responsibility firm donates not only more than the low responsibility firm, but also more than the case when information is symmetric. This indicates that asymmetric information is a source of inefficiency for the firms. But the overdonation of the high type leads to a higher level of total corporate donation, resulting in higher consumer welfare. And if consumers care a lot about corporate social responsibility, corporate donation will be really high, and this will mitigate the free-riding problem of the consumers, leading to even higher consumer welfare. So in a sense, information asymmetry, a source of inefficiency for the firms, counteracts the inefficiency with public good, and leads to a higher level of social welfare.

The channels to attract consumers should only work for firms that sell directly to consumers. Hence, I use donation data to show that firms in final consumption good industries donate more than firms in intermediate good industries. However, as mentioned, there must exist other potential reasons for profit-driven firms to donate to charities, such as complementarity between the business and a charitable cause.

If this is the case, then the outcome on public good provision and consumer welfare might be positive as well, and could be interesting to explore.

Table 2.3: Industry and Donation Category

Notes: The table presents the coefficients and marginal effects from a fractional multinomial logit model of firm's industry on firm's donation category. The sample consists of S&P 500 firms with available data. Firm's donations are classified into six categories: food, environment, education, science, health, and others. See Table 2.1 for category definition. The dependent variable is the fraction of a firm's donation in each of these six categories during the entire sample period. The independent variables are three dummies. *Food* equals 1 if firm's NAICS code is in 11, 311, 312, 445, or 722. *High-tech* equals 1 if firm's NAICS code is in 32411, 325, 33299, 3331-3333, 3336, 3339, 3353, 33599, 3361-3364, 3391, 334, 5112, 518, 519, 5413, 5415-5417, or 8812. *Medical* equals 1 if firm's NAICS code is in 3254, 3391, or 621-624. Panel A reports the coefficients of the fractional multinomial logit model. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively. Panel B reports the change in donation fractions if a firm changes from not in food, high-tech and medical industry to food industry, high-tech industry or medical industry, respectively. The last row of Panel B reports the average donation fraction in each category for all sample firms.

Panel A: Fractional Multinomial Logit Model

	Donation Categories (Other Category Omitted)				
	Food (1)	Environment (2)	Education (3)	Science (4)	Health (5)
Food	2.0662*** (0.4394)	0.7759* (0.4263)	0.5207* (0.2662)	0.0983 (0.6168)	0.3939 (0.3347)
High-tech	0.1310 (0.7156)	0.2488 (0.3101)	0.4166*** (0.1616)	1.7060*** (0.3890)	0.0359 (0.2292)
Medical	-1.6577** (0.8236)	0.7888 (0.8106)	0.2769 (0.4288)	-1.5276** (0.5931)	1.5592*** (0.4382)
Constant	-3.2983*** (0.2648)	-2.5723*** (0.1697)	-0.0132 (0.0996)	-3.9674*** (0.3203)	-0.7838*** (0.1149)
Observations	405				
Model χ^2	138.00				

Panel B: Marginal Effects

	Food (1)	Environment (2)	Education (3)	Science (4)	Health (5)	Other (6)
Food	0.0621	0.0138	0.0521	-0.0019	0.0000	-0.1262
High-tech	-0.0012	0.0001	0.0825	0.0251	-0.0300	-0.0771
Medical	-0.0128	0.0065	-0.1033	-0.0065	0.2894	-0.1733
Category average	0.02	0.03	0.41	0.02	0.18	0.34

Chapter 3

CRIME RATE, HOUSING PRICE, AND VALUE OF A STATISTICAL CASE OF HOMICIDE

3.1 Introduction

Crime could be a life-threatening issue and it is hard for individuals to fight crime. As a result, societies usually rely on the government to provide public safety as a local public good. However, the amount of money a local government should spend on reducing crime is difficult to measure. The most challenging task in the cost and benefit analysis is to estimate people's willingness to pay for a safer neighborhood. The revealed preference method and the property hedonic model are commonly used to estimate the marginal willingness to pay for a change in the level of local public goods, such as environmental quality (e.g., Kiel and McClain, 1995; Leggett and Bockstael, 2000; Chay and Greenstone, 2005). These methods can similarly be used in the study of crime. In this paper, we use the hedonic approach to explore the relationship between changes in crime rate during the 1990s and changes in housing price during that period. We then translate the estimates into people's willingness to pay for a reduction in violent crime.

During the 1990s, crime rate dropped sharply and unexpectedly among all categories of crime and across all parts of the United States (Levitt, 2004, Pope and Pope, 2012). According to data by the Federal Bureau of Investigation, from 1991 to 2000, the violent crime rate plunged by 33% percent and the property crime rate by 30%. Similar trends can also be observed in any other categories of the crime. These declines were largely unexpected and experts actually predicted an explosion in crime rate for the 1990s (Levitt, 2004). Although it must interest researchers a lot to look for the causes of the decline, we believe it is also worthwhile to look at the effect side of it.

Some early papers study the relationship between crime rate and housing price using cross sectional data (Thaler, 1978; Lynch and Rasmussen, 2001; Gibbons, 2004). More recently, researchers exploit temporal changes in crime rate and use panel data with fixed effects to control for unobservable time-invariant factors. For example, Linden and Rockoff (2008) and Pope (2008) study how move-ins of a sex offender affect the housing prices in Montgomery County, Ohio and Hillsborough County, Florida, respectively. Both papers find a negative relation between crime risk and property values, but the results pertain to the specific counties and to the one particular type of crime. In a more recent study by Pope and Pope (2012), they use data from 3,000 zip codes in five states and find the elasticity of property value with respect to crime to range from -0.15 to -0.35 by exploring the 1990s crime drop. Our paper shares some features with Pope and Pope (2012), in the sense that we also use the 1990s crime drop to find the effects on housing price. However, we believe that our study also brings some new features to the existing literature from at least the three following aspects.

First, we use a more comprehensive measure of crime as compared to earlier studies and consider a different instrumental variable from the one used in Pope and Pope (2012). The instrument we use is the state level abortion rate in the 1970s and 1980s as an instrument for MSA level crime rates in the 1990s. The validity of the instrument is based on the 10- to 20-year gap between the abortion rate and the housing price. The relevance of our instrument is proposed by Donohue and Levitt (2001), in which they argue that the unexpected crime rate drop in 1990s was largely driven by the legalization of abortion in 1970s. We find a large F-statstics from the first stage regression, which suggests that the instrument is strong judged by the criteria suggested in Bound *et al.* (1995).

Second, in almost all empirical papers in the hedonic literature, the hedonic function is assumed to be stable. But such assumption does not necessarily hold. Kuminoff and Pope (2014) point out that as the level of a public good changes (absence of crime in our case), the gradient of the hedonic function should also change. They argue that if one ignores the evolution of the coefficients, the estimate of the traditional "capitalization" effect does not necessarily say anything about people's true marginal willingness to pay. In our paper, we explicitly take into account the evolution of the coefficients by estimating a fixed effects model with time-varying coefficients.

Third, we obtain an estimate for the value of a statistical case of homicide, which is an estimate of people's willingness to pay to avoid one expected case of homicide. Since our crime data contain seven categories: murder and nonnegligent manslaughter, forcible rape, robbery, aggravate assault, burglary, larceny theft, and motor theft, we translate the marginal willingness to pay for fewer cases of murder and nonnegligent manslaughter into a value of a statistical case of homicide. [1] To do so, we first create a crime index based on the method proposed by Sellin and Wolfgang (1964), then use the index as the crime variable in the regression, and finally translate the estimated coefficient into a value of a statistical case of homicide. Based on our preferred model specification, we find that people's willingness to pay for avoiding a case of homicide is about 0.4 million dollars.

[1] Technically, homicide contains not only murder and nonnegligent manslaughter, but also other cases such as accident killing of a man. However, to simplify language, we use the term "value of a statistical case of homicide".

3.2 Data

3.2.1 Data Sources

We use data from four different sources. We collect the annual MSA level crime rate in all categories from Uniform Crime Report (UCR) published by the FBI. We obtain the annual MSA level housing data from the Housing Price Index (HPI) published by the Federal Housing Financing Agency (FHFA). Other annual MSA level demographic characteristics such as population and income are gathered and derived from IPUMS CPS. Finally, we obtain the state level abortion rate in 1970s and 1980s from the Guttmacher Institute.

We collect the annual MSA level crime rate from 1992 to 2000 from FBI's Uniform Crime Report (UCR). The years 1990 and 1991 are excluded from the regression because the crime rate was still increasing in the year 1990 following the trend in the 1980s, and also because it might take some time for people to realize the decrease and update their beliefs in the crime rate. The FBI gather and release the crime data based on reports from local law enforcement agencies through UCR annually. The data is an unbalanced panel because there is a deadline for the local agencies to report their annual crime data and some agencies miss the deadline from year to year. FBI divides crime into seven categories: murder and nonnegligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny-theft and motor vehicle theft. The first four types of crime are categorized into violent crimes while the last three are considered as property crimes. Even though we attempt to derive a value of a statistical case of homicide, we do not want to consider murder as the only type of crime in our regression, because not controlling for other types of crime could lead to omitted variable bias and inflate the coefficient on murder. Neither do we want to include all categories as separate variables since most types of crimes are highly correlated. So we compose all categories into one crime variable. However, not all categories of crime should be given the same weight, because, for example, murder is obviously more serious than theft. Therefore, we adopt the methodology suggested by Sellin and Wolfgang (1964) and calculate a weighted crime index based on the severity of the offenses. According to that method, a one unit increase in the homicide rate (per 100,000 people) increases the index by 26 units, which is approximately 13 times bigger than the contribution of a one unit increase in larceny theft.

Sellin and Wolfgang published *The Measurement of Delinquency* in 1964, in which they develop a method to measure the seriousness of different types of crimes. The method allowed researchers to better understand the "qualitative elements in criminal behavior", according to Wellford and Wiatrowski (1975) in a work following Sellin and Wolfgang. In their original work, Sellin and Wolfgang made a list of 141 cases of crime and asked various judges such as university students to rate the seriousness of all the cases. From the rating of the judges, Sellin and Wolfgang created the scale of offense seriousness that is adopted in this paper. The seven categories of crime are given the following weights: murder and nonnegligent manslaughter (26.4), forcible rape (14.7), robbery (4.6), aggravated assault (5), burglary (2.4), larceny-theft (2.1), and motor vehicle theft (3.1). The Sellin and Wolfgang scale is replicated by various studies in different parts of the world (e.g., Normandeau, 1966; Velez-Diaz and Megargee, 1970; Hsu, 1973), and the replication studies mostly prove the Sellin and Wolfgang scale to be reliable (Wellford and Wiatrowski, 1975).

The MSA level housing prices are recorded from Federal Housing Financing Agency's Housing Price Index and range from 1994 to 2002. The time scopes of the data on housing price and crime are slightly different because we assume there is a time lag between the decline in crime rate and its effect on housing price. The HPI is a broad measure of the price movement of single-family detached properties. Based on the data on conforming mortgage transactions obtained from the Federal Home Loan Mortgage Corporation (Freddie Mac) and the Federal National Mortgage Association (Fannie Mae), FHFA estimates and publishes quarterly average price changes in repeat sales or refinancings on the same properties. The estimation is based on a modified version of the weighted-repeat sales methodology proposed by Case and Shiller (1989). We use the MSA level data downloadable from the FHFA's website and convert the quarterly data into annual data by taking an unweighted average (the results are almost identical if we take a weighted average based on the standard deviation of the estimates). To convert the price index into dollar values and to calculate the value of a statistical case of homicide, we use the MSA level median housing prices from 2000 census provided by National Historical Geographic Information System.

We also use data on demographics as control variables. They include household income, age structure, education level, race composition, housing ownership, poverty and unemployment rate. Since the census only provides demographic data on 2000 and our estimation is based on annual MSA level data, we derive those information from Current Population Survey. We collect individual demographic characteristics from March CPS through IPUMS and then convert them into aggregated MSA level demographic information by using the weight suggested by the CPS.

Finally, we collect state level abortion rates in the 1970s and 1980s from the Guttmacher Institute. Abortion was legalized in the United States after 1973, so we are unable to find information on abortion rate before that year. To construct an instrument for a given MSA in a given year, we calculate an annual state level effective abortion rate based on the formula suggested in Donohue and Levitt (2001).

3.2.2 Descriptive Statistics

Figure 3.1 presents the trend in national crime index from 1992 to 2000. The index is calculated as the average of MSAs by using the Sellin and Wolfgang method. On average, crime declined dramatically over the period, with the index falling from the highest value of 15,581 in 1992 to the lowest value of 11,681 in 2000. If we only look at MSAs with crime index in 1992 above the 75th percentile of the sample, the plunge becomes even bigger. If we instead look at MSAs with low initial crime index, the drop is milder. A similar trend can also be observed if we instead use the unweighted index or any single category. These observations suggest that our data provide enough variations both across time and between MSAs in the key explanatory variable to identify the effect of crime on property value.

Figure 3.1: Crime Index from 1992 to 2000

Notes: The index is calculated as the average of MSAs in the sample and by using the Sellin and Wolfgang's weighting function. MSAs with high (low) initial crime rate are defined as MSAs whose crime index in 1992 is above (below) the 75th (25th) percentile of the sample.

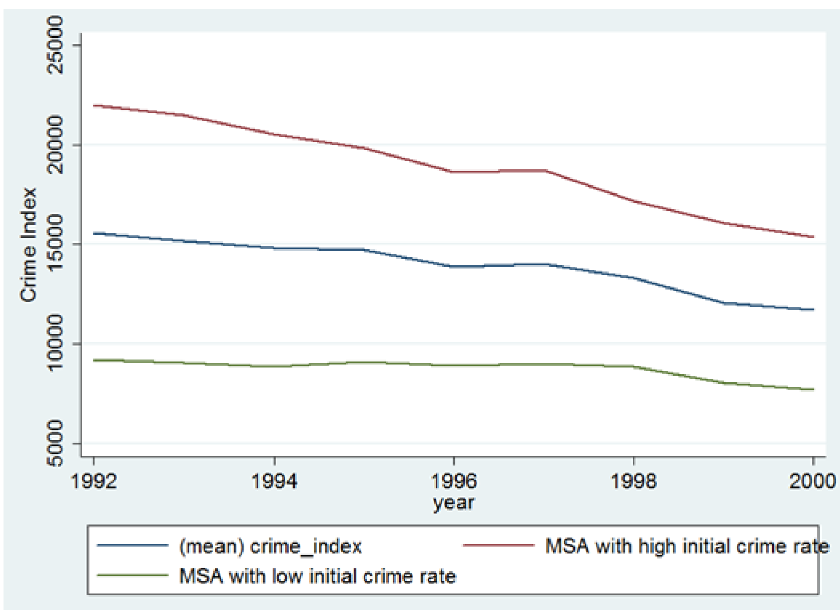


Table 3.1 presents summary statistics on variables that we use in the subsequent regressions. The means are calculated as the averages across MSAs used in the pri-

mary regressions. The monetary figures are denoted in 1999 dollars. During 1992 to 2000, the mean of MSA's housing price index (adjusted for inflation) increased roughly by 12%, whereas crime index decreased by 25%. Income per household rose approximately by 5%, unemployment rate decreased by 3%, and the population was 12% higher at the end of the period. The increase in education attainment is evident: 3% increase in high school attainment and 4% increase for college. The race composition, poverty rate and housing ownership are roughly constant at the beginning and the end of the period.

Table 3.1: Summary Statistics

Notes: All values in the table are averages across MSAs used in the primary regressions. For HPI, the first quarter of 1995 is set as the benchmark and is assigned the number of 100. The MSA HPIs in the table are adjusted by the inflation.

Mean	1992	2000
Housing price index	108.19	120.73
Crime index	15580.51	11680.97
Income per household (1999)	43983.4	46294.2
Population	642873	720947
Unemployment rate	0.08	0.05
% high school graduate	0.81	0.84
% college graduate	0.21	0.24
% white	0.86	0.85
% poverty	0.13	0.12
% owned house	0.68	0.69

3.3 Empirical Methodology

3.3.1 Fixed Effects Hedonic Model

We use hedonic regressions with MSA level panel data to estimate the effect of crime on housing price. The hedonic approach is originally developed as an individual-level model, but the aggregation to lower resolution level is common in the literature. (For instance, Chay and Greenstone (2005) consider the county level aggregation and use the hedonic model to estimate the effect of air quality on housing price.)

The model estimated is as follows:

$$p_{i,t} = \beta_t Crime_{i,t-1} + \gamma X_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{i,t} \ ,$$

where $p_{i,t}$ is the median housing price in MSA i at time t, $Crime_{i,t-1}$ is the key variable of interest, $X_{i,t-1}$ is the set of demographic control variables, α_i is the MSA fixed effect, λ_t is the national time dummy, and $\varepsilon_{i,t}$ the idiosyncratic error term. To reduce measurement error in crime and other demographic characteristics, we combine the data of two adjacent years to generate one time period. Hence, we have five periods in the panel. We assume that there is a one period lag between the realized crime rate and its effect on the housing price.

One difference between our estimation strategy and the traditional hedonic model is that we allow the coefficient β_t to be varying over time. The interpretation of the coefficients in the hedonic regression is people's marginal willingness to pay (MWTP) for the local public good (absence of crime in our case). However, MWTP does not necessarily stay constant as time goes on and as the level of public good provision changes. Therefore, if one ignores the evolution in the gradients, the hedonic model is not correctly specified and hence the estimate might not be the true MWTP (Kuminoff and Pope, 2014).

3.3.2 Instrumental Variable

In the above estimation equation, even though we include MSA fixed effects to absorb time constant unobservable effects, there might still exist some unobservable time varying factors that influence both crime and housing price. For instance, suppose a local government carries out a policy to provide housing subsidies for households with low income, then such policy might reduce the crime rate because low income families could potentially invest more time and money in their children and keep them away from violence and crime. At the same time, housing subsidies could also have an impact on the local housing market through raising demand. If this is the case and we do not include the policy as an explanatory variable, the OLS estimation will give an upward biased result. Therefore, we find it necessary to use an instrumental variable for crime to address the potential endogeneity issue.

We are not the first to propose the usage of an instrument. In the study by Pope and Pope (2012), they use a zip code with similar initial crime rate as an instrument for crime in a target zip code. However, we think the validity of such instrument might be questionable. Suppose one uses a zip code in California as the instrument for a zip code in New York because the two zip codes have the same initial crime level, and a national policy was carried out to improve school quality in these areas. Since school quality might affect young people's tendency to commit crime, the change in crime rate in the California zip code could be related to the increase in school quality in the New York zip code and have an effect on its housing price. Hence, if such national policies are not controlled, the validity of the instrument is violated.

The instrument we implement in this paper is the state level abortion rate in the 1970s and 1980s, based on Donohue and Levitt (2001). The U.S. Supreme Court's Roe v. Wade decision in 1973 announced the legalization of abortion. After 1973, there was a significant increase in abortion rates in almost all states across the country. Such trend continued until the beginning of 1980s and was then followed by a steady decrease throughout the 1980s and 1990s.

We argue for the validity of the instrument based on the time lag between the abortion data and the housing data. The abortion data we use are from the 1970s and 1980s and the housing price data are from the 1990s, so there is at least a 10 year gap between the two. The argument for the relevance of the instrument is based on two premises: unwanted children have a higher risk for crime and legalization of abortion reduces the number of unwanted birth. Such argument is proposed and verified in Donohue and Levitt (2001). They find a statistically significant negative relation between abortion rates in the 1970s and early 1980s and crime rates in the 1990s. However, Christopher L. Foote and Christopher F. Goetz point out that there is a coding error in the last model specification in Donohue and Levitt (2001) and question the endogeneity issue in all their regressions. But we think these issues do not affect our study for two reasons. First, Donohue and Levitt (2001) use five different approaches, all of which point out the same relation between abortion and crime. Foote and Goetz only found an error in the last approach, while our IV is mainly based on the fourth approach. Secondly, even if Donohue and Levitt (2001) overlook other factors that are correlated with abortion in 1970s and are determinants of crime rates in the 1990s, the relevance of our instrument only depends on correlation rather than causation, so we think it is not a big issue, either. We also look at the F-statistics from the first stage regression, and the results show that the abortion rate is a strong instrument for the crime.

To generate the instrument for crime rate at a given year t, we adopt the idea of the "effective abortion rate" in Donohue and Levitt (2001), which is the average of abortion rate across all cohorts of arrestees weighted by the cohort's share in the population of arrestees, i.e., $Eff\_Abortion_t = \sum_a Abortion_{t-a}(Arrests_a/Arrests_{total})$. The youngest group we consider is the cohort at age of 10; the oldest is the one at age of 19. We only consider people of younger age because we want to exclude the main group of home buyers. And excluding adult arrestees does not undermine the strength of the IV. To illustrate the construction of the effective abortion rate, taking year 1992 as an example, we consider the cohorts with age of 10 to 19 in that year, and correspondingly use the abortion rate from year 1973 to year 1982. [2] We use the three year national average arrest data in 1981 to 1983 from the Uniform Crime Reports to compute the weighting function and use it for all states. We then use the effective abortion rate as the instrument for the crime index and estimate the following equation:

$$p_{i,t} = \beta_t Crime_{i,t-1}(Eff\_Abortion_t) + \gamma X_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{i,t} \,.$$

The results from this equation are our preferred estimates.

3.4 Results

3.4.1 Estimation Results

Table 3.2 presents the OLS estimation results of the fixed effects hedonic model without the instrumental variable. Column 1 and column 3 only include the crime index and time dummies; column 2 and column 4 add demographic control variables. Estimates in columns 1 and 2 are the conventional capitalization effects, i.e. assuming the coefficients are stable; columns 3 and 4 allow the coefficient of the crime index in periods 4 and 5 to be different from that in periods 1 to 3. These four model specifications are also used in Table 3.4. There are two points to mention about our choice of the time-varying coefficients. First, ideally, we could generate interaction terms between the time dummy and the crime rate and estimate the coefficient on crime for all 5 periods separately; however, by doing so we would create too many interaction terms, which deteriorates the identification power of the model. Since the crime rate decreases in a stable trend throughout our data period, we think it is a natural way to cut the total time span in half. Second, we do not allow the coefficients

[2] The arrests are also divided into seven categories, and we use the same method to create the effective abortion rate index as used in weighted crime index.

of other control variables to change over time due to the same consideration for identification power. Our reasoning behind this is quite similar to what researchers usually do with regard to the endogeneity issue. Although endogeneity of any variable could lead to biased estimates for all coefficients, in reality, people often only address the endogeneity of the key variable of interest. Similarly, although varying coefficients on other variables could lead to biased coefficient on crime, we only regard it as a second order issue.

Table 3.2: Estimates of the Impact of Crime on Housing Price by Fixed Effects Hedonic Model

Notes: The control variables used in columns (2) and (4) include all demographic controls mentioned above in equation (1). Here, we only present the coefficients on population size and household income because these two coefficients are statistically significant. For columns (3) and (4), we assume different coefficients on crime for period 1 to 3 and period 4 to 5. The last row shows the difference between the coefficients for different periods. The numbers in the parenthesis are cluster robust standard errors. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Crime index ($\times 10^5$)	-0.8352***	-0.6083***		
	(0.2203)	(0.2126)		
Crime index (period 1) ($\times 10^5$)			-0.8517***	-0.6079***
			(0.2205)	(0.2143)
Crime index (period 2) ($\times 10^5$)			-1.1314***	-0.9458***
			(0.2589)	(0.2496)
Household income ($\times 10^5$)		0.2127***		0.2085***
		(0.0587)		(0.0577)
Population ($\times 10^5$)		0.0158***		0.0173***
		(0.0051)		(0.0049)
Unemployment rate		-0.0747		-0.0606
		(0.1131)		(0.1109)
Poverty %		-0.0245		-0.0476
		(0.0706)		(0.0716)
Old people %		-0.0335		-0.0268
		(0.0678)		(0.0679)
High school graduate %		-0.2081***		-0.2118***
		(0.0702)		(0.0689)
College graduate %		0.0565		0.0573
		(0.0498)		(0.0491)
White %		-0.0402		-0.0389
		(0.0531)		(0.0528)
Owned house %		-0.0192		-0.0195
		(0.0446)		(0.0445)
Other demographic controls	No	Yes	No	Yes
Time dummies	Yes	Yes	Yes	Yes
MSA dummies	Yes	Yes	Yes	Yes
Observations	1,023	1,023	1,023	1,023
Adjusted R^2	0.843	0.853	0.845	0.855
Difference of coefficients on crime index			0.280**	0.338***
			(0.112)	(0.103)

The results in column 1 show that if the crime index decreases by 1 unit, the housing price will increase by $8.35 \times 10^{-6}\%$. Adding more demographic controls reduces the elasticity to 6.08×10^{-6}.

Although the estimates in columns 1 and 2 are reasonable in terms of their signs, they could be biased because we ignore the potential change in coefficients over time. Columns 3 and 4 investigate such possibility. The results show that if we allow the coefficients to change over time, the elasticity in the later period is larger than that in the earlier period. Furthermore, the t-test result indicates that the increase in elasticity is statistically significant at 5% level. Therefore, it is problematic to interpret the conventional capitalization effect in column 1 and 2 as people's MWTP and to use them to conduct welfare analysis.

However, even though the fixed effects model in Table 3.2 controls for the time constant unobservable MSA characteristics, the crime index could still suffer from the endogeneity issue if there exist some unobservable time varying factors that influence both crime and housing price. To address this issue, we use the fixed effects instrumental variable model. Table 3.3 provides the first stage results for the model specifications without and with demographic controls, respectively. As expected, the effective abortion rate is negatively correlated with crime rate. Furthermore, the F-statistic testing the hypothesis that the coefficient on effective abortion rate equals zero is very high, indicating that the instrument is very strong. These results do not change if we add demographic control variables. Hence, based on the first stage results and the results from Donohue and Levitt (2001), it is safe to say that our instrumental variable does not suffer from the weak instrument problem.

Table 3.4 presents the main results in our paper. Column 1 shows that a one unit decrease in the crime index leads to a significant $2.88 \times 10^{-5}\%$ increase in the median housing price. This estimated elasticity is about 4 times larger than the one suggested

Table 3.3: First Stage Results of Fixed Effects Instrumental Variable Hedonic Model
Notes: The numbers are the first stage results of the two stage least square estimation of Table 3.4. The last column shows the F-statistic used to check whether the effective abortion rate suffers from the weak instrument problem. The numbers in the parenthesis are cluster robust standard errors. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)
Effective abortion rate	-6.53***	-5.83***
	(0.96)	(0.92)
Other demographic controls	No	Yes
Time dummies	Yes	Yes
Sample size	1023	1023
R^2	0.06	0.00
F-stat	45.81	39.80

in Table 3.2. The elasticity slightly reduces to 2.64×10^{-5} when we add demographic controls in column 2. To investigate whether people's MWTP for the reduction in crime changes over time, we again consider the time varying coefficients model in columns 3 and 4. The results show that the elasticity is slightly lower in the later periods than in the earlier periods, but the difference is not statistically significant, unlike the case shown in OLS estimations. However, the lower coefficient in the later period seems to point toward the intuition that when crime rate becomes lower, people's MWTP for a further reduction in crime also becomes lower as compared to when crime rate is higher. And even though the difference between the two periods are not significant, we still think the exercise in columns 3 and 4 is valuable, because we take the potential evolution into account. If there were any evolution in people's MWTP, our method would allow us to identify it.

3.4.2 Interpretations of Economic Magnitude

In this section, we use the instrumental variable regression results in Table 3.4 to interpret the economic magnitude of our findings, i.e., people's marginal willingness to pay (MWTP) for changes in crime risk. We take on two exercises as follows. First, we compare our results on the effect of crime on housing price with the ones in

Table 3.4: Estimates of the Impact of Crime on Housing Price by Fixed Effects Instrumental Variable Hedonic Model

Notes: See notes for Table 3.2. The numbers in the parenthesis are cluster robust standard errors. *,**,*** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Crime index($\times 10^5$)	-2.8795***	-2.6358***		
	(0.4865)	(0.5423)		
Crime index (period 1) ($\times 10^5$)			-2.9307***	-2.6814***
			(0.5419)	(0.6561)
Crime index (period 2) ($\times 10^5$)			-2.8048***	-2.5987***
			(0.6396)	(0.6292)
Household income ($\times 10^5$)		0.2331***		0.2346***
		(0.0614)		(0.0631)
Population ($\times 10^5$)		0.0027		0.0021
		(0.0053)		(0.0075)
Unemployment rate		0.0024		0.0007
		(0.1088)		(0.1107)
Poverty %		0.0390		-0.0460
		(0.0748)		(0.0947)
Old people %		-0.0428		-0.0447
		(0.0797)		(0.0818)
High school graduate %		-0.1567**		-0.1547**
		(0.0698)		(0.0723)
College graduate %		0.0799		0.0802
		(0.0605)		(0.0611)
White %		-0.0258		-0.0258
		(0.0534)		(0.0528)
Owned house %		-0.0283		-0.0285
		(0.0455)		(0.0459)
Other demographic controls	No	Yes	No	Yes
Time dummies	Yes	Yes	Yes	Yes
MSA dummies	Yes	Yes	Yes	Yes
Sample size	1023	1023	1023	1023
R^2	0.36	0.41	0.36	0.41
Difference of coefficients on crime index			-0.126	-0.827
			(0.639)	(0.673)

Pope and Pope (2012). Then, we further use these results to calculate the value of a statistical case of homicide.

We can calculate people's MWTP for a reduction in a certain type of crime by finding the effect of that type of crime on housing price, holding all other types of crime constant. For example, to calculate MWTP for a one-unit decrease in homicide risk, we first translate a one-unit decrease in homicide into a 26.4-unit decrease in the crime index based on the Sellin-Wolfgang weight, then use the regression coefficient of the crime index to calculate the effect on housing price, and finally use the sample average housing price to calculate the dollar value of avoiding a statistical case of homicide.

It might be useful to notice that for our above method to correctly capture the MWTP for a reduction in crime risk, the relative weights of different crime categories in Sellin-Wolfgang's crime index need to be equal to people's actual perception of seriousness for different types of crime as reflected in housing prices. The following simplified example can illustrate why such condition is necessary. Suppose there are only two types of crime, homicide and robbery, and housing price is affected by these two crimes based on the $HP = \beta_1 Homicide + \beta_2 Robbery + \varepsilon$. The crime index is formed as $Crime = \alpha_1 Homicide + \alpha_2 Robbery$. Homicide and robbery are related as $Homicide = \gamma Robbery + \xi$. In this case, the true MWTP for homicide risk is reflected by β_1. However, if homicide and robbery are highly correlated, then in practice, it is rarely possible to include both homicide and robbery in the housing price regression and identify them separately. One potential way to address this issue is to regress housing price on the crime index, which yields the coefficient $\frac{\beta_1+\beta_2\gamma}{\alpha_1+\alpha_2\gamma}$. Then, a one-unit change in homicide results in an α_1-unit change in the crime index, and an $\alpha_1\frac{\beta_1+\beta_2\gamma}{\alpha_1+\alpha_2\gamma}$-unit change in housing price. If the weights of homicide and robbery in the crime index equal their effects on housing price, i.e., $\alpha_1/\alpha_2 = \beta_1/\beta_2$, then $\alpha_1\frac{\beta_1+\beta_2\gamma}{\alpha_1+\alpha_2\gamma} = \beta_1$, which is the true MWTP. If the relative weight of homicide in the crime index is larger (smaller) than the relative effect of homicide on housing price, the MWTP calculated based on our method will overestimate (underestimate) the true MWTP to avoid homicide.

When calculating the economic magnitude of the results in Table 3.4, we also compare our findings with the ones in Pope and Pope (2012). One difference between our approach and Pope and Pope (2012) is that we classify crimes into seven categories, while Pope and Pope only divide crimes into violent crimes and property crimes. So in order to provide a meaningful comparison, we calculate the effect of homicide on housing price as an upper bound for the effect of violent crime, as homicide has the largest Sellin-Wolfgang weight in the crime index and is clearly the most serious type among the four types of violent crime. As for the lower bound of the effect of violent crime, we use robbery, which has the lowest Sellin-Wolfgang weight. Similarly, we use motor theft and larceny theft for the upper and lower bound for the effect of property crime, respectively. We obtain the following results.

For murder, a decrease of 100 cases per 10,000 people equals to a 26,400 unit ($1,000 \times 26.4$, as our crime numbers are measured per 100,000 people and the weight of homicide is 26.4) decrease in the crime index. Based on the coefficient in column 2 of Table 3.4, this is translated into a 69.7% increase in housing price. Similarly, a decrease in robberies by 100 per 10,000 people is associated with a 12.1% increase in housing price. As for property crimes, a decrease in motor theft and larceny theft by 100 cases per 10,000 people result in a 8.2% and 5.5% increase in housing price, respectively. Pope and Pope (2012) find in their study that a decrease in violent (property) crime of 100 cases per 10,000 people is associated with an increase in housing price of 4.3% (1.1%). So overall, our results suggest a three to five times larger MWTP than in Pope and Pope (2012).

An additional difference between the two studies is that our crime index method allows us to calculate the effect of one type of crime holding other types constant, while Pope and Pope include their two types of crime in two regressions separately. To be more specific, when calculating the effect of violent crime on housing price, Pope and Pope (2012) regress housing price on violent crime without controlling for property crime, so the coefficient reflects the combined effect of violent crime and property crime.

The final step we take is to use people's MWTP for a reduction in homicide to calculate the value of a statistical case of homicide. Strictly speaking, homicide contains not only murder and nonnegligent manslaughter, but also other cases such as accident killing of a person. However, to simplify language, we use the term "value of a statistical case of homicide," while it should actually be "value of a statistical case of murder and nonnegligent manslaughter."

Following the literature (e.g., Davis (2004)), we assume that housing price capitalizes the present discount value of all future homicide risk associated with living there. We further assume that people live infinitely, discount future risk at a 5% annual rate, and their perceived level of homicide risk in all future years equals to the current level of homicide risk, similar to Davis (2004). Based on these assumptions, a one-unit change in annual homicide risk is equivalent to a 21-unit change in lifetime homicide risk. Therefore, a decrease of lifetime homicide risk by one per 10,000 is associated with a 3.31×10^{-5} ($26.4/21 \times 2.64 \times 10^{-5}$) increase in housing price. As the mean housing price is \$121,681 (in 1999 dollars), this means a \$4.04 increase in housing price. Since the homicide risk is measured over per 100,000 people, the value of a statistical case of homicide is about 0.4 million dollars. This number is lower than the value of a statistical life estimated from cancer, mortality in labor, etc. (Gayer *et al.*, 2000, Viscusi and Aldy, 2003, Davis, 2004). One potential reason is that the weight of homicide relative to other crimes in Sellin-Wolfgang's index (13 times larger than larceny theft) is smaller than the actual effect of homicide relative to other crimes on housing price, so our method underestimates people's MWTP to avoid homicide. Nevertheless, the economic magnitude of our finding is still much larger than the magnitude found in related papers studying people's MWTP for crime risk such as Pope and Pope (2012).

3.5 Conclusion

This paper exploited the unexpected crime plunge during the 1990s to offer an estimate on people's willingness to pay for a safer living environment. We collect crime data and housing price on annual MSA level and calculate the Sellin and Wolfgang's weighted crime index in order to derive a comprehensive crime variable. To control for the potential upward bias in the fixed effect hedonic model, we use the effective abortion rate (Donohue and Levitt, 2001) as an instrumental variable for crime using the abortion rates in 1970s and 1980s and the age structure among arrestees. Since the conventional capitalization effect does not necessarily reflect MWTP, we also take into account the potential evolution of the gradients in the hedonic regression. Based on our preferred model specification, we conclude that a one unit increase in the crime index is associated with a 2.64×10^{-5} percent increase in the housing price. We further convert that number into a value of a statistical case of homicide and find people's willingness to pay to for the homicide reduction to be around 0.4 million in 1999 dollars.

REFERENCES

Ahern, K. R. and J. Harford, "The importance of industry links in merger waves", Journal of Finance **69**, 2, 527–576 (2014).

Andreoni, J., "Privately provided public goods in a large economy: The limits of altruism", Journal of Public Economics **35**, 1, 57–73 (1988).

Andreoni, J., "Toward a theory of charitable fund-raising", Journal of Political Economy **106**, 6, 1186–1213 (1998).

Andreoni, J., "Leadership giving in charitable fund-raising", Journal of Public Economic Theory **8**, 1, 1–22 (2006).

Andreoni, J., J. M. Rao and H. Trachtman, "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving", Journal of Political Economy (forthcoming).

Bagnoli, M. and S. G. Watts, "Selling to socially responsible consumers: Competition and the private provision of public goods", Journal of Economics & Management Strategy **12**, 3, 419–445 (2003).

Barrage, L., E. Chyn and J. Hastings, "Advertising, reputation, and environmental stewardship: Evidence from the bp oil spill", Unpublished working paper, National Bureau of Economic Research (2014).

BBB Wise Giving Alliance, "National charity reports", Better Business Bureau Http://www.bbb.org/charity-reviews/national/ (accessed July, 2013) (2011–2013).

Becker, M. J. and S. Thomas, "The spillover effects of changes in industry concentration", Unpublished working paper, University of Pittsburgh, Pittsburgh, PA (2008).

Bénabou, R. and J. Tirole, "Individual and corporate social responsibility", Economica **77**, 305, 1–19 (2010).

Bergstrom, T., L. Blume and H. Varian, "On the private provision of public goods", Journal of Public Economics **29**, 1, 25–49 (1986).

Besley, T. and M. Ghatak, "Retailing public goods: The economics of corporate social responsibility", Journal of Public Economics **91**, 9, 1645–1663 (2007).

Bound, J., D. A. Jaeger and R. M. Baker, "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak", Journal of the American statistical association **90**, 430, 443–450 (1995).

Bracha, A., M. Menietti and L. Vesterlund, "Seeds to succeed?: Sequential giving to public projects", Journal of Public Economics **95**, 5, 416–427 (2011).

Charness, G. and T. Cheung, “A restaurant field experiment in charitable contributions”, Economics Letters **119**, 1, 48–49 (2013).

Chay, K. Y. and M. Greenstone, “Does air quality matter? evidence from the housing market”, Journal of political Economy **113**, 2, 376–424 (2005).

Cornes, R. and T. Sandler, “Easy riders, joint production, and public goods”, Economic Journal **94**, 375, 580–598 (1984).

Cornes, R. and T. Sandler, “The comparative static properties of the impure public good model”, Journal of Public Economics **54**, 3, 403–421 (1994).

Davis, L. W., “The effect of health risk on housing values: Evidence from a cancer cluster”, American Economic Review **94**, 5, 1693–1704 (2004).

DellaVigna, S., J. A. List and U. Malmendier, “Testing for altruism and social pressure in charitable giving”, Quarterly Journal of Economics **127**, 1-56, qjr050 (2012).

Donohue, J. J. and S. D. Levitt, “The impact of legalized abortion on crime”, Quarterly Journal of Economics **116**, 2, 379–420 (2001).

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

Eckel, C. C., D. H. Herberich and J. Meer, “It’s not the thought that counts: A field experiment on gift exchange and giving at a public university”, Unpublished working paper, National Bureau of Economic Research (2016).

Falk, A., “Gift exchange in the field”, Econometrica **75**, 5, 1501–1511 (2007).

Gayer, T., J. T. Hamilton and W. K. Viscusi, “Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk”, Review of Economics and Statistics **82**, 3, 439–451 (2000).

Gibbons, S., “The costs of urban property crime”, Economic Journal **114**, 499 (2004).

Harbaugh, W. T., “What do donations buy?: A model of philanthropy based on prestige and warm glow”, Journal of Public Economics **67**, 2, 269–284 (1998).

Hotelling, H., “Stability in competition”, Economic Journal **39**, 153, 41–57 (1929).

Hsu, M., “Cultural and sexual differences on the judgment of criminal offenses: A replication study of the measurement of delinquency”, Journal of Criminal Law and Criminology **64**, 3, 348–353 (1973).

Huck, S., I. Rasul and A. Shephard, “Comparing charitable fundraising schemes: Evidence from a natural field experiment and a structural model”, American Economic Journal: Economic Policy **7**, 2, 326–369 (2015).

Karlan, D. and J. A. List, “Does price matter in charitable giving? evidence from a large-scale natural field experiment”, American Economic Review **97**, 5, 1774–1793 (2007).

Kiel, K. A. and K. T. McClain, “The effect of an incinerator siting on housing appreciation rates”, Journal of Urban Economics **37**, 3, 311–323 (1995).

Kotchen, M. and J. J. Moon, “Corporate social responsibility for irresponsibility”, B.E. Journal of Economic Analysis & Policy **12**, 1 (2012).

Kotchen, M. J., “Impure public goods and the comparative statics of environmentally friendly consumption”, Journal of Environmental Economics and Management **49**, 2, 281–300 (2005).

Kuminoff, N. V. and J. C. Pope, “Do capitalization effects for public goods reveal the public’s willingness to pay?”, International Economic Review **55**, 4, 1227–1250 (2014).

Leggett, C. G. and N. E. Bockstael, “Evidence of the effects of water quality on residential land prices”, Journal of Environmental Economics and Management **39**, 2, 121–144 (2000).

Levitt, S. D., “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not”, Journal of Economic Perspectives **18**, 1, 163–190 (2004).

Linden, L. and J. E. Rockoff, “Estimates of the impact of crime risk on property values from megan’s laws”, American Economic Review **98**, 3, 1103–1127 (2008).

Lynch, A. K. and D. W. Rasmussen, “Measuring the impact of crime on house prices”, Applied Economics **33**, 15, 1981–1989 (2001).

Margolis, J. D., H. A. Elfenbein and J. P. Walsh, “Does it pay to be good? a meta-analysis and redirection of research on the relationship between corporate social and financial performance”, Unpublished working paper, Harvard University, Boston, MA (2007).

Milgrom, P. and J. Roberts, “Price and advertising signals of product quality”, Journal of Political Economy **94**, 4, 796–821 (1986).

Miller, M. H. and K. Rock, “Dividend policy under asymmetric information”, Journal of Finance **40**, 4, 1031–1051 (1985).

Morgan, J., “Financing public goods by means of lotteries”, Review of Economic Studies **67**, 4, 761–784 (2000).

Normandeau, A., “Measurement of delinquency in montreal”, J. Crim. L. Criminology & Police Sci. **57**, 172 (1966).

Pope, D. G. and J. C. Pope, “Crime and property values: Evidence from the 1990s crime drop”, Regional Science and Urban Economics **42**, 1, 177–188 (2012).

Pope, J. C., "Fear of crime and housing prices: Household reactions to sex offender registries", Journal of Urban Economics **64**, 3, 601–614 (2008).

Potters, J., M. Sefton and L. Vesterlund, "Leading-by-example and signaling in voluntary contribution games: An experimental study", Economic Theory **33**, 1, 169–182 (2007).

Rose-Ackerman, S., "Charitable giving and excessive fundraising", Quarterly Journal of Economics **97**, 2, 193–212 (1982).

Sellin, T. and M. E. Wolfgang, *The measurement of delinquency* (New York: John Wiley & Sons, Inc., 1964).

Servaes, H. and A. Tamayo, "The impact of corporate social responsibility on firm value: The role of customer awareness", Management Science **59**, 5, 1045–1061 (2013).

Siegel, D. S. and D. F. Vitaliano, "An empirical analysis of the strategic use of corporate social responsibility", Journal of Economics & Management Strategy **16**, 3, 773–792 (2007).

Spence, M., "Job market signaling", Quarterly Journal of Economics **87**, 3, 355–374 (1973).

Thaler, R., "A note on the value of crime control: Evidence from the property market", Journal of Urban Economics **5**, 1, 137–145 (1978).

Velez-Diaz, A. and E. I. Megargee, "An investigation of differences in value judgments between youthful offenders and non-offenders in Puerto Rico", Journal of Criminal Law, Criminology, and Police Science **61**, 4, 549–553 (1970).

Vesterlund, L., "The informational value of sequential fundraising", Journal of Public Economics **87**, 3, 627–657 (2003).

Viscusi, W. K. and J. E. Aldy, "The value of a statistical life: a critical review of market estimates throughout the world", Journal of risk and uncertainty **27**, 1, 5–76 (2003).

Wellford, C. F. and M. Wiatrowski, "On the measurement of delinquency", Journal of Criminal Law and Criminology **66**, 2, 175–188 (1975).

APPENDIX A

PROOFS IN CHAPTER 1

A.1 Proof of Proposition 1

Step 1. Claim: When K is sufficiently small, if open is reached, both consumers search. Moreover, if charity type is $highA$, equilibrium outcome is $(g^*_{A|A}, g^*_{B|A}) = (1,0)$; if charity type is $highB$, equilibrium outcome is $(g^*_{A|B}, g^*_{B|B}) = (0,1)$.

Given B searches and $g^*_{B|A} = 0$, $g^*_{B|B} = 1$,

(i) if A does not deviate, Expected Payoff (search) is: $\frac{1}{2}(w-1) + \frac{1}{2}w - K$.

(ii) if A deviates to no search, the highest payoff he can get depends on the solution to:

$$\max_{g_A} \frac{1}{2}(w - g_A + log(g_A + g^*_{B|A})) + \frac{1}{2}(w - g_A + plog(g_A + g^*_{B|B}))\,.$$

The solution would be $g_A = (p - 1 + \sqrt{p^2 - 2p + 9})/4$. Hence, Expected Payoff (no search) is: $w - g_A + \frac{1}{2}log(g_A) + \frac{1}{2}plog(g_A + 1)$.

Expected Payoff (search) $-$ Expected Payoff (no search) $\approx g_A - \frac{1}{2} - \frac{1}{2}log(g_A) - \frac{1}{2}plog(g_A+1) - K$. Therefore, when $K^U \leq g_A - \frac{1}{2} - \frac{1}{2}log(g_A) - \frac{1}{2}plog(g_A+1)$, A would not deviate to no search.

(iii) when charity type is $highA$, $g^*_{A|A} = 1$ maximizes A's utility. Therefore, at open, given B's equilibrium strategy, A would not deviate. Since the Consumers' problems are symmetric, given A's equilibrium strategy, B also would not deviate.

Step 2. Claim: When K is sufficiently large, when epsilon is reached, both consumers take on no search, given belief described in Proposition. Furthermore, with $\varepsilon = \frac{q+qp}{2}log(2)$, both consumers would choose $g^*_i = \varepsilon$.

Given consumers take on no search and $g^*_B = \varepsilon$,

(i) if A does not deviate, Expected Payoff (no search, $g_A = \varepsilon$) is: $\frac{1}{2}(w - \varepsilon + qlog(\varepsilon + \varepsilon)) + \frac{1}{2}(w - \varepsilon + qplog(\varepsilon + \varepsilon))$.

(ii) if A deviates to $g_A = 0$, Expected Payoff (no search, $g_A = 0$) is: $\frac{1}{2}(w + qlog(\varepsilon)) + \frac{1}{2}(w + qplog(\varepsilon))$.

Expected Payoff (no search, $g_A = \varepsilon$) $\geq$ Expected Payoff (no search, $g_A = 0$) if and only if $\varepsilon \leq \frac{q+qp}{2}log(2)$.

The exact same argument applies to Consumer B.

Given Consumer B takes on no search and donates ε, if A deviates to search, and if charity type is $lowA$, A would donate. And if charity type is $lowB$, A would not donate. Hence, Expected Payoff (search) is: $\frac{1}{2}(w - \varepsilon + qlog(\varepsilon + \varepsilon)) + \frac{1}{2}(w + qplog(\varepsilon)) - K$. Hence, when $K^L \geq \frac{q-qp}{4}log(2)$, Expected Payoff (no search, $g_A = \varepsilon$) $\geq$Expected Payoff (search).

The exact same argument applies to Consumer B. It is also checked that $K^L < K^U$.

Step 3. Claim: Given consumers' belief and strategy, when $p \geq \frac{1}{log(2)} - 1 \approx 0.4427$, a low type charity would choose strategy epsilon with $\varepsilon^* = \frac{q+qp}{2}log(2)$.

If a low type deviates to open, then consumers will search and find out the charity's type. Then the consumer who cares more about the purpose donates δ, and the consumer who cares less does not donate. Then the charity's payoff, i.e. total fund raised, is $G_{bad}(open) = q$.

If the low type does not deviate, it could raise $G_{bad}(\varepsilon^*) = 2\varepsilon^* = (q + qp)log(2)$. It holds for $p \geq \frac{1}{log(2)} - 1$ that $G_{bad}(\varepsilon^*) \geq G_{bad}(open)$.

If the low type deviates to an epsilon amount larger than $\frac{q+qp}{2}log(2)$, as shown in Step 2, consumer would donate zero, given the other consumer donates the epsilon amount. Since the epsilon amount cannot exceed q, the charity is better off asking for $\varepsilon^* = \frac{q+qp}{2}log(2)$. Moreover, it is straightforward that the low type would not choose a smaller epsilon amount than ε^*. Hence, the claim holds.

Step 4. Claim: Given consumers' belief and strategy, when $q(1 + p)log(2) \leq 1$, a high type would choose strategy open.

If a high type deviates to ε^*, then consumers will not search. So $G_{good}(\varepsilon^*) = 2\varepsilon^*$.

And as calculated in Step 1, $G_{good}(open) = 1$. $G_{good}(open) \geq G_{good}(\varepsilon^*)$ if and only if $(q + qp)log(2) \leq 1$.

A.2 Proof of Proposition 2

Step 1. Claim: There does not exist an equilibrium where consumers do not search.

If consumers do not search, then they donate $\tilde{g}$ where $\tilde{g}$ is the solution to g that maximizes: $w - g + (1+q)(1+p)log(g+\tilde{g})$. $\tilde{g} = \frac{1}{8}(1+q)(1+p)$. This gives an expected payoff of $w - \frac{1}{8}(1+q)(1+p) + \frac{1}{4}(1+q)(1+p)log(\frac{(1+q)(1+p)}{4})$.

If a consumer deviates to search, he fully learns the type and maximize utility under each type. The ex ante expected payoff would be $\frac{1}{4}(w-(1-\frac{(1+q)(1+p)}{8})) + \frac{1}{4}(w-(p-\frac{(1+q)(1+p)}{8}) + plog(p)) + \frac{1}{4}(w - max\{q - \frac{(1+q)(1+p)}{8}, 0\} + qlog(max\{q - \frac{(1+q)(1+p)}{8}, 0\} + \frac{(1+q)(1+p)}{8})) + \frac{1}{4}(w - max\{qp - \frac{(1+q)(1+p)}{8}, 0\} + qlog(max\{qp - \frac{(1+q)(1+p)}{8}, 0\} + \frac{(1+q)(1+p)}{8}))$, which is higher than the expected payoff without deviation, for all q, p that satisfy the conditions in Proposition 1. Therefore, there is no equilibrium where both consumers do not search.

Step 2. Claim: There exists an equilibrium where consumers search.

If a consumer searches, he donates 1, if the charity is of high quality and matching type, and donates q, if the charity is of low quality and matching type. If the purpose of charity does not match the consumer's interest, then he does not donate. So the expected payoff from searching is $w - \frac{1}{4}(1 + q) + \frac{1}{4}q(1 + p)log(q) - K$. If the consumer deviates to no search, then it can be shown that for search costs small enough, deviation leads to a lower expected payoff, for all q, p that satisfy the conditions in Proposition 1. Hence, consumers would not deviate, and both searching is an equilibrium outcome.

Step 3. Claim: The equilibrium in Step 2 yields lower ex ante payoff than the baseline separating equilibrium.

The expected payoff from the baseline equilibrium is $-\frac{1}{4}+\frac{1}{4}q(1+p)log(\frac{q(1+p)}{2}log(2))-\frac{K}{2}$. So Expected Payoff (Baseline) $-$ Expected Payoff (No Epsilon) is $\frac{1}{4}q + \frac{1}{4}q(1+p)log(\frac{1+p}{2}log(2)) + \frac{K}{2}$. Since q, p satisfy the conditions in Proposition 1, $\frac{1}{4}q + \frac{1}{4}q(1+p)log(\frac{1+p}{2}log(2)) \geq 0$. And since search cost is strictly positive, Expected Payoff (Baseline) $>$ Expected Payoff (No Epsilon).

In sum, the baseline equilibrium yields higher consumer welfare than the case where epsilon is banned.

APPENDIX B

PROOFS IN CHAPTER 2

B.1 Proof of Proposition 0

Let $g_i(d_i, d_j)$ denote firm $i's$ first order condition of maximization total profit with respect to donation d_i given firm $j's$ donation d_j under symmetric information. Specifically, $g_i(d_i, d_j) = \frac{2\beta}{3}[\frac{l}{2} + \frac{\alpha}{6}(r_i - r_j) + \frac{\beta}{6}(d_i - d_j)] - k_d(d_i, r_i)$, $i = 1, 2$ and $i \neq j$. Let $d_i(d_j, \beta)$ denote the solution of $g_i(d_i(d_{j,}\beta), d_j) = 0$.

I first show that $d_1^{SI} > d_2^{SI}$. Combining $g_1(d_1^{SI}, d_2^{SI}) = 0$ and $g_2(d_2^{SI}, d_1^{SI}) = 0$, I obtain $\frac{2\beta}{9}\alpha(r_1 - r_2) = k_d(d_1^{SI}, r_1) - k_d(d_2^{SI}, r_2) - \frac{2\beta^2}{9}(d_1^{SI} - d_2^{SI})$. Strict single crossing property of $k(d, r)$ implies that $\frac{2\beta}{9}\alpha(r_1 - r_2) < k_d(d_1^{SI}, r_2) - k_d(d_2^{SI}, r_2) - \frac{2\beta^2}{9}(d_1^{SI} - d_2^{SI}) = \int_{d_2^{SI}}^{d_1^{SI}}[k_{dd}(x, r_2) - \frac{2\beta^2}{9}]dx$. Note that the steadiness of equilibrium under symmetric information requires $-\frac{\partial g_2(d_2,d_1)}{\partial d_2} > -\frac{\partial g_1(d_1,d_2)}{\partial d_2}$, so $k_{dd}(d, r_2) - \frac{\beta^2}{9} > \frac{\beta^2}{9}$. As $\frac{2\beta}{9}\alpha(r_1 - r_2) > 0$ and $k_{dd}(x, r_2) - \frac{2\beta^2}{9} > 0$, it must be $d_1^{SI} > d_2^{SI}$.

Next, I show that $\frac{d(d_1^{SI} + d_2^{SI})}{d\beta} > 0$. $g_1(d_1^{SI}, d_2^{SI}) = 0$ and $g_2(d_2^{SI}, d_1^{SI}) = 0$ imply that $\frac{2\beta}{3}l = k_d(d_1^{SI}, r_1) + k_d(d_2^{SI}, r_2)$. Differentiating both sides with respect to β, I obtain $k_{dd}(d_1^{SI}, r_1)\frac{d(d_1^{SI})}{d\beta} + k_{dd}(d_2^{SI}, r_2)\frac{d(d_2^{SI})}{d\beta} = \frac{2}{3}l > 0$. As $k_{dd}(d_i, r_i) > 0$, it must be $\frac{d(d_1^{SI})}{d\beta} > 0$, or $\frac{d(d_2^{SI})}{d\beta} > 0$, or both. Suppose $\frac{d(d_2^{SI})}{d\beta} > 0$. Note that $\frac{d(d_1^{SI})}{d\beta} = \frac{d(d_1(d_2^{SI}, \beta))}{d\beta} = \frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}}\frac{d(d_2^{SI})}{d\beta} + \frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}}$, so it is sufficient to show $\frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}} > -1$ and $\frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}} > 0$. Taking the partial derivative on both sides of $g_1(d_1^{SI}, d_2^{SI}) = 0$ with respect to d_2^{SI} and β, I get $\frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}} = -\frac{\beta^2}{9}/(k_{dd}(d_1, r_1) - \frac{\beta^2}{9}) > -1$ and $\frac{\partial(d_1(d_2^{SI}, \beta))}{\partial d_2^{SI}} = \frac{2}{3}[\frac{l}{2} + \frac{\alpha}{6}(r_i - r_j) + \frac{\beta}{3}(d_1^{SI} - d_2^{SI})]/(k_{dd}(d_1, r_1) - \frac{\beta^2}{9}) > 0$. Therefore, $\frac{d(d_1^{SI} + d_2^{SI})}{d\beta} > 0$. Similar argument can be applied to the case in which $\frac{d(d_1^{SI})}{d\beta} > 0$.

B.2 Proof of Claim 1

I prove the claim using contradiction. Suppose that $\widetilde{d}_1 < d_1^{SI}$.

I continue to use the notation of $g_i(d_i, d_j)$ as in the proof of claim 1. The goal is to prove that $g_1(\widetilde{d}_1, \widetilde{d}_2) > 0$. Since firm 1 is still considered as high type if it increases its donation, $g_1(\widetilde{d}_1, \widetilde{d}_2) > 0$ indicates that firm 1 can increase its profit by increasing donation under asymmetric information, which contradicts $(\widetilde{d}_1, \widetilde{d}_2)$ being an equilibrium.

Note that by the definition of g_i, $g_1(d_1^{SI}, d_2^{SI}) = 0$. Thus, $g_1(\widetilde{d}_1, \widetilde{d}_2) > 0$ is equivalent to $\int_{\widetilde{d}_1}^{d_1^{SI}} [k_{dd}(d_1, r_1) - \frac{\beta^2}{9}]dd_1 > \int_{d_2^{SI}}^{\widetilde{d}_2} \frac{\beta^2}{9} dd_2$. Note that the steadiness of equilibrium under symmetric information requires $-\frac{\partial g_1(d_1,d_2)}{\partial d_1} > -\frac{\partial g_2(d_1,d_2)}{\partial d_1}$, so $k_{dd}(d_1, r_1) - \frac{\beta^2}{9} > \frac{\beta^2}{9}$. As $d_1^{SI} - \widetilde{d}_1$¿0, a sufficient condition for $g_1(\widetilde{d}_1, \widetilde{d}_2) > 0$ is $d_1^{SI} - \widetilde{d}_1 > \widetilde{d}_2 - d_2^{SI}$. As firm 2 is considered as low type for at least an interval around $\widetilde{d}_2$, it must be the case that $g_2(\widetilde{d}_2, \widetilde{d}_1) = 0$. Otherwise, $\widetilde{d}_2$ is not the optimal choice of firm 2 under asymmetric information. Combining $g_2(d_2^{SI}, d_1^{SI}) = 0$ and $g_2(\widetilde{d}_2, \widetilde{d}_1) = 0$, one can easily show that $\int_{\widetilde{d}_1}^{d_1^{SI}} \frac{\beta^2}{9} dd_1 = \int_{d_2^{SI}}^{\widetilde{d}_2} [k_{dd}(d_2, r_2) - \frac{\beta^2}{9}]dd_2$, which implies $d_1^{SI} - \widetilde{d}_1 > \widetilde{d}_2 - d_2^{SI}$. This completes the proof of $g_1(\widetilde{d}_1, \widetilde{d}_2) > 0$ and results in contradiction.

B.3 Proof of Claim 2

I continue to use the notation of $g_i(d_i, d_j)$ as in the proof of claim 1. By the argument in the proof above, $g_2(d_2^{SI}, d_1^{SI}) = 0$ and $g_2(\widetilde{d}_2, \widetilde{d}_1) = 0$. Therefore, $\int_{d_1^{SI}}^{\widetilde{d}_1} \frac{\beta^2}{9} dd_1 = \int_{\widetilde{d}_2}^{d_2^{SI}} [k_{dd}(d_2, r_2) - \frac{\beta^2}{9}]dd_2$. As $\widetilde{d}_1 - d_1^{SI} > 0$ and $k_{dd}(d_2, r_2) - \frac{\beta^2}{9} > \frac{\beta^2}{9}$, I obtain $\widetilde{d}_1 - d_1^{SI} > d_2^{SI} - \widetilde{d}_2$, which is $\widetilde{d}_1 + \widetilde{d}_2 > d_1^{SI} + d_2^{SI}$. The equation also implies $d_2^{SI} - \widetilde{d}_2 > 0$ as $\widetilde{d}_1 - d_1^{SI} > 0$. Thus, $\widetilde{d}_1 - d_1^{SI} > 0 > \widetilde{d}_2 - d_2^{SI}$, which indicates $\widetilde{d}_1 - \widetilde{d}_2 > d_1^{SI} - d_2^{SI}$.

B.4 Proof of Proposition 1

I first prove the existence of a separating equilibrium, and then show that (d_1^{AI}, d_2^{AI}) characterized by conditions (1) and (2) is an equilibrium.

I continue to use the notation of $g_i(d_i, d_j)$ and $d_2(d_1)$ as in the proof of claim 1. Note that $d_2^{AI} = d_2(d_1^{AI})$ by condition (2).

Existence

For any $\widetilde{d}_1 \geq d_1^{SI}$, let $B_2^{dev}(\widetilde{d}_1)$ denote firm 2's benefit of deviating from $d_2(\widetilde{d}_1)$ to $\widetilde{d}_1$ under asymmetric information. As firm 1 and firm 2 are considered as high type and low type respectively under $(\widetilde{d}_1,\ d_2(\widetilde{d}_1))$, firm 2's profit is $2[\frac{l}{2} + \frac{\alpha}{6}(r_2 - r_1) + \frac{\beta}{6}(d_2(\widetilde{d}_1) - \widetilde{d}_1)]^2 - k(d_2(\widetilde{d}_1), r_2)$. If firm 2 deviates to $\widetilde{d}_1$, then both firms are considered as high type, and firm 2's profit becomes $2(\frac{l}{2})^2 - k(\widetilde{d}_1, r_2)$. [1] Thus, $B_2^{dev}(\widetilde{d}_1) = 2(\frac{l}{2})^2 - k(\widetilde{d}_1, r_2) - 2[\frac{l}{2} + \frac{\alpha}{6}(r_2 - r_1) + \frac{\beta}{6}(d_2(\widetilde{d}_1) - \widetilde{d}_1)]^2 + k(d_2(\widetilde{d}_1), r_2)$. Note that $B_2^{dev}(\widetilde{d}_1) = 0$ is condition (1) in the proposition, and it pins down d_1^{AI} and $d_2(d_1^{AI})$.

A sufficient condition for separating equilibrium to exist is $dB_2^{dev}(\widetilde{d}_1)/d(\widetilde{d}_1) < 0$. If $B_2^{dev}(d_1^{SI}) \leq 0$, then (d_1^{SI}, d_2^{SI}) is a separating equilibrium. If $B_2^{dev}(d_1^{SI}) > 0$, then $dB_2^{dev}(\widetilde{d}_1)/d(\widetilde{d}_1) < 0$ indicates that there exists $\widetilde{d}_1$ such that $B_2^{dev}(\widetilde{d}_1) \leq 0$. Since $\partial B_2^{dev}(\widetilde{d}_1)/\partial d_2 = g_2(d_2(\widetilde{d}_1), \widetilde{d}_1) = 0$, I have $dB_2^{dev}(\widetilde{d}_1)/d(\widetilde{d}_1) = \partial B_2^{dev}(\widetilde{d}_1)/\partial \widetilde{d}_1 = -k_d(\widetilde{d}_1, r_2) + \frac{2\beta}{3}[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) - \frac{\beta}{6}(\widetilde{d}_1 - d_2(\widetilde{d}_1))]$. Using $g_1(d_1^{SI}, d_2^{SI}) = 0$ and noticing that $k_{dd} > 0$, $\widetilde{d}_1 \geq d_1^{SI}$, $r_1 \geq r_2$ and $\widetilde{d}_1 \geq d_2(\widetilde{d}_1)$, one can easily obtain that $dB_2^{dev}(\widetilde{d}_1)/d(\widetilde{d}_1) < 0$.

(d_1^{AI}, d_2^{AI}) characterized by conditions (1) and (2) is an equilibrium

I prove that (d_1^{AI}, d_2^{AI}) is an equilibrium by ruling out all possible deviations in the following five cases.

(i) Firm 1 does not have incentive to deviate to $\hat{d}_1 > d_1^{AI}$.

If firm 1 deviates upwards, it is still considered as high type. Analogous to the proof of claim 1, a sufficient condition for $g_1(\hat{d}_1, d_2^{AI}) < 0$ for any $\hat{d}_1 \geq d_1^{AI}$ is $\hat{d}_1 - d_2^{SI} >$

[1] If firm 2 deviates to $\hat{d}_2 > d_1^{AI}$, it is considered as a high type firm based on consumer's belief function. In this case, both firms are considered as high type, or more generally both firms are considered to have the same responsibility distribution. With regard to consumers' choice between firm 1 and firm 2, it does not matter whether consumers believe that both firms have high responsibility (more consistent with the belief function), or that firm 1 has high responsibility and firm 2 has low responsibility with 1/2 probability and the other way around for the remaining 1/2 probability (more consistent with the model setup of one high responsibility firm and one low responsibility firm).

$d_2^{SI} - d_2^{AI}$. $g_2(d_2^{SI}, d_1^{SI}) = 0$ and $g_2(d_2^{AI}, d_1^{AI}) = 0$ together imply that $d_1^{AI} - d_2^{SI} > d_2^{SI} - d_2^{AI}$. As $\hat{d}_1 > d_1^{AI}$, I obtain $g_1(\hat{d}_1, d_2^{AI}) < 0$, which indicates it is not beneficial for firm 1 to deviate upwards.

(ii) Firm 2 does not have incentive to deviate to $\hat{d}_2 < d_1^{AI}$.

If firm 2 deviates to any $\hat{d}_2 < d_1^{AI}$, it is still considered as low type. Since $d_2^{AI} = d_2(d_1^{AI})$ is firm 2's optimal donation level given being a low type, it is not beneficial for firm 2 to deviate to $\hat{d}_2 < d_1^{AI}$.

(iii) Firm 2 does not have incentive to deviate to $\hat{d}_2 = d_1^{AI}$.

By condition (1), firm 2 generates same profit under d_2^{AI} and d_1^{AI}, so it has no incentive to deviate.

(iv) Firm 2 does not have incentive to deviate to $\hat{d}_2 > d_1^{AI}$.

Suppose firm 2 deviates to $\hat{d}_2 > d_1^{AI}$. Then firm 2 is considered as high type. Since firm 2's total profit under (d_1^{AI}, d_2^{AI}) is same as its profit under (d_1^{AI}, d_1^{AI}) by condition (1), I only need to show that firm 2's profit under $(d_1^{AI}, \hat{d}_2)$ is lower than its profit under (d_1^{AI}, d_1^{AI}) to construct a contradiction. The idea is that if firm 1 as a good firm has no incentive to donate more than d_1^{AI} (as shown in (i)), firm 2 as a bad firm should also have no incentive to do so. Mathematically, it is easy to show that $2[\frac{l}{2} + \frac{\beta}{6}(\hat{d}_2 - d_1^{AI})]^2 - 2(\frac{l}{2})^2 < 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(\hat{d}_2 - d_2^{AI})]^2 - 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2$, and the single crossing property of the cost function implies $k(\hat{d}_2, r_1) - k(d_1^{AI}, r_1) < k(\hat{d}_2, r_2) - k(d_1^{AI}, r_2)$. Also, (i) indicates that $2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(\hat{d}_2 - d_2^{AI})]^2 - 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 < k(\hat{d}_2, r_1) - k(d_1^{AI}, r_1)$. These inequalities together imply that $2[\frac{l}{2} + \frac{\beta}{6}(\hat{d}_2 - d_1^{AI})]^2 - k(\hat{d}_2, r_2) < 2(\frac{l}{2})^2 - k(d_1^{AI}, r_2)$, so firm 2's profit under $(d_1^{AI}, \hat{d}_2)$ is lower than its profit under (d_1^{AI}, d_1^{AI}). Thus, firm 2 has no incentive to deviate to $\hat{d}_2 > d_1^{AI}$.

(iv) Firm 1 does not have incentive to deviate to $\hat{d}_1 < d_1^{AI}$.

Suppose firm 1 deviates to $\hat{d}_1 < d_1^{AI}$. Then firm 1 is considered as low type. Let

$\pi_1^{d_1^{AI}}$ denote firm 1's profit with no deviation. $\pi_1^{d_1^{AI}} = 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - k(d_1^{AI}, r_1)$. Since firm 2's benefit of deviating to d_1^{AI} is zero by condition (1), $\pi_1^{d_1^{AI}} = 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - 2(\frac{l}{2})^2 + 2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 + k(d_1^{AI}, r_2) - k(d_2^{AI}, r_2) - k(d_1^{AI}, r_1)$. The single crossing property of the cost function implies $k(d_1^{AI} r_2) - k(d_1^{AI}, r_1) > k(\hat{d}_1, r_2) - k(\hat{d}_1, r_1)$. So $\pi_1^{d_1^{AI}} > 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - 2(\frac{l}{2})^2 + 2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 + k(\hat{d}_1, r_2) - k(\hat{d}_1, r_1) - k(d_2^{AI}, r_2)$. As d_2^{AI} is firm 2's optimal donation level under $d_2 < d_1^{AI}$ given firm 1's donation d_1^{AI}, $2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - k(d_2^{AI}, r_2) > 2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) - \frac{\beta}{6}(d_1^{AI} - \hat{d}_1)]^2 - k(\hat{d}_1, r_2)$. So $\pi_1^{d_1^{AI}} > 2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - 2(\frac{l}{2})^2 + 2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) - \frac{\beta}{6}(d_1^{AI} - \hat{d}_1)]^2 - k(\hat{d}_1, r_1)$. It can be easily shown that $2[\frac{l}{2} + \frac{\alpha}{6}(r_1 - r_2) + \frac{\beta}{6}(d_1^{AI} - d_2^{AI})]^2 - 2(\frac{l}{2})^2 > 2[\frac{l}{2} + \frac{\beta}{6}(\hat{d}_1 - d_2^{AI})]^2 - 2[\frac{l}{2} - \frac{\alpha}{6}(r_1 - r_2) - \frac{\beta}{6}(d_1^{AI} - \hat{d}_1)]^2$. Therefore, $\pi_1^{d_1^{AI}} > 2[\frac{l}{2} + \frac{\beta}{6}(\hat{d}_1 - d_2^{AI})]^2 - k(\hat{d}_1, r_1)$. As the right-hand-side of the inequality is firm 1's profit with deviation $\hat{d}_1$, the inequality indicates that firm 1 does not have incentive to deviate downwards.

B.5 Proof of Proposition 2

Since $d_1^{AI} + d_2^{AI} > d_1^{SI} + d_2^{SI}$, it is sufficient to show that $C^{AI} + \Pi^{AI} > C^{SI} + \Pi^{SI}$.

For both symmetric and asymmetric information cases, $C + \Pi = c_1 + c_2 + \pi_1 + \pi_2$, where c_i denotes consumers' utility from purchasing goods from firm i, and π_i denotes firm $i's$ total profit. Let x^* denote the consumer who is just indifferent between purchasing from firm 1 or firm 2. Then, $\pi_1 + \pi_2 = x^*(p_1 - c) - k(d_1, r_1) + (l - x^*)(p_2 - c) - k(d_2, r_2)$, and $c_1 + c_2 = \int_0^{x^*} [-x - p_1 + \alpha r_1 + \beta(d_1 - d_2)]dx + \int_0^{l - x^*} [-x - p_2 + \alpha r_2]dx$. After simplifications, $C + \Pi$ can be written as $-\frac{1}{2}(x^*)^2 - \frac{1}{2}(l - x^*)^2 + [\alpha(r_1 - r_2) + \beta(d_1 - d_2)]x^* - k(d_1, r_1) - k(d_2, r_2) + l(\alpha r_2 - c)$.

Since $d_2^{AI} < d_2^{SI}$, a sufficient condition for $C^{AI} + \Pi^{AI} > C^{SI} + \Pi^{SI}$ is $-\frac{1}{2}(x^{AI})^2 - \frac{1}{2}(l - x^{AI})^2 + [\alpha(r_1 - r_2) + \beta(d_1^{AI} - d_2^{AI})]x^{AI} - k(d_1^{AI}, r_1) > -\frac{1}{2}(x^{SI})^2 - \frac{1}{2}(l - x^{SI})^2 + [\alpha(r_1 - r_2) + \beta(d_1^{SI} - d_2^{SI})]x^{SI} - k(d_1^{SI}, r_1)$, where $x^{AI} = \frac{1}{2}l + \frac{\beta}{6}(d_1^{AI} - d_2^{AI}) + \frac{\alpha}{6}(r_1 - r_2)$, and $x^{SI} = \frac{1}{2}l + \frac{\beta}{6}(d_1^{SI} - d_2^{SI}) + \frac{\alpha}{6}(r_1 - r_2)$. Note that (d_1^{AI}, d_2^{AI}) is a separating equilibrium, so firm 1's profit is higher under (d_1^{AI}, d_2^{AI}) than under (d_1^{SI}, d_2^{AI}). Thus, $2(x^{AI})^2 - k(d_1^{AI}, r_1) > 2[\frac{l}{2} + \frac{\beta}{6}(d_1^{SI} - d_2^{AI})]^2 - k(d_1^{SI}, r_1) > 2[\frac{l}{2} + \frac{\beta}{6}(d_1^{SI} - d_2^{SI})]^2 - k(d_1^{SI}, r_1)$. Using this equality, the sufficient condition for the proposition becomes $2(x^{AI})^2 - 2[\frac{l}{2} + \frac{\beta}{6}(d_1^{SI} - d_2^{SI})]^2 < \frac{1}{2}(x^{SI})^2 + \frac{1}{2}(l - x^{SI})^2 - \frac{1}{2}(x^{AI})^2 - \frac{1}{2}(l - x^{AI})^2 + [\alpha(r_1 - r_2) + \beta(d_1^{AI} - d_2^{AI})]x^{AI} - [\alpha(r_1 - r_2) + \beta(d_1^{SI} - d_2^{SI})]x^{SI}$.

The LHS of the above inequality can be simplified to $[(d_1^{AI} - d_2^{AI}) - (d_1^{SI} - d_2^{SI})][\frac{1}{3}l\beta + \frac{1}{18}\alpha\beta(r_1 - r_2) + \frac{1}{18}\beta^2(d_1^{SI} - d_2^{SI} + d_1^{AI} - d_2^{AI})]$, and the RHS can be simplified to $[(d_1^{AI} - d_2^{AI}) - (d_1^{SI} - d_2^{SI})][\frac{1}{2}l\beta + \frac{5}{18}\alpha\beta(r_1 - r_2) + \frac{5}{36}\beta^2(d_1^{SI} - d_2^{SI} + d_1^{AI} - d_2^{AI})]$. Thus, the inequality holds. This completes the proof of proposition 2.

B.6 Welfare Calculation

I show how total social welfare is calculated under social planner's problem, symmetric information, and asymmetric information. These calculations are used in the numerical example.

Social Planner's Problem

Social planner maximizes total welfare $C + W - G + N\alpha log(G + d_1 + d_2) + \Pi$ by choosing p_1, p_2, d_1, d_2 and G. Based on the proof of Proposition 2, $C + \Pi = -\frac{1}{2}(x^*)^2 - \frac{1}{2}(l - x^*)^2 + [(r_1 - r_2) + \beta(d_1 - d_2)]x^* - k(d_1, r_1) - k(d_2, r_2) + l(\alpha r_2 - c)$, where $x^* = \frac{1}{2}l - \frac{1}{2}(p_1 - p_2) + \frac{\alpha}{2}(r_1 - r_2) + \frac{\beta}{2}(d_1 - d_2)$. Here, I suppose $d_1 \geq d_2$. Later I will show that $d_1 < d_2$ cannot be socially optimal. I solve the problem by first finding the optimal choice of p_1, p_2 for any given d_1, d_2 and G, and then solve the full maximization problem. Since p_1, p_2 is one-to-one determined by x^*, finding the optimal p_1, p_2 is equivalent to finding the optimal x^*. One can easily show that social planner's optimal x^* is $\frac{1}{2}l + \frac{\alpha}{2}(r_1 - r_2) + \frac{\beta}{2}(d_1 - d_2)$, and the objective function can written as $[\frac{1}{2}l + \frac{\alpha}{2}(r_1 - r_2) + \frac{\beta}{2}(d_1 - d_2)]^2 - \frac{1}{2}l^2 - lc + l\alpha r_2 - k(d_1, r_1) - k(d_2, r_2) - G +$

$N\alpha log(G+d_1+d_2)$. d_1^{SP}, d_2^{SP} is either pinned down by the FOC of d_1, d_2 if $d_2^{SP}>0$, or d_1^{SP} is the solution to its FOC and $d_2^{SP}=0$. $G^{SP}=N\alpha-d_1^{SP}-d_2^{SP}$.

If social planner chooses $d_1<d_2$, the objective becomes $[\frac{1}{2}l-\frac{\alpha}{2}(r_1-r_2)+\frac{\beta}{2}(d_2-d_1)]^2-\frac{1}{2}l^2-lc+l\alpha r_1-k(d_1,r_1)-k(d_2,r_2)-G+N\alpha log(G+d_1+d_2)$, which can be shown to be smaller than $[\frac{1}{2}l+\frac{\alpha}{2}(r_1-r_2)+\frac{\beta}{2}(d_2-d_1)]^2-\frac{1}{2}l^2-lc+l\alpha r_2-k(d_1,r_1)-k(d_2,r_2)-G+N\alpha log(G+d_1+d_2)$ for any given d_1, d_2 and G. This means that for any choice of $d_1<d_2$, social planner can increase welfare by making firm 1 donate the higher amount d_2 and firm 2 denote d_1. Thus, $d_1<d_2$ cannot be socially optimal.

Symmetric Information

Firms' choice of donation d_1^{SI} and d_2^{SI} under symmetric information is solved by $g_1(d_1^{SI},d_2^{SI})=0$ and $g_2(d_2^{SI},d_1^{SI})=0$. For each individual consumer i, the optimal choice of $g_i=max\{0,(\alpha-d_1^{SI}-d_2^{SI})/N\}$.

Asymmetric Information

Firms' choice of donation d_1^{AI} and d_2^{AI} under asymmetric information depends on whether (d_1^{SI},d_2^{SI}) can be supported as a separating equilibrium. If Δr is small enough such that firm 2 does not have incentive to mimic firm 1's donation at d_1^{SI} (see proof of Proposition 1 for the calculation of deviation benefit), then $(d_1^{AI},d_2^{AI})=(d_1^{SI},d_2^{SI})$. Otherwise, (d_1^{AI},d_2^{AI}) is solved by conditions (1) and (2) in Proposition (1). For each individual consumer i, the optimal choice of $g_i=max\{0,(\alpha-d_1^{AI}-d_2^{AI})/N\}$.

B.7 Personal Consumption Percentage

I describe in details how to use the benchmark input-output (IO) matrix from the Bureau of Economic Analysis to calculate personal consumption percentage. The procedures follow Becker and Thomas (2008) and Ahern and Harford (2014).

IO matrix provides the summary of producing and purchasing activities in U.S. based on data from the Economic Census. It is consisted of a make table and a use table. The make table is an industry by commodity matrix which gives the value in each commodity produced by each industry based on producer prices. The use table is a commodity by industry matrix which gives the value in each commodity used by each industry or final consumer (personal consumption, government) also based on producer prices. The link between the make table and the use table is commodity.

To measure personal consumption percentage of each industry, I start with the make table and calculate the market share of each commodity c that industry i produces. Industry $i's$ market share of commodity c is $share_{i,c} = make_{i,c}/\sum\limits_{k} make_{k,c}$, where $make_{k,c}$ is the value of commodity c produced by industry k from the make table. I then use the use table to calculate the total revenue that producer industry i generated from each user industry j (including final consumer). For industry i, the revenue from industry j is $rev_{i,j} = \sum\limits_{c} share_{i,c} \times use_{c,j}$, where $use_{c,j}$ is the value of commodity c used by industry j from the use table. Next, I calculate the percentage of industry $i's$ revenue from industry j as $rev\_pct_{i,j} = rev_{i,j}/\sum\limits_{i} make_{i,c}$. Finally, the personal consumption percentage of industry i is defined as $rev\_pct_{i,j}$ where j is personal consumption, which takes the value of F01000.

The industry defined in IO matrix is based on BEA's industry classification. BEA defines industries at two levels of aggregation, detailed and summary. I use detailed classification to calculate PCP for each industry. To convert BEA's industry to 6-digit NAICS, I use the concordance tables reported with the IO tables. For NAICS industries that have multiple corresponding BEA detailed industries, I calculate the PCP as the weighted average of PCP of corresponding BEA industries where the industry total output is used as weights.