University of
St Andrews

School of Economics & Finance Discussion Papers

Coordination in Public Good Provision: How Individual Volunteering is Impacted by the Volunteering of Others

Theodoros M. Diasakos and Florence Neymotin

School of Economics and Finance Discussion Paper No. 1402
7 Feb 2014
JEL Classification: H4, D8
Keywords: stochastic coordination, volunteer work, public goods

School of Economics & Finance
Discussion Paper Series
issn 0962-4031
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

Coordination in Public Good Provision: How Individual Volunteering is Impacted by the Volunteering of Others

Theodoros M. Diasakos * Florence Neymotin†

July 26, 2013

Abstract

In this analysis, we examine the relationship between an individual's decision to volunteer and the average level of volunteering in the community where the individual resides. Our theoretical model is based on a coordination game , in which volunteering by others is informative regarding the benefit from volunteering. We demonstrate that the interaction between this information and one's private information makes it more likely that he or she will volunteer, given a higher level of contributions by his or her peers. We complement this theoretical work with an empirical analysis using Census 2000 Summary File 3 and Current Population Survey (CPS) 2004-2007 September supplement file data. We control for various individual and community characteristics, and employ robustness checks to verify the results of the baseline analysis. We additionally use an innovative instrumental variables strategy to account for reflection bias and endogeneity caused by selective sorting by individuals into neighborhoods, which allows us to argue for a causal interpretation. The empirical results in the baseline, as well as all robustness analyses, verify the main result of our theoretical model, and we employ a more general structure to further strengthen our results.

Keywords: stochastic coordination, volunteer work, public goods

JEL Classification Numbers: H4, D8

*School of Economics and Finance, Castlecliffe, The Scores, University of St. Andrews, Fife, UK, KY169AR. Ph: +44(0) 1334 462425, td31@st-andrews.ac.uk

†Department of Finance and Economics, Nova Southeastern University, 3301 College Ave., Fort Lauderdale, FL, 33324, USA (corresponding author: fneymotin@nova.edu, 954-262-5339). We thank Rebecca Blank, David Card, Daniel Hamermesh, Esteban Jaimovic, Dong Li, Ignacio Monzon, Mario Pagliero, James Powell, and especially Giovanni Mastrobuoni for many helpful comments and suggestions. We are also grateful to Urmimala Sen for outstanding research assistance. All mistakes are our own.

1 Introduction

Economists have become increasingly interested in determining what motivates individuals to engage in volunteer work and thereby participate in what has been termed the "unpaid labor" sector (see Freeman[1], Govekar and Govekar [2], and Greisling [30] as a few examples). In contrast to earlier research, which considered unpaid labor to be indistinguishable from leisure, it has become increasingly common to model volunteering and related activities as a decision distinct in type and nature from the more general choice of engaging in leisure. Problems have also been empirically documented, primarily in small experiment settings, with an explanation of the choice to volunteer employing a purely public goods framework. A particular challenge comes from mounting evidence that, although individuals often engage in activities which are costly to themselves and, seemingly, primarily beneficial to others, the presence of individuals with other-regarding preferences does not suffice as an explanation. In the modern literature, this has been interpreted as suggesting the presence of an underlying private value from volunteering. This private value provides an individual motivation to acquire what has been termed variously "a warm glow," "prestige," or "self-worth."

Broadening the set of motives that influence pro-social behavior has led economists to ask how these motives interact with one another and with the underlying economic environment.[1] In some situations, the intrinsic motives of individuals seem to play such a vital role that providing rewards and punishments to foster pro-social conduct can actually crowd out individual contributions, the result being a reduction in both the size and number of individual contributions. In other settings, however, social pressure and norms favoring pro-social deeds and punishing selfish acts do seem to achieve their stated goal without causing crowd-out. To complicate matters further, most people value the perceived opinion others have of them (sometimes referred to as "self-image"), while also striving to maintain at least a minimal amount of consistency between their actions and their core values or beliefs. Clearly, the multi-dimensionality in how motivations interact with one other and with the economic environment dictate the need for additional research in this fairly complex arena.

Establishing a clear understanding of the aforementioned relationships is also fundamentally important to the task of identifying the personal and communal determinants of pro-social behavior, as well as comparing their relative strengths. Unfortunately, the actions and interactions of these drivers are contingent not only upon the context of the pro-social behavior in question (for example, contributing towards a public good, or volunteering one's labor or other resources, etc.) but also upon the specific social group from which they emanate and within which they operate. An examination of the effect the social group itself

[1]See the introduction of Benabou and Tirole [12] for a review of these issues.

exerts upon the pro-social conduct of its members is, therefore, equally imperative. This goal is the primary focus of our analysis within the context of engaging in volunteer work.

In the present study, we theoretically outline and empirically investigate an alternative explanation for the individual volunteering decision. Volunteering, as we see it, is an activity imposing costs on the individual who volunteers, but primarily benefiting other members of the community. In order to be observed at sufficiently high levels within a community, members of the community must successfully coordinate in a way that makes it more likely for any particular individual to contribute given higher expected levels of contributions by others in the community. We formalize this view using a model in which a group of agents each individually and simultaneously face the decision of whether or not to make a contribution to a public good. The need for coordination arises because the public good production technology requires a minimum amount of contributions for the good to be supplied. The usual economic problem with decentralized provision of public goods stems from the fact that, from the individual perspective, contributing does not make sense if that individual expects either too few, or more than enough, total contributions by other members in his or her community.

Despite this public good dilemma, we find that the presence of uncertainty regarding the required minimum amount of contributions suffices to support a symmetric pure-strategyε-Nash equilibrium in which the public good is supplied with positive probability. Perhaps even more interestingly, in this equilibrium the total number of individual contributions is positively correlated with the typical agent's decision to contribute. More precisely, the higher the number of contributions within her community, the more probable it is that any individual will choose to contribute. We see that, within the social coordination that generically takes place in our model, the expected number of individual contributions plays a crucial role. In conjunction with her private information, it allows the typical agent to form her posterior belief regarding the likelihood of provision of the public good. This in turn determines her expected marginal rate of substitution between contributing and not contributing. Since one's net benefit from contributing is not a priori known, beliefs about the total number of contributions are of course fundamentally important. In our model, an individual's net benefit from contributing depends on the actual desirability of the public good in question. This desirability is itself a function of a variety of factors including the degree of efficiency in the provision of the good (depicted in our setting by the required minimum amount of contributions), the good's value relative to other private or public projects that may be funded by individual resources, and the economic value of these resources, to name just a few examples. These are factors that our simple theoretical model fails to capture explicitly and this necessitates including a more general empirical investigation.

The empirical section focuses on the interaction between the characteristics of the community in which an individual resides and her decision to volunteer for a public good. The positive relationship between an individual's propensity to volunteer and the total amount of contributions in the community is clearly supported by the data. Our empirical results are robust to various (empirical) modeling specifications, and the inclusion or exclusion of community and individual control factors. We address the issue of reflection bias and endogeneity caused by selective sorting of individuals by neighborhood through the use of an innovative instrumental variables strategy allowing for a causal interpretation of our results. Our instruments make use of both previous levels of volunteering in the area to instrument for current levels, as well as instrumenting nonreligious individual volunteering with average levels of religious volunteering in the community. Taken together, in the main analysis as well as in our various robustness checks, our findings are clearly consistent with our theoretical premise. We take this as evidence for the consistency of results between our theoretical model and a more general and less stylized empirical reality.

To the best of our knowledge, we are the first within the realm of the volunteering and public goods literature to use an informational perspective to model the interaction between an individual's propensity to engage in pro-social behavior and the propensity to engage in pro-social behavior of others in the individual's community. Our theoretical framework is very similar to a standard structure used when discussing speculative exchange-rate attacks and bank- or asset-runs. One key difference between our model and the aforementioned models in the speculative exchange-rate and bank- or asset-runs literature relates to how the actions of others affect an individual's payoffs. In the aforementioned literature, the actions of others only affect one's payoff indirectly by carrying information regarding the underlying state of the world. In contrast, in our structure the actions of others also matter directly, since sufficient contributions by others in the community to provide the public good mean that an individual should choose not to contribute. Empirically, our use of Census data linked to CPS data allows us to employ the largest and most representative sample of recent volunteering which is presently available. It is also true that our use of an innovative instrumental variables structure and clear statistical testing allow for a more rigorous and causal interpretation than previously existing in the empirical literature in this area. In this sense, the novelty of our investigation lies in the extent to which it is comprehensive, causally interpretable, and allows us to be confident in our results.[2]

[2] See Brudney and Gazley [16] for a discussion of the problems that plague data-gathering regarding volunteering.

The remainder of our analysis is organized as follows: We summarize and review the relevant literature in Section 2. We outline our theoretical analysis and its implications in Section 3. For the empirical portion of our analysis, we present and discuss our data, structure and results in Section 4. We summarize our results and overarching conclusions in Section 5.

2 Related Literature

In the current analysis, we chose the model and empirical structure with the goal of understanding how the volunteering of others affects an individual's propensity to volunteer. Before we can fully convey the implications of our analysis, however, we must first place it in context of a discussion of the literature on individually-driven but socially-affected motivations for individual volunteering.[3] The economics literature in this area has made some headway, and now offers some fairly established theories regarding why individuals volunteer in a general sense, as well as why they volunteer specifically in the provision of public goods. [4]

Traditionally, economists considered volunteering to be a purely public good and, as such, noted that it should be subject to the problem of the commons. Namely, an individual's propensity to volunteer should decrease as the number of contributions by others increase. For this same reason, private contributions should be crowded out by governmental provision of a public good (see for instance Bergstrom et al. [11] for early work on this topic).

The empirical evidence to date has not borne out this simplistic view of the act of volunteering as a purely public good. Specifically, neither of the previous assertions regarding the problem of the commons and its results for the provision of public goods has been

[3]We do not address the issue of prescribed volunteering, such as service learning requirements often instituted in junior high or high school. This area of analysis is currently absent in the existing literature and would serve as an interesting and useful extension to the current work. For an introduction to service learning see for example McGoldrick et al. [35]

[4]We note as we begin this discussion that many of the motivations to be discussed could easily refer to donations of one's material resources to charity, however, they could equally well relate to donations of one's time or effort to a common good or cause. This is not particularly surprising, since volunteering has been analyzed as being a substitute for, and also a complement to charitable monetary contributions. Nonetheless, some of the reasons an individual might volunteer his or her time are quite distinct from those dictating the choice to give charitable monetary contributions. For instance, she may want to volunteer for certain tasks in order to increase her human capital by either providing her with skills complementing those she uses while employed in the paid labor sector or, more importantly, by providing her with contacts and work experience during times of unemployment. This reason is clearly unique to volunteering of one's time rather than monetary contributions. Other reasons given for volunteering, in contrast to reasons given for monetary contributions, are increases in self-worth through higher levels of productivity and decreases in social isolation-particularly after retirement.

demonstrated in a consistent fashion. Recent empirical evidence even shows that, consistent with this paper's structure, albeit using small experimental studies, there may in fact be a positive relationship between the number of individual contributions and an individual's propensity to contribute. Indeed, the well-known studies in Andreoni [10] and [8] document using small experimental studies, that the provision of a public good appears to increase with the concurrent level of giving in the community, and that private contributions are not necessarily crowded out by governmental provision of the public good. One explanation given for this type of result, albeit with respect to a very specific kind of public good, is offered by Bilodeau and Slivinski [13]. In their work, these authors use the premise that private valuations of the public good may differ across individual members of a group.

A more general possibility which supports these initial empirical results, however, is the possibility that individuals are in fact impure altruists. In this view, the individual contributing to the public good not only benefits from using the public good which is ultimately created, but the very act of contributing to the public good offers a private benefit. Several

of these additional private valuations are discussed in the sense of individuals gaining a "warm-glow," "prestige" or "moral superiority" from the very act of contributing to a public good.

The "warm glow" characterization of volunteering is fairly straightforward in assigning individuals a private benefit from the act of volunteering, but it is not very complex in how or why this warm-glow is formed or in relating the benefits as intrinsically socially determined [6]-[9]. In contrast, some scholars take the approach that individuals contribute (money) because they would like to be considered "philanthropists" and, thereby, gain prestige within their community [31] [28]. The number of other contributors as well as the timing of when an individual contributes relative to others in the community will also affect the private prestige benefit. Higher contributions by others render the cause for which the individual is called upon to donate more prestigious, and, thereby, increase the possible level of prestige an individual can gain by contributing. Additionally, being an early contributor would gain the individual more prestige and attention than falling within the dense ranks of later, less-visible donors.[5] Finally, in the "moral stance" theory, individuals are viewed as volunteering because they believe free-riding to be morally wrong and, consequently, feel an obligation to contribute to the public good. In this structure, put forth by Sugden [46], donations to the public good offer individuals the socially-determined, private benefit of establishing themselves as moral human beings. It is also true that each of these explanations is related to the feeling of "making a difference," which seems to be linked to the number of others

[5] For a general discussion on this topic, see Carman [18] as well as Knox [33]. See also Andreoni [6] for an excellent introduction to the issues and research on volunteering and charitable giving.

who acting in a similar fashion (Francois [26]). The above overview of the possible reasons for private valuations within the very act of contributing highlights the importance of social influences in establishing private valuations.[6] We next discuss a variation which is more closely related theoretically to the form employed in the present analysis.

The signaling structure examines the social context of volunteering in terms of the information conveyed by the volunteering of others. For instance, instead of being motivated to gain prestige by early contributions to a charity, individuals may instead choose to be later contributors in order to learn more about the intrinsic value of the public good in question. In this context, later contributions make sense since previously observed donations can serve as a signal regarding the intrinsic value of the public good. This type of signaling effect was documented experimentally by Potters et al. [42] who found that sequential contributions to a public good can be useful when there is uncertainty regarding its true value. More precisely, this study suggests that agents who are informed regarding the underlying quality of the public good should make their decisions first so as to provide the uninformed agents with a signal (see also Andreoni [6] for a related finding). This type of signaling framework also explains why charities which tend to announce early contributions along with the related amounts of these contributions generally receive more total contributions than charities which do not make these types of announcements (Vesterlund [51]).

In our model, we abstract from an explicit account of intrinsic private valuations in order to instead focus on the underlying social interactions in the decision to volunteer. Our theoretical setup is very similar to the one in Nirei et al. [41]. Specifically, our agents also employ Bayesian learning and, in equilibrium, choose whether or not to take an action based on the information conveyed by their own private signal as well as the expected actions of others. The equilibrium strategy exhibits complementarity, since each agent is more likely to contribute when the aggregate number of contributions in the community is larger. As a result, some fraction of the agents ends up synchronizing their actions. The size of this fraction is determined by the realization of the agents' signals via a threshold-based switching strategy, as in Morris and Shin [38]. Here, however, the threshold in question is affected by the actions of others because these actions affect one's expected payoff. In fact, one of the goods being public in the present context, the underlying complementarities are not only strategic, but also ones that affect the payoffs directly (which makes the analysis rather more involved).

[6] For a comprehensive overview of the economic relations within social interactions, see Blume et al. [15].

3 The Model

In what follows, we consider a community of $N \in \mathbb{N}^*$ individuals (indexed by $n = 1, \ldots, N$) in which a public good may be provided according to the production technology $h_0 : \mathbb{R}_+ \mapsto \mathbb{R}_+$ given by

$$h_0(X) = \begin{cases} 0 & \text{if } X < e^{\theta_0} \\ h(X) & \text{if } X \geq e^{\theta_0} \end{cases}$$

for some function $h : \mathbb{R}_{++} \mapsto \mathbb{R}_+$. Here, X denotes a single input which must be available in at least some minimum quantity for the public good to be produced. To keep things analytically tractable, we take this minimum to be log-normally distributed: $\theta_0 \sim \mathcal{N}(\vartheta, \sigma^2)$. Regarding h, we make the following assumptions:

Assumption A.1 The function $h : \mathbb{R}_{++} \mapsto \mathbb{R}_+$ is twice continuously differentiable, and satisfies

- (i) $\lim_{X \to 0} h(X) = 0$
- (ii) Strict monotonicity: $h'(X) > 0 \;\; \forall X \in \mathbb{R}_+$
- (iii) Convexity: $h''(X) \geq 0 \;\; \forall X \in \mathbb{R}_+$

The input in question results from the collective contributions of the community members who all have identical von-Neumann Morgenstern preferences for consumption of the public good and all other commodities, and an endowment w units of private consumption. The individual preferences are given by the Bernoulli utility function $U : \mathbb{R}^2_+ \mapsto \mathbb{R}$, which satisfies the following conditions:

Assumption A.2

- (i) The function U is strictly increasing in either argument on $\mathbb{R}^2_+$
- (ii) As the input to the public good production process becomes arbitrarily large, the marginal rate of substitution between private and public consumption tends to a limit which is dominated by that of the marginal physical product of X

$$\lim_{X \to +\infty} \left[h'(X) - \frac{U_1(y, h(X))}{U_2(y, h(X))} \right] > 0 \quad \forall y \in (0, w)$$

- (iii) As the input to the public good production process becomes arbitrarily large, the typical agent views private and public consumption as (weak) substitutes, and is (weakly) risk-loving with respect to the latter

$$\lim_{X \to +\infty} U_{12}(y, h(X)) \leq 0 \leq \lim_{X \to +\infty} U_{22}(y, h(X)) \quad \forall y \in (0, w)$$

These restrictions refer to a setting in which the public good production technology exhibits non-decreasing marginal product and, at least in the limit (for a sufficiently large amount of the public good), the marginal rate of substitution between private and public consumption becomes non-increasing with respect to the latter. To fix ideas, suppose that the individual members of a community are called upon to volunteer their time and effort in order to clean up some commons (a neighborhood park, a beach, etc.), to improve the quality and effectiveness of the educational, cultural, or recreational activities their offspring enjoy at a local center (school, summer camp, church, etc.), or to enhance the prospects of a socio-political or cultural campaign. In each of these cases, we focus on the interaction between the public good in question and a private one which is a substitute, at least as long as the public good is available in large enough quantities. For instance, when the latter condition represents a public park that is sufficiently large, beautiful, accessible, and well-maintained, our theoretical platform views one's private garden within a nearby property as an amenity that becomes increasingly dispensable. Similarly, when the frame of reference is a grassroots socio-political movement, we restrict our attention to those characteristics that render participation rates as an input with increasing marginal product. For instance, the movement can lead to significant voter registration for a political party locally, which in turn may help at the state and national levels.

In each of these contexts, we assume that the members of the community are called upon to decide individually but simultaneously whether or not to make a fixed contribution of x units of private consumption towards the production of the public good. When the typical agent is called upon to act, but prior to her actual choice, she receives some information about the minimum amount of the productive input the available technology requires. Specifically, she observes the private signal $\theta_n = \theta_0 + \epsilon_n$ where the error terms are i.i.d. across the agents, according to the normal distribution $\mathcal{N}(0, \sigma_\epsilon^2)$ truncated at the interval $[-c, c]$, for some given $c \in \mathbb{R}^*$, and such that $\mathrm{Cov}(\theta_0, \epsilon_n) = 0$.[7]

[7] In other words, the probability density function of the private signal is given by $\frac{\phi\left(\frac{\theta_n - \theta_0}{\sqrt{\sigma_\epsilon^2}}\right)}{\sqrt{\sigma_\epsilon^2}\left[\Phi\left(\frac{c}{\sqrt{\sigma_\epsilon^2}}\right) - \Phi\left(\frac{-c}{\sqrt{\sigma_\epsilon^2}}\right)\right]}$ where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

The Equilibrium

Obviously, in this simple setting, the typical individual who expects exactly k other agents to contribute will choose to make her own contribution if and only if she believes it is pivotal for the production of the public good. More precisely, given her signal θ_n and her belief that k other individuals will contribute, her payoff is given by

$$\Pr\left[e^{\theta} \leq (k+1)\,x|\theta_n, k\right] U\left(w_n - x, h\left((k+1)\,x\right)\right) + \Pr\left[e^{\theta} > (k+1)\,x|\theta_n, k\right] U\left(w_n - x, 0\right)$$

if she contributes and

$$\Pr\left[e^{\theta} \leq kx|\theta_n, k\right] U\left(w_n, h\left(kx\right)\right) + \Pr\left[e^{\theta} > kx|\theta_n, k\right] U\left(w_n, 0\right)$$

if she doesn't. She will contribute, therefore, if and only if

$$\begin{aligned} 0 \;\leq\; & \Pr\left[e^{\theta} \leq kx|\theta_n, k\right] \left[U\left(w_n - x, h\left((k+1)\,x\right)\right) - U\left(w_n, h\left(kx\right)\right)\right] \\ +\; & \Pr\left[e^{\theta} > kx|\theta_n, k\right] \left[U\left(w_n - x, 0\right) - U\left(w_n, 0\right)\right] \\ +\; & \Pr\left[kx \leq e^{\theta} \leq (k+1)\,x|\theta_n, k\right] \left[U\left(w_n - x, h\left((k+1)\,x\right)\right) - U\left(w_n - x, 0\right)\right] \end{aligned} \tag{1}$$

Equivalently (recall Assumptions A.1(ii) and A.2(i)), if and only if

$$V\left(\theta_n, k\right) + f\left(\theta_n, k\right) \geq 0$$

where

$$\begin{aligned} V\left(\theta_n, k\right) \;=\; & \frac{\left[U\left(w_n - x, h\left((k+1)\,x\right)\right) - U\left(w_n, h\left(kx\right)\right)\right]}{\left[U\left(w_n - x, h\left((k+1)\,x\right)\right) - U\left(w_n - x, 0\right)\right]} \\ +\; & \frac{\Pr\left[e^{\theta} > kx|\theta_n, k\right]}{\Pr\left[e^{\theta} \leq kx|\theta_n, k\right]} \left(\frac{U\left(w_n - x, 0\right) - U\left(w_n, 0\right)}{U\left(w_n - x, h\left((k+1)\,x\right)\right) - U\left(w_n - x, 0\right)}\right) \end{aligned}$$

and

$$f\left(\theta_n, k\right) = \frac{\Pr\left[kx \leq e^{\theta} \leq (k+1)\,x|\theta_n, k\right]}{\Pr\left[e^{\theta} \leq kx|\theta_n, k\right]}$$

To investigate the typical agent's strategic behavior, we restrict our attention to obtaining a symmetric ε-Nash equilibrium in which she follows the strategy of contributing if and only if her private signal realization does not exceed a critical value. Of course, the cutoff in question cannot help but depend on the surmised number of individual contributions so that, if a total of k agents make contributions in equilibrium, it is because their signals and only their signals fell below some value $\theta(k)$. It follows, therefore, that we have to verify the following premise: given her realization θ_n and her belief that exactly k other agents will contribute, it is optimal for the typical agent to contribute if and only if $\theta_n \leq \theta(k)$, as long as every other player follows the same strategy.

We do this by reasoning in steps. First, we show that the premise in question is valid in the limit, for a sufficiently large number of contributions. Then, we establish that, also in the limit, the cut-off function $\theta(\cdot)$ is monotone. Given, however, that either of the previous steps is valid for an arbitrary fixed size x of individual contributions, the problem at hand can be transformed to one in which $\theta(\cdot)$ is monotone over all possible values of k. As a result, the desired equilibrium exists even without the $k \to +\infty$ qualification.

Regarding the first step of the reasoning outlined above, it turns out that we may ignore the last term on the right-hand side of (1) without loss of generality. Intuitively, this term evaluates the net welfare gain of a typical contributor when the public good gets produced. The model weights it, however, according to the likelihood that she is pivotal for the production process, an event which becomes null in the limit as the number of contributions grows.

Proposition 1 *Let the variables θ_n, θ_0 and the quantity $V(\theta_n, k)$ be defined as above. The following are equivalent*

- *For a given $\varepsilon > 0$, there exists $k_1 \in \mathbb{N}$ s.t.* $V(\theta_n, k) + f(\theta_n, k) \geq -\varepsilon \quad \forall k \in \mathbb{N} : k \geq k_1$
- *For a given $\varepsilon > 0$, there exists $k_2 \in \mathbb{N}$ s.t.* $V(\theta_n, k) \geq -\varepsilon \quad \forall k \in \mathbb{N} : k \geq k_2$

In the limit, therefore, we may restrict attention to an ε-Nash equilibrium in which the typical individual contributes when she expects k other members of her community to do so if and only if $V(\theta_n, k) \geq -\varepsilon$. And this limit is independent of the individual's private signal.[8] Moreover, given the following result

Lemma 1 *Let*

$$\Delta U(w, x, k) = U(w - x, h((k+1)x)) - U(w, h(kx))$$

Under assumptions A.1(ii) and A.2(ii), we have $\lim_{k \to +\infty} \Delta U(w, x, k) > 0$.

Proof. See Appendix A.

[8]Indeed, neither k_1 nor k_2 depend on the realization θ_n

there will be some $k_3 \in \mathbb{N}$ with $k_3 \geq max\left[k_1, k_2\right]$ such that, for all $k \geq k_3$, the decision rule in question may be re-written as follows

$$\frac{\Pr\left[e^{\theta} \leq kx|\theta_n, k\right]}{\Pr\left[e^{\theta} > kx|\theta_n, k\right]} \geq \frac{U\left(w_n, 0\right) - U\left(w_n - x, 0\right)}{\Delta U\left(w, x, k\right) + \varepsilon\left[U\left(w_n - x, h\left(\left(k+1\right)x\right)\right) - U\left(w_n - x, 0\right)\right]} \tag{2}$$

Let us turn now to the left-hand side of (2). On the one hand, Bayes' rule gives

$$\frac{\Pr\left[e^{\theta_0} \leq kx|\theta_n, k\right]}{\Pr\left[e^{\theta_0} > kx|\theta_n, k\right]} = \frac{\Pr\left[e^{\theta_0} \leq kx, \theta_n, k\right]}{\Pr\left[e^{\theta_0} > kx, \theta_n, k\right]} = \frac{\Pr\left[k|\theta_n, e^{\theta_0} \leq kx\right]\Pr\left[\theta_n, e^{\theta_0} \leq kx\right]}{\Pr\left[k|\theta_n, e^{\theta_0} > kx\right]\Pr\left[\theta_n, e^{\theta_0} > kx\right]}$$

while, on the other hand, the hypothesized typical response rule requires that

$$\Pr\left[k|\theta_n, e^{\theta_0} \leq kx\right] = F\left(\theta\left(k\right), k\right)^{N-k-1}\widetilde{F}\left(\theta\left(k\right), k\right)^k$$

where we deploy the following likelihoods

$$\begin{aligned}
F\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n > \theta\left(k\right)|e^{\theta_0} \leq kx\right] \\
\widetilde{F}\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n \leq \theta\left(k\right)|e^{\theta_0} \leq kx\right] = \Pr\left[e^{\theta_0} \leq kx\right] - F\left(\theta\left(k\right), k\right)
\end{aligned}$$

And similarly,

$$\Pr\left[k|\theta_n, e^{\theta_0} > kx\right] = G\left(\theta\left(k\right), k\right)^{N-k-1}\widetilde{G}\left(\theta\left(k\right), k\right)^k$$

where

$$\begin{aligned}
G\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n > \theta\left(k\right)|e^{\theta_0} > kx\right] \\
\widetilde{G}\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n \leq \theta\left(k\right)|e^{\theta_0} > kx\right] = 1 - \Pr\left[e^{\theta_0} \leq kx\right] - G\left(\theta\left(k\right), k\right)
\end{aligned}$$

It then follows that the left-hand side of (2) can be expressed more succinctly as

$$\frac{\Pr\left[e^{\theta_0} \leq kx, \theta_n, k\right]}{\Pr\left[e^{\theta_0} > kx, \theta_n, k\right]} = \delta\left(\theta_n, k\right)A\left(\theta\left(k\right), k\right)^{N-k-1}B\left(\theta\left(k\right), k\right)^k$$

with

$$\begin{aligned}
A\left(\theta\left(k\right), k\right) &= \frac{F\left(\theta\left(k\right), k\right)}{G\left(\theta\left(k\right), k\right)} \qquad B\left(\theta\left(k\right), k\right) = \frac{\widetilde{F}\left(\theta\left(k\right), k\right)}{\widetilde{G}\left(\theta\left(k\right), k\right)} \\
\delta\left(\theta_n, k\right) &= \frac{\Pr\left[\theta_n, e^{\theta_0} \leq kx\right]}{\Pr\left[\theta_n, e^{\theta_0} > kx\right]}
\end{aligned}$$

To verify the typical agent's best response as dictating that she contributes if and only if her signal does not exceed the cutoff value, it suffices to show that, other things being equal in terms of k and the other players' signal realizations $\{\theta_1, \ldots, \theta_{n-1}, \theta_{n+1}, \ldots, \theta_N\}$, the likelihood ratio $\frac{\Pr\left[e^{\theta_0} \leq kx, \theta_n, k\right]}{\Pr\left[e^{\theta_0} > kx, \theta_n, k\right]}$ is decreasing in θ_n. Yet, the latter variable affects this ratio only through the likelihood ratio $\delta(\theta_n, k)$. It is enough, therefore, to show that $\delta(\theta_n, k)$ is decreasing in θ_n. Which is indeed the case, in fact, for any given $k \in \mathbb{N}$.

Proposition 2 *Let $kx \in \mathbb{R}_{++}$ be fixed and θ_n, θ_0 be defined as above. The likelihood ratio*

$$\delta(\theta_n, k) = \frac{\Pr\left[\theta_n, e^{\theta_0} \leq kx\right]}{\Pr\left[\theta_n, e^{\theta_0} > kx\right]}$$

is such that $\delta_1(\theta_n, k) < 0$.

Proof. See Appendix A.

Next, we need to identify the cutoff schedule $\theta : \mathbb{N} \mapsto \mathbb{R}$ and show that the strategic rule in question supports an ε-Nash equilibrium. Adopting the usual convention that, if an agent is indifferent between making a contribution or not, she will choose to contribute, $\theta(\cdot)$ gets defined implicitly by the binding version of (2) for $\theta_n = \theta(k)$. Formally,

$$\begin{aligned} & \frac{U(w_n, 0) - U(w_n - x, 0)}{\Delta U(w, x, k) + \varepsilon\left[U(w_n - x, h((k+1)x)) - U(w_n - x, 0)\right]} \\ = \; & \delta(\theta(k), k) A(\theta(k), k)^{N-k-1} B(\theta(k), k)^k \end{aligned} \tag{3}$$

Recall now that we are seeking a symmetric ε-Nash equilibrium under the strategic premise that each player who expects k contributions from the rest of the community will herself contribute if and only if her private signal does not exceed the value $\theta(k)$. Since this scenario entails by construction each and every agent playing her best-response to the strategic profile of the rest of the community, to obtain an equilibrium, it remains only to ensure that the players' beliefs about the total number of contributions are correct. We need to show, therefore, that, for any realization of signals $\widetilde{\theta} = (\theta_1, \ldots, \theta_N) \in \mathbb{R}^N$, there exists some $k_{\widetilde{\theta}} \in \mathbb{N}$ such that exactly $k_{\widetilde{\theta}} \in \mathbb{N}$ entries from $\widetilde{\theta}$ do not exceed the cutoff $\theta\left(k_{\widetilde{\theta}}\right)$. To this end, we will make use of another result and a well-known fixed point theorem.

Proposition 3 *Let the mapping $\theta : \mathbb{N} \mapsto \mathbb{R}$ be defined implicitly by (3). Then, $\exists k_4 \in \mathbb{N}$ s.t. the restricted mapping $\theta : \mathbb{N} \setminus \{1, \ldots, k_4\} \mapsto \mathbb{R}$ is strictly increasing.*

Proof. See Appendix A.

To fix ideas, our analysis thus far ensures that, for $k_5 = \max\{k_i\}_{i=1}^4$ and *any* signal realization $\widetilde{\theta}$, the symmetric strategy profile in which each individual contributes to the public good if and only if her private signal does not exceed the cut-off $\theta(k)$ consists of best-responses, while the restriction of the mapping θ to the set $\mathbb{N}\setminus\{1,\ldots,k_5\}$ is guaranteed to be strictly increasing. Moreover, we established these facts for arbitrary fixed, size x of individual contributions. Hence, to obtain monotonicity over the entire domain $\mathbb{N}$, we may consider a change of variables $x' = k_5 x$. Under this transformation, since $(n+1)x' = (n+1)k_5 x \geq k_5 x$ for all $n \in \mathbb{N}$, the entire cutoff mapping $\theta : \mathbb{N} \mapsto \mathbb{R}$ becomes strictly increasing.

We may then define the correspondence $\Theta : \mathbb{R}^N \times \{0,\ldots,N\} \mapsto \{0,\ldots,N\}$ so that, for any signal realization $\widetilde{\theta}$, $\Theta\left(\widetilde{\theta},k\right)$ is the number of entries in $\widetilde{\theta}$ that do not exceed $\theta(k)$. Obviously, since $\theta(\cdot)$ is increasing, so is $\Theta\left(\widetilde{\theta},\cdot\right)$ which, in turn, must have a fixed point by Tarski's theorem (see Tarski [47]). For any given $\widetilde{\theta}$, therefore, there will indeed be some $k_{\widetilde{\theta}} \in \{0,\ldots,N\}$ such that exactly $k_{\widetilde{\theta}}$ signals are not larger than $\theta\left(k_{\widetilde{\theta}}\right)$. Put differently, there will indeed be an equilibrium, for every realization of signals.

We are finally in a position to conclude our theoretical investigation by pointing out our most important implication of the equilibrium under examination. Namely, that the cutoff mapping $\theta(\cdot)$ being strictly increasing necessitates an ex-ante probability of typical contribution that is strictly increasing in the total number of contributions across the community. Indeed, the probability in question is given by

$$\begin{aligned}
\Pr\left[\theta_n \leq \theta(k)\right] &= \Pr\left[\theta_n \leq \theta(k), e^{\theta_0} \leq kx\right] + \Pr\left[\theta_n \leq \theta(k), e^{\theta_0} > kx\right] \\
&= \int_{\theta_0 - c}^{\theta(k)} \int_{-\infty}^{\ln kx} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)} \mathrm{d}\theta_0 \mathrm{d}\theta_n \\
&\quad + \int_{\theta_0 - c}^{\theta(k)} \int_{\ln kx}^{+\infty} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)} \mathrm{d}\theta_0 \mathrm{d}\theta_n \\
&= \int_{\theta_0 - c}^{\theta(k)} \int_{-\infty}^{+\infty} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)} \mathrm{d}\theta_0 \mathrm{d}\theta_n \\
&= \int_{-\infty}^{+\infty} \left(\int_{\theta_0 - c}^{\theta(k)} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)} \mathrm{d}\theta_n \right) \mathrm{d}\theta_0
\end{aligned}$$

and the inner-integral is strictly increasing in k given that so is $\theta(k)$.

4 Empirical Analysis

Our theoretical model delivers strong implications. It provides a structure in which communities may form stochastic herds of public-good contributions based on the fact that an individual expects many others in the community to contribute, and this expectation enhances the individual's perceived expected benefit from themselves contributing. It is also true, however, that our model is fairly stylized, hinging heavily on the premise that total, community-wide contributions for the production of the public good are sufficiently large. More precisely, we require that either the total number or the average size of individual contributions are sufficiently large. These conditions, which are duals of one another, presuppose that the community under study is either sufficiently numerous or else willing and able to contribute sufficiently large amounts to the public good. Because of this, a complementary, non-structural, empirical investigation is required.

4.1 Data

The empirical portion of our analysis made use of two data sets: the 2004-2007 September supplements to the Current Population Survey (CPS), and the Census 2000 Summary Files (STF3). The CPS and Census files were matched using Core Based Statistical Areas (CBSA's) via a county-level matching procedure. Geographic information after the match was retained for the CBSA level of analysis and may be more properly considered as the "CBSA-area" level of analysis.[9] An additional, and easily met, restriction which we used was that at least 10 individuals had to be found in the CBSA-area within our sample.

The CPS September supplement is unique with respect to other CPS supplements in its focus on questions related to individual volunteering. It contains information on individual-level demographic characteristics as well as indicators for whether an individual chose to

[9]At first, we attempted to match the two datasets entirely at the county level using county of residence information in the CPS. This strategy was abandoned due to the large number of individuals for whom county information in the CPS was unavailable. Instead, we adopted the method of matching the CBSA of residence in the CPS data to the corresponding county(ies) in the Census data. This matching technique alleviated the problem of missing CPS residence information significantly, allowing us to keep a substantially larger fraction of the data. We use the term "CBSA area" to reflect the fact that we averaged information across all counties within a particular CBSA. Unfortunately, in order to maintain consistency of measurements for the merged years, we had to focus only on the years 2004-2007, as these are the ones for which CBSA information is available in the CPS. The 2002-2003 September Volunteering Supplements do have county and Metropolitan Statistical Area (MSA) information, however, they do not provide information on CBSA of residence, since CBSA was not yet in use at that time in the supplements. It should also be noted that county information was generally unavailable for the CPS observations in the New England states. For these states, a New England City and Town Area (NECTA) to CBSA match was virtually impossible since there is a many-to-many relationship between NECTA's and CBSA's, even before accounting for county locations.

volunteer, along with the type of volunteer organization that was chosen. Our analysis focused on working-age adults (ages 25 to 65). We used this restriction because, as Mutchler et.al. [40] pointed out, it is possible that retirees or students volunteer for different reasons than do individuals currently in the workforce.[10] Regarding its time span, our baseline analysis focused primarily on the last two years of our data (2006-2007). These two years were used in the baseline regressions in order to make the time frame comparable between the baseline and the instrumental variables regressions. The instrumental variables regression, by their nature, employed all four years of data in the analysis.[11]

The CPS information employed in our empirical investigation included individual-level demographic characteristics and information on whether an individual chose to volunteer, as well as the type of organization for which he or she chose to volunteer. Specifically, from the CPS data, we used information on an individual's gender, age, race, educational attainment, family income, family structure and size, as well as marital and employment status.[12] These variables will henceforth be collectively referred to as $DEMOG$. Also from the CPS data set, we imported a binary variable denoting whether or not an individual volunteered, and noted also the particular organizations for which he or she volunteered, paying particular attention to whether or not he or she volunteered for a religious organization.[13] Our empirical analysis did not employ the number of hours volunteered. This dimension of an individual's contribution decision is not captured by our theoretical structure and, to maintain rigor and consistency, our empirical investigation stays as much as is possible within the limits of our theoretical intuition. Specifically, the extensive, rather than intensive, margin poses more interest for our particular theoretical analysis.

Our intuition for including the $DEMOG$ control characteristics in our analysis is supported by substantial evidence in the literature. We next consider a few of these points, although this list is far from exhaustive in nature. Gender: It has been documented that, on average, females volunteer more and are thought to be differentially altruistic, at least in some settings, than males (Simmons and Emanuele [45]). Wealth and Education: These are considered to be intimately associated with an individual's likelihood of being civic-minded

[10]Further analyses showed similar results (available upon request) for the full population of individuals aged 15 and older.

[11]Employing the larger four-year sample in the baseline regressions did not change the substance of our results.

[12]To proxy family structure, we employed variables indicating the presence of children of various ages in the home, as well as marital status. The variables used to depict marital status are coded as: never married, married and spouse present, married and spouse absent, and divorced or separated.

[13]More precisely, we used the responses to the question in the survey asking about the organization for which the individual had volunteered. Answering that one had volunteered but not for a religious organization renders one a "nonreligious" volunteer.

and engaging in volunteer activities (Dee [21]). Family Structure and Size: Individuals with children, and in particular small children, are more likely to volunteer, at least for some types of organizations. Employment Status: This characteristic is important because of the possibility that individuals are either trading off volunteering with paid work in an effort to increase human capital or life satisfaction or instead, according to average statistics, such as in Kulik [34], the employed are also volunteering at higher rates. Race: This variable is often implicated as affecting not only the choice of volunteering but also the type of the chosen volunteer activity. This effect is especially strong in studies of religious volunteering, such as in Musick *et.al.* [39].

From the Census data, we gathered information on the average characteristics of the CBSA area. Specifically, we included as controls the total population size (and its square), the average CPI-adjusted income level (and its square), and the fraction living in an urban location in the CBSA area.[14] We also employed the fraction of various racial groups in the CBSA-area as part of the area-level controls. In what follows, these location-related variables will be collectively referred to as $CBSACHAR$.[15]

Finally, regarding the average level of volunteering in the community, we created several CBSA-area measures of this variable. We note that these measures of average volunteering in the community correspond to the focal variable of our theoretical analysis (the number k of contributions from others adjusted for community size). We created the average (leave-out-mean) of individuals volunteering "generally," i.e. religious or non-religious, using current year data, and, alternatively, using a two-year lagged version. We repeated this construction of the average, as well as the two-year lag, for both the specific type of volunteering, which we considered "religious" i.e. volunteering for religious organizations, as well as for volunteering which was "non-religious" in nature. We chose a two-year lag as our time-frame for several reasons. First, we wanted to employ a measure which did not confound the effects from repeated observation from the outgoing rotation groups in the CPS, and second, we wanted to use a lag which was plausible as an instrumental variable strategy but retained more than

[14]We used variants of our fraction urbanized vs. total population size and found no substantive difference from including both in the regressions. We chose to include both population size and fraction urbanized so that we could distinguish larger areas from simply different types of areas by urban-rural distinction.

[15]We experimented with regression specifications including measures of fractionalization (by race, immigrant, and income inequality using Gini coefficients) at the CBSA area-level as control variables. These were constructed in the same way as the ethno-linguistic type of fractionalization measures. Empirically, these required a slightly less favored measure of race for the individual-level, since our probit models using the same racial structure as in this analysis did not achieve numerical convergence. Nevertheless, also in these alternative race-coded models, the average level of volunteering in the community has a positive effect on an individual's decision to volunteer, both in the baseline OLS/probit regressions as well as in their respective instrumented versions. This was true with respect to both general volunteering and for non-religious volunteering. Notice that general volunteering includes both religious and non-religious volunteering.

a single year of data for the present analysis.[16] [17]

Summary Statistics

Table 1 displays the mean, minimum, and maximum of each variable in our analysis for the full population of individuals in the CPS September Supplement Files (2004-2007).[18] Means are shown for the 2004-2007 merged-year sample as well as separately by each of the four included years. Annual results are quite similar to the merged-year results, albeit with a few exceptions. In particular, the fraction of individuals with a family income above $75,000 varies from a low of 29% in 2004 to a high of 35% in 2007.[19] Variation over time is also present in the fraction who volunteer, going from a high of 30% in 2004 to a low of 27% in 2007. We feel that there may have been a break in volunteer activities with a greater number of individuals choosing to volunteer at the beginning of the 2004-2007 period due to outside environmental factors. There is also some small variation in the composition of individuals based on education level and age. Overall, however, the similarities in the summary statistics of our sample over the relevant time period are such that they do not point to any significant biases in the data.

Table 2 displays the breakdown of volunteering by category among the individuals who chose to volunteer. [20] It is clear from this table that volunteering for religious organizations,

[16]See for example Isham *et.al.* [32], Tao and Yeh [48], and Segal *et.al.* [44].

[17]In our theoretical model, we assume that once an individual decides to volunteer his or her time for a common good or cause, the amount of time which is offered is exogenously given. Although the intensive margin is obviously important, we have chosen not to focus on it because it depends on factors such as the opportunity cost of one's time as one example. The concern is that these types of factors may not be directly related to information regarding the quality of the public good or the importance of the volunteer cause. Within our theoretical framework, this information in particular is what helps us discern the positive relationship between average community levels of volunteering as affecting individual choices to volunteer. Therefore, we have chosen to focus on the extensive, rather than the intensive margin, in order to allow the theoretical and empirical aspects of the analysis to retain a closer compatibility.

[18]The regressions in Tables 3-5 and Appendix Tables 1-2 refer to an age-restricted sample (age 25-65). We chose to focus on the full sample in this table in order to show generality in the reported results. It is also true, however, that the age-restricted sample shows a similar pattern of results to those shown in Table 1.

[19]Since it was reported in a categorical fashion, individual income data from the CPS files could not reasonably employ a CPI-adjustment. This partially explains the variation we see over time in the summary statistics of individual income at the highest brackets. Average area-level income was coded in a continuous fashion so that a CPI-adjustment was, in principle, a possibility. We were concerned about a lack of conformity in having the personal income not CPI adjusted and average area-level income CPI-adjusted, so we ran our regression specifications with and without CPI-adjusted average income in the CBSA area. The results were not distinguishably different in any way. The effects of income were also similar when our regressions were run with yearly data rather than using the full sample, further showing little reliance of our results on the CPI adjustment.

[20]Of course, there are discrepancies in the percentages for religious volunteering between this table and the preceding one because the respective ratios use different denominators. Specifically, this table shows the

children's educational or sports groups, as well as social and community service is generally the largest portion of volunteer work in this survey. Together, these three categories account for between 71% and 76% of the volunteering activities in any year of the sample. The observed high rate of participation in religious organizational volunteering is an additional reason for considering this type of volunteer work separately in the regression portion of our analysis. We note that, despite a few fluctuations, the time period under study is characterized by a relatively constant flow of volunteer work by type of activity.[21]

4.2 Regression Analysis

The primary goal of our empirical investigation was to determine the effect of average volunteering by others on one's probability of engaging in volunteer work. We additionally accounted for other factors at the individual and area-level of analysis which may affect one's volunteering decision. The general specification for individual i in CBSA area j is:

$$VOL_{i,j} = f\left(DEMOG_i, CBSACHAR_j, AVGVOL_j\right)$$

where $VOL_{i,j}$, $DEMOG_i$, $CBSACHAR_j$, and $AVGVOL_j$ depict, respectively, the individual's binary decision of whether or not to volunteer, individual demographic characteristics, community characteristics of the CBSA-area, and the average level of volunteering in that area. We also note that when $VOL_{i,j}$ is "general" volunteering, $AVGVOL_j$ measures average "general" volunteering in the area. Similarly, when $VOL_{i,j}$ is non-religious volunteering, $AVGVOL_j$measures average non-religious volunteering in the area.

The Baseline Probit-model Structure

We instantiate this general model with a probit structure (and a differenced probit for marginal effects). Specifically, we view each individual as having some inherent desire to volunteer and actually doing so once this desire exceeds an unknown threshold α. This can be represented with a latent variable structure where $VOL^*_{i,j}$ and $VOL_{i,j}$ are, respectively, the latent desire to volunteer, and the observed decision of whether or not to do so. Formally, we estimate the specification

percentage who do volunteer work of a particular type, using, as the denominator representations of different types of volunteering in the community. In contrast, Table 1 uses the full population as the denominator.

[21]Notice that changes in the identification of volunteering organizations over time (the inclusion of immigration volunteering being one of the most important) altered the way these questions are coded. This is why an "other" category is missing from our 2007 data and explains part of the variation by category.

$$
\begin{aligned}
VOL^*_{i,j} &= \beta_0 + \beta_1 DEMOG_i + \beta_2 CBSACHAR_j + \beta_3 AVGVOL_j + \epsilon_{i,j} \\
VOL_{i,j} &= \begin{cases} 1 & \text{if } VOL^*_{i,j} > \alpha \\ 0 & \text{if } VOL^*_{i,j} \leq \alpha \end{cases}
\end{aligned}
$$

for some value α. All of our regressions in the baseline and all robustness analyses additionally include year fixed effects, with robust standard errors clustered on CBSA-area, and probability weighting for sample inclusion.

Ordinary Least Squares Robustness Test

To further test the robustness of our probit results, we compare them with those from an Ordinary Least Squares (OLS) model. Namely, for the ith individual in the jth CBSA, our OLS specification is as follows:

$$VOL_{i,j} = \alpha_0 + \alpha_1 DEMOG_i + \alpha_2 CBSACHAR_j + \alpha_3 AVGVOL_j + \varepsilon_{i,j}$$

Even though the linearized probability model is clearly an inferior fit compared to the probit one given a binary dependent variable structure, we present OLS results for reasons of completeness in our exposition. A more important reason for employing the OLS structure is due to well-known problems in the use of instrumental variables (IV) analysis for binary dependent variables. Employing an OLS structure for the baseline analysis facilitates the ease of comparison with associated linearized instrumental variables regressions which do not suffer from the aforementioned problems with binary dependent variables.

Instrumental Variables Regressions

In considering the effect of the average volunteering in one's community on an individual's decision to volunteer, two main confounding forces are considered to be of chief concern. The first regards individuals sorting by personal proclivity into communities with matching associated volunteering patterns. In particular, endogeneity may arise through the following mechanism: individuals inclined to volunteer choose to settle in communities where others are like them and, therefore, choose locations with higher average rates of volunteering. As a result, observed positive relationships between higher average volunteering rates at the community-level and individual volunteering may reflect, at least partially, this selective sorting mechanism. The second issue regards reflection bias. Reflection bias will occur when employing the average of the dependent variable in the set of control variables. This issue will be partially mitigated by the present analysis's focus on a leave-out mean. However, there does remain some concern on the matter and necessitates the use of an alternative approach.

In order to account for these two types of endogeneity, namely, reflection bias and selective sorting by community average volunteering, we undertook an IV analysis choosing two different types of instruments. The two year lagged value of average volunteering was chosen as an instrument for current average volunteering in the community.[22] [23] This choice of instrumentation strategy alleviated concerns regarding reflection bias. To a lesser extent, it alleviated concerns regarding selective sorting in the case where individuals may have chosen their communities more recently.

In order to address selective sorting in particular, we used the two year lagged average of religious volunteering as an instrument for current average nonreligious volunteering. We chose this because of its posited relevance and excludability. Specifically, the reasoning behind this second instrumentation strategy was as follows: individuals engaged in nonreligious volunteering may follow the pattern of nonreligious volunteering in the community. The "religious" level of volunteering, while correlated with other types of (nonreligious) volunteering in the community (i.e. instrument relevance), should not directly influence the selective sorting of individuals into the community (i.e. excludable instrument). Additionally, the reasons that individuals engage in religious volunteering are quite different compared to general non-religious types of volunteering.

Both instrumentation strategies were employed in isolation as well as in a multi-instrument regression strategy. The combined instrumentation strategy, as opposed to the single-instrument approach, allowed us to test for instrument validity, per the exclusion restriction using Hansen-J over-identification statistics.[24] In this way, the power of the exclusion restriction was addressed through over-identification tests in the multi-instrumentation strategy.[25]

[22]We employed a two-year lagged value of the level of average volunteering because of the sampling design of the CPS. The outgoing rotation group structure had individuals resampled over time so the current two-year lag strategy alleviated the concern that individuals would show up in the instrument.

[23]The time-span of our baseline analysis focused primarily on the last two years of our data set (2006-2007) because of the use of this instrumentation strategy. This allowed us to obtain a comparable time-frame between the baseline and the instrumental variables regressions. When we employed all four years of data for the baseline analysis, results were extremely similar to those presented here and are available upon request.

[24]The Hansen J Statistic was appropriate rather than the Sargan F statistic due to the nature of our standard errors. The concept, however, is identical.

[25]We would note the similarity of results using each of the instrumentation strategies in our analysis. Each of our instruments passes the test of relevance, since first stage F-statistics are well beyond conventional levels of significance (in the single instrument case, they greatly exceeded the "rule of thumb" of 10. In the multi-instrument strategy, the Cragg Donald F also showed clear relevance for the set of instruments). While we

Regression Results

Table 3 displays the effect of average levels of volunteering in the CBSA on an individual's likelihood of engaging in volunteer work. All regressions in this and later tables employ controls for year, as well as probability weighting for sample inclusion and robust standard errors clustered at the CBSA-level. Controls for demographic and local area characteristics were added into the regressions in a progressive fashion.

Columns 1-3 display results from regressions examining the effect of average (all types) volunteering on individual volunteering of any type. Columns 4-6 display results from regressions examining the effect of area-level average non-religious volunteering on individual non-religious volunteering. Averages are taken from the current year relative to the individual's decision to volunteer. The results are shown with probit regressions in the top half of the table and OLS linearized regressions on the bottom half of the table. There are three regressions for each type of volunteering outcome employed here and, therefore, a total of twelve regressions shown in this table. Coefficients are shown with t-statistics in brackets. Marginal probit coefficients are shown in italics below the probit coefficients and above the associated t-statistics. [26]

Table 3 demonstrates a clear positive relationship between average area-level volunteering and an individual's decision to volunteer. This relationship is statistically significant at the 1% level throughout the table. This is true for both the general volunteering as well as the non-religious volunteering regressions. It is also clear that controlling for measures of individual and community characteristics, while reducing the effect of average volunteering (because presumably there will be some relationship between other characteristics of communities and individuals and the volunteering rate), does not manage to entirely erase the relationship. The coefficient on the average level of volunteering in the community remains positive and highly significant. Specifically, in the regressions in both columns 3 and 6 which employ all individual and local-area controls, the magnitude of the relationship between area-level volunteering and an individual's decision to volunteer remains statistically significant at the 1% level.[27] In order to determine whether selection and endogeneity drive this

have provided multi-instrument over-identification evidence, we would note that the instruments served very different purposes in our analysis and, for this reason, it may be argued that individual instrumentation strategies were more appropriate than the multi-instrument strategy. We note this here as a caveat.

[26] Controls for individual characteristics, area characteristics and year are included at various points in the table as explained above. Coefficients on these variables are suppressed due to space constraints, but are available upon request.

[27] *A priori*, one might have assumed that religious volunteering was more strongly affected by the actions of one's peers. As evidenced from this table, at least from a baseline analysis, there is little evidence for this hypothesis.

positive relationship, Tables 4 and 5 employ an instrumental variables regression approach.

Table 4 replicates the regression analysis in Table 3, however, in Table 4, only the probit and associated instrumental variables probit regressions are shown, rather than the additional OLS regressions shown in Table 3.[28] Columns 1-3 of Table 4 once again show the effect of average (all types) volunteering on individual volunteering of any type, while columns 4-6 examine the effect of area-level average non-religious volunteering on individual non-religious volunteering. The instrument for average current levels of volunteering is the two year lag of this same measure.

Table 4 once again demonstrates a positive relationship between average volunteering in the area and an individual's decision to volunteer. This is true for both the general type of volunteering as well as for non-religious volunteering.[29] Appendix Table 1 further confirms this positive relationship between average area-level and individual volunteering by using the same general setup as in Table 4 but employing a linearized instrumental variables regression rather than a probit structure.

In both Tables 4 and Appendix Table 1, it is clear that the coefficients on average area-level volunteering are somewhat larger and display smaller t-statistics in the instrumented, relative to the baseline, probit regressions. The decrease in t-statistics is to be expected because we are using an instrumental variables approach while the increase in the magnitude of coefficients shows evidence that we may indeed see some problems of bias in our initial baseline regression relationships. Specifically, these increased coefficients in the instrumented, relative to the baseline, regressions point to an underestimation of the true positive effect of average on individual levels of volunteering. We find additional evidence for endogeneity in our baseline relationship from the chi-squared for the IVPROBIT regressions, which tests for exogeneity and rejects the null hypothesis of exogeneity at the 1% level.

In addition to employing an instrumental variables structure designed chiefly to address reflection bias, i.e. the two-year lagged values of average volunteering and non-religious volunteering as instruments for their respective current values which was just described, we also examined an instrumentation strategy designed to address endogeneity arising from selective sorting on location. Results from this additional instrumentation strategy, shown in Table 5, verify the same pattern of evidence indicating a clear positive effect of average area-level volunteering on the individual volunteering choice.

In Table 5, due to the nature of the instrument, only the outcome of individual non-

[28]The top half of Table 4 is a replica of the bottom half of Table 3, allowing for a direct comparison of the baseline and instrumental variables probit regressions.

[29]The effect of community volunteering on individual decisions appears stronger when examining non-religious volunteering rather than when examining volunteering more generally. This result is not robust to other regression specifications, however, so we do not over-interpret the importance of this particular result.

religious volunteering is examined. To facilitate comparison with results from our previous instrumentation strategy, columns 1-3 redisplay the instrumental variables regression results available from Table 4 (IVPROBIT) and Appendix Table 1 (IVREG) in the top and bottom halves of the table, respectively. Columns 4-6 depict the results from employing the new instrumental variables strategy, where the average two-year lagged value of religious volunteering is used as an instrument for average current values of non-religious volunteering. Columns 7-9 examine effects from including both instruments in a joint specification structure. It is evident from this table that regardless of the instrumental variables chosen or the underlying structural model, average volunteering in the area causes an increase in individual volunteering rates. Specifically, throughout the table, the effect of average volunteering on individual volunteering choices is positive and statistically significant at the 1% level.

Appendix Table 2 supplies empirical evidence for the relevance and validity of each of our chosen instruments. Once again, controls for individual and area-level characteristics are added into the regression in a progressive fashion so that the third column includes all relevant control characteristics necessary in a first stage regression. The four specifications shown in this table are, moving down vertically:

- 1. Outcome is the area "general" average volunteering, associated instrument is the 2-year lag of "general" average volunteering
- 2. Outcome is the area non-religious average volunteering, associated instrument is the 2-year lag of non-religious average volunteering
- 3. Outcome is the area non-religious average volunteering, associated instrument is the 2-year lag of religious average volunteering
- 4. Outcome is the area non-religious average volunteering, associated instruments are both the 2-year lag of religious average volunteering, and the 2-year lag of non-religious average volunteering.

In terms of instrument relevance, we can rule out weak instruments in each of specifications (1)-(3) above using the single-instrument rule-of-thumb threshold that our F-statistic exceeds 10. In the fourth and final specification in this table where two instruments are present, we consult the Cragg-Donald F-Statistic to determine instrument strength. The Cragg-Donald F of 27.25 in the fully-controlled regression with the associated p-value of 0.000 allows us to comfortably reject the null hypothesis of joint instrument irrelevance (i.e. weak instruments).

More importantly perhaps, the p-value of the Hansen J-Statistic is 0.53062 in this fourth specification. Clearly then we have provided some evidence for validity with this result, since we have no evidence supporting a failure of our instruments to conform to the exclusion restriction.

5 Concluding Remarks

In this analysis, we have improved upon the existing literature's understanding of how social interactions impact individual decision making in the context of volunteering for a public good. We have done so through the use of a coordination-game framework. More specifically, we have constructed a framework in which individual pro-social behavior is a function of the expected actions of other community members, not only due to specific community characteristics or intra-community relationships, but, more importantly, because of the underlying coordination required to realize communal goals.

Our paper is, therefore, a step forward in helping us to understand how the volunteering of individuals in the community as a whole affects the choice of whether or not any specific individual will volunteer. Our analytical examination has highlighted the importance of the community-wide level of volunteering as a stochastic coordination device to help people decide whether or not they will individually choose to volunteer. Our theoretical model predicted that, given his or her private information about the benefits of volunteering, an individual will choose to volunteer with higher probability when the level of volunteering in his or her community is also higher.

We substantiated our theoretical prediction with a complementary, and more general, empirical analysis. We found clear evidence for the positive relationship asserted in our theoretical model between the individual likelihood of volunteering and community levels of volunteering. The empirical analysis employed Census 2000 Summary File data geographically matched with CPS September Supplement data for the years 2004-2007 providing the largest and most representative current dataset on volunteering. Our baseline analysis, as well as a series of robustness checks employing various empirical specifications, verified our theoretical prediction. Our results were similar regardless of whether we used a probit or linearized regression, and we argued for a causal interpretation of results based on the statistical analysis of our instrumental variables regressions.

Taken together, our analysis has provided a clearer understanding of the relationship between individual volunteering and the volunteering of others. We have documented our results theoretically and substantiated, using a unique instrumental variables strategy, a more general version empirically. Our empirical analysis has also made a contribution in providing rigorous statistical testing of the stated relationships using a unique choice of instrumental variables in a representative dataset.

References

[1] Freeman R.B. (1997): "Working for Nothing: The Supply of Volunteer Labor," *Journal of Labor Economics*, **15**1:140-66.

[2] Govekar P.L. and M.A. Govekar (2002): "Using Economic Theory and Research to Better Understand Volunteer Behavior," *Nonprofit Management and Leadership*, **13**1:33-48.

[3] Alesina A., Baqir R., and C. Hoxby (2004): "Political Jurisdictions in Heterogeneous Communities," *Journal of Political Economy*, **112**(2):

[4] Alesina A. and E. La Ferrara (2005): "Ethnic Diversity and Economic Performance," *Journal of Economic Literature,* **43**(3):762-800.

[5] Alesina A. and E. La Ferrara (2000): "Participation in Heterogeneous Communities," *Quarterly Journal of Economics,* **115**(3):857-904.

[6] Andreoni J. (2006): "Philanthropy," in *The Handbook of Giving, Reciprocity and Altruism*, Gerard-Varet L.A., Serge C., Kolm, and J.M. Ythier (ed), Elsevier/North-Holland

[7] Andreoni J. and J.K. Scholz (1998): "An Econometric Analysis of Charitable Giving With Interdependent Preferences," *Economic Inquiry*, **36**(3):410-29.

[8] Andreoni J. (1993): "An Experimental Test of the Public Goods Crowding-Out Hypothesis," *American Economic Review*, **83**(5):1317-27.

[9] Andreoni J. (1990): "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, **100**:464-77.

[10] Andreoni J. (1988): "Privately Provided Public Goods in a Large Economy: The Limits of Altruism," *Journal of Public Economics*, **35**:57-73.

[11] Bergstrom T., Blume L.E., and H. Varian (1986): "On the Private Provision of Public Goods," *Journal of Public Economics,* **29**:25-49.

[12] Benabou R. and J. Tirole (2006): "Incentives and Prosocial Behavior," *American Economic Review*, **96**:1652-78.

[13] Bilodeau M. and A. Slivinski (1996): "Toilet cleaning and department chairing: Volunteering a public service," *Journal of Public Economics*, **59**(2):299-308.

[14] Bossert W., D'Ambrosio C., and E. La Ferrara (2010): "A Generalized Index of Fractionalization," *Economica,* **78**(312): 723-50.

[15] Blume L.E., Brock W.A., Durlauf S.N., and Y.M. Ioannides (2010): "Identification of Social Interactions," Available at SSRN: http://dx.doi.org/10.2139/ssrn.1660002

[16] Brudney J.L. and Gazley B.(2006): "Moving Ahead or Falling Behind? Volunteer Promotion and Data Collection" *NonProfit Management and Leadership*, **16**(3):259-76.

[17] Card D., Hallock K.F., and E. Moretti (2010): "The Geography of Giving: The Effect of Corporate Headquarters on Local Charities," *Journal of Public Economics,* **94**: 220-34.

[18] Carman K.G. (2004): "Social Influences and the Private Provision of Public Goods: Evidence from Charitable Contributions in the Workplace," Working Paper, Harvard University

[19] Cattan M., White M., Bond J., and A. Learmouth (2005): "Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions," *Ageing & Society*, **25**: 41-67.

[20] Chotikapanich D. and W. Griffiths (2001): "On Calculation of the Extended Gini Coefficient," *Review of Income and Wealth*, **47**(4):541-47.

[21] Dee T.S. (2004): "Are There Civic Returns to Education?" *Journal of Public Economics,* **88**(9-10):1697-720.

[22] Dorfman R. (1979): "A Formula for the Gini Coefficient," *The Review of Economics and Statistics,* **61**(1):146-49.

[23] Dosman D., Fast J., Chapman S.A., and N. Keating (2006): "Retirement and Productive Activity in Later Life," *Journal of Family and Economic Issues*, **27**: 401-19.

[24] Fedderke J., Luiz J., and R. De Kadt (2008): "Using Fractionalization Indexes: Deriving Methodological Principles for Growth Studies from Time Series Evidence," *Social Indicators Research* **85**:257-78.

[25] Feldman E.E. (2010): "Time is Money: Choosing Between Charitable Activities," *American Economic Journal,* **2**(1):103-30.

[26] Francois P. (2005): "Making A Difference," *CEPR Discussion Paper No. 5158*

[27] Garcia-Mainar I. and C. Marcuell (2007): "Members, Volunteers, and Donors in Nonprofit Organizations in Spain," *Voluntary and Nonprofit Sector Quarterly,* **36**(1):100-20.

[28] Glazer A. and K.A. Konrad (1996): "A Signaling Explanation for Charity," *The American Economic Review,* **86**(4):1019-28.

[29] Grafton R.Q., Knowles S., and P.D. Owen (2001): "Social Divergence and Economic Performance," Department of Economics, University of Ottawa, Working Paper #0103E

[30] Greisling D. (1998): "Anyone Want to Lend a Hand?" *Business week*, Issue 3587, pp. 110

[31] Harbaugh W.T. (1998): "The Prestige Motive for Making Charitable Transfers" *American Economic Review: Papers and Proceedings*, **88**(2):277-82.

[32] Isham J., Kolodinsky J., and G. Kimberly (2006): "The Effects of Volunteering for Nonprofit Organizations on Social Capital Formation: Evidence from a Statewide Survey," *Nonprofit and Voluntary Sector Quarterly,* **35**(3):367-83.

[33] Knox T.M. (1999): "The Volunteer's folly and socio-economic man: Some Thoughts on Altruism, rationality, and community," *Journal of Socio-Economics,* **28**:475-92.

[34] Kulik L. (2007): "Explaining Responses to Volunteering: An Ecological Model," *Nonprofit and Voluntary Sector Quarterly,* **36**(2):239-55.

[35] McGoldrick K., A. Battle, and G. Suzanne (2000): "Service-Learning and the Economics Course: Theory and Practice," *The American Economist*, **44**(1):43-52.

[36] Michalopoulos S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review,* **102**(4):1508-39.

[37] Miguel E. (2004) "Nation Building and Public Goods in Kenya Versus Tanzania," *World Politics,* **56**:327-62.

[38] Morris S. and H. Shin (1998) "Unique Equilibrium in a Model of Self-fulfilling Currency Attacks," *American Economic Review,* **88**:587-97.

[39] Musick M., Wilson J., and W. Bynum (2000): "Race and Formal Volunteering: The Differential Effects of Class and Religion," *Social Forces,* **78**(4):1539-70.

[40] Mutchler J.E., Burr J.A., and F.G. Caro (2003): "From Paid Worker to Volunteer: Leaving the Paid Workforce and Volunteering in Later Life," *Social Forces*, **81**(4):1267-93.

[41] Nirei M., Stamatiou T., and V. Sushko (2012): "Stochastic Herding in Financial Markets: Evidence from Investor Equity Portfolios," Bank of International Settlements, Working Paper #371.

[42] Potters J., Sefton M., and L. Vesterlund (2005): "After-You Endogenous Sequencing in Voluntary Contribution Games," *Journal of Public Economics,* **89**(8): 1399-1419.

[43] Ribar and Wilhelm (2002): "Altruistic and Joy-of-Giving Motivations in Charitable Behavior," *Journal of Political Economy*, **110**(2):425-57.

[44] Segal L.M. and B.A. Weisbrod (2002): "Volunteer Labor Sorting Across Industries" *Journal of Policy Analysis and Management*, **21**(3):427-47.

[45] Simmons W.O. and R. Emanuele (2007): "Male-female Giving differentials: Are Women More Altruistic?" *Journal of Economic Studies,* **34**(6):534-50.

[46] Sugden R. (1984): "Reciprocity: The Supply of Public Goods Through Voluntary Contributions," *The Economic Journal*, **94**: 772-87.

[47] Tarski A. (1955): "A Lattice-theoretic Fixed Point Theorem and its Applications," *Pacific Journal of Mathematics*, **5**: 285-308.

[48] Tao H.L. and P. Yeh (2007): "Religion as an Investment: Comparing the Contributions and Volunteer Frequency among Christicans, Buddhists, and Folk Religionists," *Southern Economic Journal,* **73**(3):770-90.

[49] Thoits P.A. and L.N. Hewitt (2001): "Volunteer Work and Well-Being," *Journal of Health and Social Behavior,* **42**(2):115-31.

[50] Van Dihk J. and R. Boin "Volunteer Labor Supply in the Netherlands," *De Economist,* **141**(3):402-18.

[51] Vesterlund L. (2003): "The Informational Value of Sequential Fundraising," *Journal of Public Economics*, **87**:627-57.

[52] Vigdor J.L. (2002): "Interpreting Ethnic Fragmentation Effects," *Economic Letters,* **75**:271-76.

A Proofs

Proof of Lemma 1

For the given utility difference, we have

$$\begin{aligned}
\Delta U(w,k) &= U(w-x,h((k+1)x)) - U(w-x,h(kx)) - [U(w,h(kx)) - U(w-x,h(kx))] \\
&= \int_0^x U_2(w-x,h(kx+z))\,h'(kx+z)\,\mathrm{d}z - \int_{-x}^0 U_1(w+\widetilde{z},h(kx))\,\mathrm{d}\widetilde{z} \\
&= \int_0^x [U_2(w-x,h(kx+z))\,h'(kx+z) - U_1(w-x+z,h(kx))]\,\mathrm{d}z \\
&\geq \int_0^x [U_2(w-x,h(kx))\,h'(kx) - U_1(w-x,h(kx))]\,\mathrm{d}z \\
&= [U_2(w-x,h(kx))\,h'(kx) - U_1(w-x,h(kx))]\,x
\end{aligned}$$

Here, the second and last equalities above apply the fundamental theorem of calculus, while the third follows from a change of variables in the second integral. By contrast, the inequality is due to Assumption A.1(iii). Given this result, the claim follows immediately from Assumption A.2(ii). □

Proof of Proposition 2

Since

$$\exp\left(-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)\right) = \exp\left(-\frac{(\theta_0-\mu(\theta_n))^2}{2\sigma_n^2}+\xi(\theta_n)\right)$$

where

$$\sigma_n^2 = \frac{\sigma_\epsilon^2\sigma^2}{\sigma_\epsilon^2+\sigma^2}, \qquad \mu(\theta_n) = \frac{\theta_n\sigma^2+\vartheta\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma^2}, \quad \text{and } \xi(\theta_n) = \frac{\mu(\theta_n)^2}{2\sigma_n^2} - \frac{\theta_n^2}{2\sigma_\epsilon^2} - \frac{\vartheta^2}{2\sigma^2}$$

we ought to have

$$\begin{aligned}
\delta(\theta_n,k) &= \frac{\int_{-\infty}^{\ln kx} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)}\mathrm{d}\theta_0}{\int_{\ln kx}^{+\infty} e^{-\left(\frac{(\theta_n-\theta_0)^2}{2\sigma_\epsilon^2}+\frac{(\theta_0-\vartheta)^2}{2\sigma^2}\right)}\mathrm{d}\theta_0} = \frac{\int_{-\infty}^{\ln kx} e^{-\frac{(\theta_0-\mu(\theta_n))^2}{2\sigma_n^2}}\mathrm{d}\theta_0}{\int_{\ln kx}^{+\infty} e^{-\frac{(\theta_0-\mu(\theta_n))^2}{2\sigma_n^2}}\mathrm{d}\theta_0} \\
&= \frac{\int_{-\infty}^{\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}} e^{-\frac{\widetilde{\theta}^2}{2}}\mathrm{d}\widetilde{\theta}}{\int_{\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}}^{+\infty} e^{-\frac{\widetilde{\theta}^2}{2}}\mathrm{d}\widetilde{\theta}} = \frac{\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{1-\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}
\end{aligned}$$

The claim now follows from the fact that $\mu'(\theta_n) = \frac{\sigma^2}{\sigma_\epsilon^2+\sigma^2} > 0$. $\square$

Proof of Proposition 1

The *if* direction is trivial. By hypothesis, given an arbitrary $\varepsilon > 0$, there exists $k_2 \in \mathbb{N}$ s.t. for all $k \geq k_2$

$$-\varepsilon \leq V(\theta_n,k) < V(\theta_n,k) + f(\theta_n,k)$$

where the second inequality is true since $f(\theta_n,k) > 0$ for all $(\theta_n,k) \in \mathbb{R}\times\mathbb{N}$. The result immediately follows and we may in fact take $k_1 = k_2$.
For the *only if* direction, we will make use of an intermediary result.

Lemma 2 *Consider the function* $f:\mathbb{R}\times\mathbb{N}\mapsto\mathbb{R}_{++}$ *defined by* $f(\theta_n,k) = \frac{\Pr\left[kx\leq e^{\theta}\leq(k+1)x|\theta_n,k\right]}{\Pr\left[e^{\theta}\leq kx|\theta_n,k\right]}$. *For any given* $\theta_n\in\mathbb{R}$, *the following hold*

- *(i)* $\lim_{k\to+\infty} f(\theta_n,k) = 0$
- *(ii)* $f_1(\theta_n,k) > 0$ *for any* $k\in\mathbb{N}$ *s.t.* $k \geq e^{\mu(\theta_n)}/x$

Proof. Part (i) is trivial.[30] For (ii), fix the signal θ_n and recall the various likelihoods we have defined in the main text. Since

$$\begin{aligned}
f(\theta_n,k) &= \frac{\Pr\left[kx\leq e^{\theta_0}\leq(k+1)x|\theta_n,k\right]}{\Pr\left[e^{\theta_0}\leq kx|\theta_n,k\right]} \\
&= \frac{\Pr\left[k|\theta_n,kx\leq e^{\theta_0}\leq(k+1)x\right]\Pr\left[\theta_n,kx\leq e^{\theta_0}\leq(k+1)x\right]}{\Pr\left[k|\theta_n,e^{\theta_0}\leq kx\right]\Pr\left[\theta_n,e^{\theta_0}\leq kx\right]}
\end{aligned}$$

[30] Obviously, the events $\{\theta_0\in\mathbb{R}:\ln kx\leq\theta_0\leq\ln(k+1)x\}$ and $\{\theta_0\in\mathbb{R}:\theta_0\leq\ln(k+1)x\}$ become, respectively, arbitrarily small and large as $k\to+\infty$. And as the random variable θ_0 is continuous, so are the nominator and denominator of $f(\theta_n,k)$. That is, the nominator [resp. denominator] of $f(\theta_n,k)$ gets arbitrarily small [resp. close to 1] as $k\to+\infty$.

under the hypothesized typical response rule, we may also define

$$\Pr\left[k|\theta_n, kx \le e^{\theta_0} \le (k+1)\,x\right] = H\left(\theta\left(k\right), k\right)^{N-k-1} \widetilde{H}\left(\theta\left(k\right), k\right)^{k}$$

where

$$\begin{aligned} H\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n > \theta\left(k\right)|kx \le e^{\theta_0} \le (k+1)\,x\right] \\ \widetilde{H}\left(\theta\left(k\right), k\right) &= \Pr\left[\theta_n \le \theta\left(k\right)|kx \le e^{\theta_0} \le (k+1)\,x\right] \end{aligned}$$

and write

$$\begin{aligned} \frac{f\left(\theta_n, k\right)}{\widetilde{A}\left(\theta\left(k\right), k\right)^{N-k-1} \widetilde{B}\left(\theta\left(k\right), k\right)^{k}} &= \frac{\Pr\left[\theta_n, kx \le e^{\theta_0} \le (k+1)\,x\right]}{\Pr\left[\theta_n, e^{\theta_0} \le kx\right]} \\ &= \frac{\Phi\left(\frac{\ln(k+1)x-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)} - 1 \end{aligned}$$

where

$$\widetilde{A}\left(\theta\left(k\right), k\right) = \frac{H\left(\theta\left(k\right), k\right)}{F\left(\theta\left(k\right), k\right)} \qquad \widetilde{B}\left(\theta\left(k\right), k\right) = \frac{\widetilde{H}\left(\theta\left(k\right), k\right)}{\widetilde{F}\left(\theta\left(k\right), k\right)}$$

Hence,

$$
\begin{aligned}
& \frac{f_1(\theta_n,k)}{\widetilde{A}(\theta(k),k)^{N-k-1}\,\widetilde{B}(\theta(k),k)^k} \\
= \ & \frac{\mu'(\theta_n)}{\sqrt{\sigma_n^2}\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}\left[\phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\frac{\Phi\left(\frac{\ln(k+1)x-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}-\phi\left(\frac{\ln(k+1)x-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\right] \\
> \ & \frac{\mu'(\theta_n)\phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\sqrt{\sigma_n^2}\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}\left[1-\frac{\phi\left(\frac{\ln(k+1)x-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}\right] \\
= \ & \frac{\mu'(\theta_n)\phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\sqrt{\sigma_n^2}\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}\left[1-e^{-\frac{(\ln(k+1)x-\mu(\theta_n))^2-(\ln kx-\mu(\theta_n))^2}{2\pi\sigma_n^2}}\right] \\
= \ & \frac{\mu'(\theta_n)\phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\sqrt{\sigma_n^2}\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}\left[1-e^{-\frac{\left(\ln k(k+1)x^2-2\mu(\theta_n)\right)\ln\left(\frac{k+1}{k}\right)}{2\pi\sigma_n^2}}\right]
\end{aligned}
$$

and the hypothesis ensures that the last bracketed quantity above is positive. ■

We are now in a position to establish the *only if* part of the proposition. By hypothesis, for an arbitrary $\varepsilon > 0$, there will be some $k_1 \in \mathbb{N}$, such that

$$V(\theta_n,k) \geq -\left[\frac{\varepsilon}{2}+f(\theta_n,k)\right] \geq -\left[\frac{\varepsilon}{2}+f(\theta_0+c,k)\right]$$

for all $k \geq \widetilde{k} = \max\left\{k_1, e^{\mu(\theta_0+c)}/x\right\}$. Here, the second inequality is due to part (ii) of the preceding lemma. The result follows since part (i) of the lemma ensures that for some $k(\theta_0+c) \in \mathbb{N}$ and for all $k \geq k_2 = \max\left\{\widetilde{k}, k(\theta_0+c)\right\}$ the quantity $f(\theta_0+c,k)$ will be smaller than $\varepsilon/2$. □

Proof of Proposition 3

Step 1 With respect to the left-hand side of (3), $U(w-x,h((k+1)x))$ is increasing in k given assumptions A.1(ii) and A.2(i). The same type of monotonicity, moreover, also characterizes the utility difference $\Delta U(w,x,k)$ since, for any $k, k' \in \mathbb{N}$ with $k' > k$, we get

$$\begin{aligned}\Delta U(w,x,k) &= \int_0^x \left(U_2(w-x, h(kx+z)) f'(kx+z) - U_1(w-x+z, h(kx))\right) \mathrm{d}z \\ &\leq \int_0^x \left(U_2(w-x, h(kx+z)) f'(k'x+z) - U_1(w-x+z, h(k'x))\right) \mathrm{d}z \\ &= \Delta U(w,x,k')\end{aligned}$$

the inequality being due to assumptions A.1(iii) and A.2(iii). For sufficiently large k, therefore, the quantity of interest is non-increasing with respect to the number of contributions.

Step 2 The right-hand side of (3), on the other hand, has the following total differential with respect to k

$$\begin{aligned}&\delta_2(\theta(k),k) A(\theta(k),k)^{N-k-1} B(\theta(\lambda),k)^k \\ +\ &\delta(\theta(k),k) \frac{\partial}{\partial k}\left(A(\theta(k),k)^{N-k-1} B(\theta(\lambda),k)^k\right) \\ +\ &\theta'(k)\begin{pmatrix} \delta_1(\theta(k),k) A(\theta(k),k)^{N-k-1} B(\theta(\lambda),k)^k \\ +\delta(\theta(k),k)(N-k) A(\theta(k),k)^{N-k-1} A_1(\theta(k),k) B(\theta(\lambda),k)^k \\ +\delta(\theta(k),k) A(\theta(k),k)^{N-k-1} k B_1(\theta(\lambda),k)^{k-2}\end{pmatrix}\end{aligned} \tag{4}$$

Step 2.1(i) We will begin by showing that the first two terms in (4) sum to a positive number. The first term is always positive by itself since

$$\delta_2(\theta,k) = \frac{\phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\delta(\theta,k)}{\Phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\left(1 - \Phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\right)} > 0 \tag{5}$$

Step 2.1(ii) With respect to the second term, we have

$$\frac{A(\theta,k+1)^{N-k-2} B(\theta,k+1)^{k+1}}{A(\theta,k)^{N-k-1} B(\theta,k)^k} = \left(\frac{A(\theta,k+1)}{A(\theta,k)}\right)^{N-1} \frac{\left(\frac{B(\theta,k+1)}{A(\theta,k+1)}\right)^{k+1}}{\left(\frac{B(\theta,k)}{A(\theta,k)}\right)^k} \tag{6}$$

and, given the following result,

Lemma 3 *Let γ, x_0, y_0, and s be reals with the latter one non-zero. Let also $x \sim \mathcal{N}(x_0, s^2)$ and $y \sim \mathcal{N}(y_0, s^2)$. Then, $x_0 < y_0$ only if $\Pr[x \geq \gamma] < \Pr[y \geq \gamma]$.*

Proof. Since

$$\Pr\left[x \geq \gamma\right] = \int_{\gamma}^{+\infty} e^{-\frac{(x-x_0)^2}{2s^2}} \mathrm{d}x = \frac{1}{\sqrt{2s^2}} \int_{\frac{\gamma - x_0}{\sqrt{s^2}}}^{+\infty} e^{-\frac{\widetilde{x}^2}{2}} \mathrm{d}\widetilde{x} = \frac{1 - \Phi\left(\frac{\gamma - x_0}{\sqrt{s^2}}\right)}{\sqrt{2s^2}}$$

the claim follows immediately from the fact that $\Phi\left(\frac{\gamma - x}{\sqrt{s^2}}\right)$ is strictly decreasing in x. ■

it must be

$$A\left(\theta, k\right) = \frac{\Pr\left[\theta_n > \theta\left(k\right) | e^{\theta_0} \leq kx\right]}{\Pr\left[\theta_n > \theta\left(k\right) | e^{\theta_0} > kx\right]} < 1 < \frac{\Pr\left[\theta_n \leq \theta\left(k\right) | e^{\theta_0} \leq kx\right]}{\Pr\left[\theta_n \leq \theta\left(k\right) | e^{\theta_0} > kx\right]} = B\left(\theta, k\right) \tag{7}$$

Therefore,

$$\frac{A\left(\theta, k+1\right)^{N-k-2} B\left(\theta, k+1\right)^{k+1}}{A\left(\theta, k\right)^{N-k-1} B\left(\theta, k\right)^{k}} > \left(\frac{B\left(\theta, k+1\right)}{A\left(\theta, k+1\right)}\right)^{k+1} / \left(\frac{B\left(\theta, k\right)}{A\left(\theta, k\right)}\right)^{k} \tag{8}$$

and to show that the second term in (4) is positive for sufficiently large k it is enough to establish that the right-hand side of this inequality tends to 1 as $k \to +\infty$.
To this end, observe that

$$\begin{aligned} B\left(\theta, k\right) &= \frac{\Pr\left[\theta_n < \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[\theta_n < \theta, e^{\theta_0} > kx\right]} \frac{\Pr\left[e^{\theta_0} > kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} \\ &= \left(\frac{\int_{-c}^{\theta} \int_{-\infty}^{\ln kx} e^{-\frac{(\theta_n - \theta_0)^2}{2\sigma_\epsilon^2}} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0 \mathrm{d}\theta_n}{\int_{-c}^{\theta} \int_{\ln kx}^{+\infty} e^{-\frac{(\theta_n - \theta_0)^2}{2\sigma_\epsilon^2}} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta \mathrm{d}\theta_n}\right) \left(\frac{\int_{\ln kx}^{+\infty} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0}{\int_{-\infty}^{\ln kx} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0}\right) \end{aligned}$$

with

$$\begin{aligned} \frac{\partial}{\partial k} \int_{-c}^{\theta} \int_{\ln kx}^{+\infty} e^{-\frac{(\theta_n - \theta_0)^2}{2\sigma_\epsilon^2}} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0 \mathrm{d}\theta_n &= -\int_{-c}^{\theta} e^{-\frac{(\theta_n - kx)^2}{2\sigma_\epsilon^2}} e^{-\frac{(kx - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_n \\ &= -e^{-\frac{(\ln kx - \vartheta)^2}{2\sigma_0^2}} \int_{-c}^{\theta} e^{-\frac{(\theta_n - \ln kx)^2}{2\sigma_\epsilon^2}} \mathrm{d}\theta_n \\ &= -\frac{\partial}{\partial k} \int_{-c}^{\theta} \int_{-\infty}^{\ln kx} e^{-\frac{(\theta_n - \theta_0)^2}{2\sigma_\epsilon^2}} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0 \mathrm{d}\theta_n \end{aligned}$$

and

$$\frac{\partial}{\partial k} \int_{\ln kx}^{+\infty} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0 = -e^{-\frac{(\ln kx - \vartheta)^2}{2\sigma_0^2}} = -\frac{\partial}{\partial k} \int_{-\infty}^{\ln kx} e^{-\frac{(\theta_0 - \vartheta)^2}{2\sigma_0^2}} \mathrm{d}\theta_0$$

By L'Hôpital's rule, therefore, either ratio in the product defining $B(\theta, k)$ above converges to -1 as $k \to +\infty$. In other words, $\lim_{k\to+\infty} B(\theta, k) = 1$ and a trivially similar argument shows that $\lim_{k\to+\infty} A(\theta, k) = 1$ as well.
Notice also that

$$A_2(\theta, k) = \frac{G_2(\theta, k)}{G(\theta, k)} \left(\frac{F_2(\theta, k)}{G_2(\theta, k)} - A(\theta, k) \right)$$

with

$$\begin{aligned} F_2(\theta, k) = \frac{\partial}{\partial k} \left(\frac{\Pr\left[\theta_n \geq \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} \right) \\ = \quad & \frac{\Pr\left[\theta_n \geq \theta, e^{\theta_0} = kx\right] \Pr\left[e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]^2} \\ & - \frac{\Pr\left[e^{\theta_0} = kx\right] \Pr\left[\theta_n \geq \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]^2} \\ = \quad & \frac{\Pr\left[e^{\theta_0} = kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} \left(\frac{\Pr\left[\theta_n \geq \theta, e^{\theta_0} = kx\right]}{\Pr\left[e^{\theta_0} = kx\right]} - \frac{\Pr\left[\theta_n \geq \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} \right) \\ = \quad & \frac{\Pr\left[e^{\theta_0} = kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} \left(\Pr\left[\theta_n \geq \theta | e^{\theta_0} = kx\right] - \Pr\left[\theta_n \geq \theta | e^{\theta_0} \leq kx\right] \right) \end{aligned}$$

and where, abusing notation slightly, we meant to denote

$$\begin{aligned} \Pr\left[e^{\theta_0} = kx\right] &= \frac{1}{\sqrt{\sigma^2}} \phi \left(\frac{\ln kx - \vartheta}{\sqrt{\sigma^2}} \right) \\ \Pr\left[\theta_n \geq \theta, e^{\theta_0} = kx\right] &= \int_{\theta}^{\ln kx + c} e^{-\frac{(\theta_n - \ln kx)^2}{2\sigma_\epsilon^2}} \mathrm{d}\theta_n \end{aligned}$$

Similarly,

$$\begin{aligned} B_2(\theta, k) &= -\frac{\widetilde{G}_2(\theta, k)}{\widetilde{G}(\theta, k)} \left(\frac{\widetilde{F}_2(\theta, k)}{\widetilde{G}_2(\theta, k)} - B(\theta, k) \right) \\ &= \frac{\Pr\left[e^{\theta_0} = kx\right] + G_2(\theta, k)}{1 - \Pr\left[e^{\theta_0} \leq kx\right] - G(\theta, k)} \left(\frac{\Pr\left[e^{\theta_0} = kx\right] - F_2(\theta, k)}{-\Pr\left[e^{\theta_0} = kx\right] - G_2(\theta, k)} - B(\theta, k) \right) \end{aligned}$$

It is now straightforward to check the following: On the one hand, $\lim_{k\to+\infty} \frac{A_2(\theta,k)}{A(\theta,k)} - \frac{B_2(\theta,k)}{B(\theta,k)} = 0$. On the other hand, due to the property of the normal distribution we observed previously, $F_2(\theta, k) > 0$ while $G_2(\theta, k) < 0$. That is, $A_2(\theta, k) > 0$ and, by a similar analysis, $\widetilde{F}_2(\theta, k) < 0 < \widetilde{G}_2(\theta, k)$ and $B_2(\theta, k) < 0$.

To complete the argument, consider the sequence $\left\{\alpha_k = \left(\frac{A(\theta,k)}{B(\theta,k)}\right)^k\right\}_{k\in\mathbb{N}^*}$. Under (7), it must be $\alpha_k \in (0,1)$ for all k. By the mean value theorem, moreover,

$$\begin{aligned}
\frac{\alpha_{k+1}}{\alpha_k} - 1 &= \ln\left(\frac{A(\theta,\kappa)}{B(\theta,\kappa)}\right) + \kappa\left(\frac{A_2(\theta,\kappa)}{A(\theta,\kappa)} - \frac{B_2(\theta,\kappa)}{B(\theta,\kappa)}\right) \\
&> \ln\left(\frac{A(\theta,\kappa)}{B(\theta,\kappa)}\right) + \frac{A_2(\theta,\kappa)}{A(\theta,\kappa)} - \frac{B_2(\theta,\kappa)}{B(\theta,\kappa)}
\end{aligned}$$

for some $\kappa \in (k, k+1)$. And since the last quantity above vanishes as $k \to +\infty$, it must be $\lim_{k\to+\infty} \frac{\alpha_{k+1}}{\alpha_k} = 1^+$.[31] This means that the right-hand side of (8) converges to 1^-. Which in turn suffices to ensure that the sum of the first two terms in (4) is positive since successive applications of L'Hopital's rule give

$$\begin{aligned}
\lim_{k\to+\infty} \delta_2(\theta,k) &= \lim_{k\to+\infty} \frac{\phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\left(1 - \Phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\right)^2} = \lim_{k\to+\infty} \frac{k^{-1}(\ln kx - \mu(\theta_n))}{2\sigma_n^2\left(1 - \Phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\right)} \\
&= \lim_{k\to+\infty} \frac{k^{-2}(\ln kx - \mu(\theta_n) - 1)}{2\sigma_n^2\phi\left(\frac{\ln kx - \mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)} = +\infty
\end{aligned}$$

Step 2.2 Next, we determine the sign of the bracketed quantity in the last term of (4). In Proposition 2, we have already established that $\delta_1(\theta,k) < 0$. Moreover,

$$A_1(\theta,k) = \frac{G_1(\theta,k)}{G(\theta,k)}\left(\frac{F_1(\theta,k)}{G_1(\theta,k)} - A(\theta,k)\right)$$

where

$$\begin{aligned}
F_1(\theta,k) &= \frac{\partial}{\partial\theta}\left(\frac{\Pr\left[\theta_n \geq \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]}\right) \\
&= -\frac{\Pr\left[\theta_n = \theta, e^{\theta_0} \leq kx\right]}{\Pr\left[e^{\theta_0} \leq kx\right]} = -\frac{\Phi\left(\frac{\ln kx - \mu(\theta)}{\sqrt{\sigma_n^2}}\right)}{\Phi\left(\frac{\ln kx - \vartheta}{\sqrt{\sigma^2}}\right)}
\end{aligned}$$

[31] More precisely, we just established that, taking a subsequence if necessary, $\{\alpha_k\}$ becomes non-decreasing. As it is also bounded, it must converge to some $l \in (0,1]$.

and, similarly, $G_1(\theta,k) = -\frac{1-\Phi\left(\frac{\ln kx-\mu(\theta)}{\sqrt{\sigma_n^2}}\right)}{1-\Phi\left(\frac{\ln kx-\vartheta}{\sqrt{\sigma^2}}\right)}$. Hence,

$$\frac{F_1(\theta,k)}{G_1(\theta,k)} = \frac{\Phi\left(\frac{\ln kx-\mu(\theta)}{\sqrt{\sigma_n^2}}\right)}{1-\Phi\left(\frac{\ln kx-\mu(\theta)}{\sqrt{\sigma_n^2}}\right)} \frac{1-\Phi\left(\frac{\ln kx-\vartheta}{\sqrt{\sigma^2}}\right)}{\Phi\left(\frac{\ln kx-\vartheta}{\sqrt{\sigma^2}}\right)}$$

and since

$$A(\theta,k) = \frac{\int_\theta^c \Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right) e^{\xi(\theta_n)}\mathrm{d}\theta_n}{\int_\theta^c \left[1-\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)\right] e^{\xi(\theta_n)}\mathrm{d}\theta_n} \frac{1-\Phi\left(\frac{\ln kx-\vartheta}{\sqrt{\sigma^2}}\right)}{\Phi\left(\frac{\ln kx-\vartheta}{\sqrt{\sigma^2}}\right)}$$

it ought to be

$$\begin{aligned}
\frac{A(\theta,k)}{F_1(\theta,k)/G_1(\theta,k)} &= \int_\theta^c \frac{\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{\Phi\left(\frac{\ln kx-\mu(\theta)}{\sqrt{\sigma_n^2}}\right)} e^{\xi(\theta_n)}\mathrm{d}\theta_n \Big/ \int_\theta^c \frac{1-\Phi\left(\frac{\ln kx-\mu(\theta_n)}{\sqrt{\sigma_n^2}}\right)}{1-\Phi\left(\frac{\ln kx-\mu(\theta)}{\sqrt{\sigma_n^2}}\right)} e^{\xi(\theta_n)}\mathrm{d}\theta_n \\
&< 1
\end{aligned}$$

where the inequality is due to the fact that $\Phi\left(\frac{\ln kx-\mu(\cdot)}{\sqrt{\sigma_n^2}}\right)$ is decreasing. We therefore conclude that $A_1(\theta,k) < 0$. Similarly,

$$B_1(\theta,k) = -\frac{G_1(\theta,k)}{1-\Pr\left[e^{\theta_0} \le kx\right] - G(\theta,k)} \left(\frac{F_1(\theta,k)}{G_1(\theta,k)} - B(\theta,k)\right)$$

but

$$\begin{aligned}
B(\theta,k) &= \frac{\Pr\left[\theta_n \ge \theta, \theta < kx\right]}{\Pr\left[\theta_n \ge \theta, \theta \ge kx\right]} \frac{\Pr\left[\theta \ge kx\right]}{\Pr\left[\theta < kx\right]} \\
&= \frac{\int_{-c}^\theta \Phi\left(\frac{kx-\mu(\theta_n)}{\sigma^2}\right) e^{\xi(\theta_n)}\mathrm{d}\theta_n}{\int_{-c}^\theta \left[1-\Phi\left(\frac{kx-\mu(\theta_n)}{\sigma^2}\right)\right] e^{\xi(\theta_n)}\mathrm{d}\theta_n} \frac{\Pr\left[\theta \ge kx\right]}{\Pr\left[\theta < kx\right]}
\end{aligned}$$

and, thus,

$$\begin{aligned}\frac{B(\theta,k)}{F_1(\theta,k)/G_1(\theta,k)} &= \int_{-c}^{\theta}\frac{\Phi\left(\frac{kx-\mu(\theta_n)}{\sigma^2}\right)}{\Phi\left(\frac{kx-\mu(\theta)}{\sigma^2}\right)}e^{\xi(\theta_n)}\mathrm{d}\theta_n\Big/\int_{-c}^{\theta}\frac{1-\Phi\left(\frac{kx-\mu(\theta_n)}{\sigma^2}\right)}{1-\Phi\left(\frac{kx-\mu(\theta)}{\sigma^2}\right)}e^{\xi(\theta_n)}\mathrm{d}\theta_n \\ &> 1\end{aligned}$$

Since, therefore, $B_1(\theta,k)<0$ as well, the bracketed quantity in question is clearly negative. Step 3 Recall once again the defining equation for the cut-off $\theta(k)$. Since for sufficiently large k its left-hand side is non-increasing with k, the implicit function theorem dictates that the total differential of its right-hand side with respect to k ought to be non-positive. This differential is given by (4), however, and we just showed that, as $k\to+\infty$, its first two terms are positive whereas the bracketed part of its last term is negative. Hence, for large enough k, it cannot help but be the case that $\theta'(k)>0$ as required. □

Table 1: Summary Statistics

	Full Sample (04-07)						
	Mean	Min	Max	2004	2005	2006	2007
Probability Individual Volunteers	0.28	0	1	0.30	0.30	0.28	0.27
Non Church Volunteering Probability	0.13	0	1	0.14	0.14	0.13	0.13
Age	36.99	1	85	36.72	36.83	37.05	37.36
Male	0.48	0	1	0.48	0.48	0.48	0.48
White	0.81	0	1	0.81	0.81	0.81	0.80
Black	0.13	0	1	0.13	0.13	0.13	0.13
American Indian	0.01	0	1	0.01	0.01	0.01	0.01
Hispanic	0.05	0	1	0.05	0.05	0.06	0.06
Less than First Grade Education	0.00	0	1	0.00	0.00	0.00	0.00
1st-4th Grade Education	0.01	0	1	0.01	0.01	0.01	0.01
5th-6th Grade Education	0.02	0	1	0.02	0.02	0.02	0.02
7th-8th Grade Education	0.02	0	1	0.03	0.02	0.02	0.02
9th Grade Education	0.03	0	1	0.03	0.03	0.03	0.03
10th Grade Education	0.04	0	1	0.04	0.04	0.04	0.04
11th Grade Education	0.04	0	1	0.04	0.04	0.04	0.04
12th Grade Education (no Diploma)	0.01	0	1	0.01	0.01	0.02	0.01
HS Grad (or GED)	0.29	0	1	0.29	0.29	0.28	0.29
Some College (No Degree)	0.18	0	1	0.18	0.18	0.18	0.18
Associate Degree (Vocational)	0.04	0	1	0.04	0.04	0.04	0.04
Associate Degree (Academic)	0.04	0	1	0.03	0.04	0.04	0.04
Bachelor's Degree	0.18	0	1	0.18	0.18	0.18	0.19
Master's Degree	0.07	0	1	0.06	0.06	0.07	0.07
Professional School Degree	0.02	0	1	0.02	0.02	0.01	0.02
Doctorate Degree	0.01	0	1	0.01	0.01	0.01	0.01
Income < $5000	0.03	0	1	0.03	0.03	0.02	0.02
$5000<Income<$7499	0.02	0	1	0.02	0.02	0.02	0.01
$7500<Income<$9999	0.02	0	1	0.02	0.02	0.02	0.02
$10000<Income<$12499	0.03	0	1	0.03	0.03	0.03	0.02
$12500<Income<$14999	0.03	0	1	0.03	0.03	0.03	0.02
$15000<Income<$19999	0.04	0	1	0.04	0.04	0.04	0.04
$20000<Income<$24999	0.05	0	1	0.06	0.06	0.05	0.05
$25000<Income<$29999	0.06	0	1	0.06	0.06	0.06	0.05
$30000<Income<$34999	0.06	0	1	0.07	0.06	0.06	0.06
$35000<Income<$39999	0.05	0	1	0.06	0.06	0.05	0.05
$40000<Income<$49999	0.09	0	1	0.09	0.09	0.09	0.09
$50000<Income<$59999	0.09	0	1	0.09	0.09	0.09	0.09
$60000<Income<$74999	0.11	0	1	0.12	0.11	0.11	0.11
Income>$75000	0.32	0	1	0.29	0.31	0.33	0.35
Married-Spouse Present	0.41	0	1	0.41	0.41	0.41	0.41
Married-Spouse Absent	0.01	0	1	0.01	0.01	0.01	0.01
Widowed	0.05	0	1	0.05	0.05	0.05	0.05
Divorced	0.08	0	1	0.08	0.08	0.08	0.08
Separated	0.02	0	1	0.02	0.02	0.02	0.02
Never Married	0.23	0	1	0.22	0.23	0.23	0.24
All Own Kids 0-2 years old	0.06	0	1	0.06	0.06	0.06	0.06
All Own kids 3-5 years old	0.06	0	1	0.06	0.06	0.06	0.06
All Own kids 6-13 years old	0.12	0	1	0.13	0.12	0.12	0.12
All Own kids 14-17 years old	0.08	0	1	0.08	0.08	0.08	0.08
Employed	0.63	0	1	0.62	0.63	0.63	0.63
Unemployed	0.03	0	1	0.03	0.03	0.03	0.03
Not in the Labor Force	0.34	0	1	0.35	0.34	0.34	0.34
Average Income (by CBSA)	53724	26009	81803	53678	53732	53782	53782
Fraction Urbanized (by CBSA)	0.88	0	1	0.88	0.88	0.88	0.88

Table 2: Volunteering by Organization in the CPS

	2004	2005	2006	2007
Religious Organization	34.5%	35.9%	38.2%	36.4%
Children's Education, sports or Rec.	21.6%	22.4%	23.9%	22.7%
Other Educational Group	4.3%	4.5%	4.8%	4.5%
Social and Community Service	17.3%	17.9%	14.3%	18.2%
Civic Organization	4.3%	4.5%	4.8%	4.5%
Cultural or Arts Organization	2.2%	2.2%	2.4%	2.7%
Environmental or Animal Care	2.2%	2.2%	2.4%	2.7%
International Organization	0.4%	0.4%	0.2%	0.5%
Business or Professional Group	1.3%	0.9%	0.5%	0.9%
Political Group	1.7%	1.3%	1.4%	0.9%
Public Safety Group	1.3%	1.3%	1.4%	0.9%
Sports or Hobby Group	2.2%	0.9%	1.0%	1.8%
Youth Services Group	3.0%	2.2%	1.9%	3.2%
Other Type of Organization	3.9%	3.1%	2.9%	0.0%

Table 3: Area Volunteering and Individual Choices

Baseline Regressions

INDIVIDUAL CONTROLS	NO	YES	YES	NO	YES	YES
AREA CONTROLS	NO	NO	YES	NO	NO	YES
	(1)	(2)	(3)	(4)	(5)	(6)
	ALL VOLUNTEERING			**NONCHURCH**		
		OLS			OLS	
Area-Level Volunteering	0.895	0.689	0.675	0.86	0.631	0.607
	[27.40]**	[15.14]**	[18.84]**	[23.53]**	[13.59]**	[15.90]**
N	81882	69378	69378	81026	68652	68652
		PROBIT			PROBIT	
Area-Level Volunteering	2.699	2.23	2.182	3.084	2.382	2.266
	(0.898)	(0.729)	(0.713)	(0.839)	(0.608)	(0.577)
	[19.93]**	[12.88]**	[16.96]**	[17.27]**	[11.06]**	[13.78]**
N	81882	69378	69378	81026	68652	68652

Note: All regressions additionally include controls for year, probability weighting for sample inclusion, robust standard errors, and clustering of standard errors at the state-cbsa level. Coefficients are included with t-statistics in square brackets. Marginal effects for the Probit regression are included in parenthesis. The first three columns of regressions use the outcome of whether individuals volunteered in general while the second three columns use the outcome of whether individuals volunteered for a non-religious institution. * indicates significance at the 5% level, ** indicates significance at the 1% level.

Table 4: Area Effects on Individual Volunteering

Probit and IVProbit Regressions

INDIVIDUAL CONTROLS	NO	YES	YES	NO	YES	YES
AREA LEVEL CONTROLS	NO	NO	YES	NO	NO	YES
	(1)	(2)	(3)	(4)	(5)	(6)
	ALL VOLUNTEERING			**NONCHURCH**		
	PROBIT			PROBIT		
Area-Level Volunteering	2.699	2.23	2.182	3.084	2.382	2.266
	(0.898)	(0.729)	(0.713)	(0.839)	(0.608)	(0.577)
	[19.93]**	[12.88]**	[16.96]**	[17.27]**	[11.06]**	[13.78]**
N	81882	69378	69378	81026	68652	68652
	IVPROBIT			IVPROBIT		
Instrument:	2-Lag All Volunteering			2-Lag Non-Church Volunt.		
Area-Level Volunteering	3.140	2.540	2.861	3.741	2.798	3.257
	[27.87]**	[12.82]**	[13.84]**	[23.18]**	[10.23]**	[9.98]**
N	81882	69378	69378	81026	68652	68652
chi-squared	67.28	10.94	15.85	57.25	9	12.97
p-value	0.000	0.001	0.000	0.000	0.003	0.000

Note: All regressions additionally include controls for year, probability weighting for sample inclusion, robust standard errors, and clustering of standard errors at the state-cbsa level. Coefficients are included with t-statistics in square brackets. The first three columns of regressions use the outcome of whether individuals volunteered in general while the second three columns use the outcome of whether individuals volunteered for a non-religious institution. Instrumental Probit regressions show the Chi-squared of exogeneity as well as the associated p-value.* indicates significance at the 5% level, ** indicates significance at the 1% level.

Table 5: Area Effects on Individual Volunteering

IV Probit and IV Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
INDIVIDUAL CONTROLS	NO	YES	YES	NO	YES	YES	NO	YES	YES
AREA LEVEL CONTROLS	NO	NO	YES	NO	NO	YES	NO	NO	YES
OUTCOME: INDIVIDUAL NON-CHURCH VOLUNTEERING									
	IVREG			IVREG			IVREG		
Instrument:	2-Lag Non-Church Vol.			2-Lag Church Vol.			2-Lag Church & Non-Ch.		
Area Non-Church Vol.	1.035	0.74	0.864	1.083	0.785	0.946	1.042	0.746	0.876
	[37.07]**	[13.08]**	[11.18]**	[25.55]**	[10.86]**	[7.07]**	[38.22]**	[13.27]**	[11.68]**
N	81026	68652	68652	81026	68652	68652	81026	68652	68652
	IVPROBIT			IVPROBIT			IVPROBIT		
Instrument:	2-Lag Non-Church Vol.			2-Lag Church Vol.			2-Lag Church & Non-Ch.		
Area Non-Church Vol.	3.741	2.798	3.257	3.939	2.974	3.626	3.774	2.825	3.314
	[23.18]**	[10.23]**	[9.98]**	[20.66]**	[8.45]**	[6.37]**	[22.85]**	[10.10]**	[10.17]**
N	81026	68652	68652	81026	68652	68652	81026	68652	68652
chi-squared	57.25	9	12.97	35.32	5.83	6.14	61.74	9.71	14.49
p-value	0.000	0.003	0.000	0.000	0.016	0.013	0.000	0.002	0.000

Note: All regressions additionally include controls for year, probability weighting for sample inclusion, robust standard errors, and clustering of standard errors at the state-cbsa level. Coefficients are included with t-statistics in square brackets. All regressions use non-church volunteering as the outcome. The first three columns of regressions use the two-year lag of average non-church volunteering, the second three columns use the two-year lag of average church volunteering, and the last three columns use both the previously mentioned instruments in combination. Instrumental Probit regressions show the Chi-squared of exogeneity as well as the associated p-value.* indicates significance at the 5% level, ** indicates significance at the 1% level.

Appendix Table 1: Area Effects on Individual Volunteering

OLS and IV Regressions

INDIVIDUAL CONTROLS	NO	YES	YES	NO	YES	YES
AREA LEVEL CONTROLS	NO	NO	YES	NO	NO	YES
	(1)	(2)	(3)	(4)	(5)	(6)
	ALL VOLUNTEERING			**NONCHURCH**		
	OLS			OLS		
Area-Level Volunteering	0.895	0.689	0.675	0.86	0.631	0.607
	[27.40]**	[15.14]**	[18.84]**	[23.53]**	[13.59]**	[15.90]**
N	81882	69378	69378	81026	68652	68652
	IVREG			IVREG		
Instrument:	2-Lag All Volunteering			2-Lag Non-Church Volunt.		
Area-Level Volunteering	1.038	0.785	0.878	1.035	0.74	0.864
	[46.44]**	[16.17]**	[15.43]**	[37.07]**	[13.08]**	[11.18]**
N	81882	69378	69378	81026	68652	68652

Note: All regressions additionally include controls for year, probability weighting for sample inclusion, robust standard errors, and clustering of standard errors at the state-cbsa level. Coefficients are included with t-statistics in square brackets. Marginal effects for the Probit regression are included in parenthesis. The first three columns of regressions use the outcome of whether individuals volunteered in general while the second three columns use the outcome of whether individuals volunteered for a non-religious institution. * indicates significance at the 5% level, ** indicates significance at the 1% level.

Appendix Table 2: Relationship between Area Effects
First Stage and Relevant Statistics

INDIVIDUAL CONTROLS	NO	YES	YES
AREA LEVEL CONTROLS	NO	NO	YES
	(1)	(2)	(3)
OUTCOME:		Area Volunteering	
2-Lag All Volunteering	0.672	0.646	0.432
	[10.87]**	[11.06]**	[7.74]**
N	94046	75805	75805
OUTCOME:		Area Non-Church Vol.	
2-Lag Non-Church Vol.	0.611	0.583	0.337
	[9.89]**	[10.06]**	[6.97]**
N	94046	75805	75805
OUTCOME:		Area Non-Church Vol.	
2-Lag Church Vol.	0.468	0.432	0.201
	[4.58]**	[4.64]**	[4.23]**
N	94046	75805	75805
OUTCOME:		Area Non-Church Vol.	
2-Lag Non-Church Vol.	0.543	0.522	0.308
	[10.35]**	[10.31]**	[6.28]**
2-Lag Church Vol.	0.167	0.158	0.099
	[3.08]**	[3.06]**	[2.44]*
N	94046	75805	75805
Hansen J-Statistic	*1.361*	*0.657*	*0.393*
p-value (Hansen J)	*0.23431*	*0.41755*	*0.53062*
Cragg Donald F-Statistic	*60.23*	*59.68*	*27.25*
p-value (Cragg Donald F)	*0.000*	*0.000*	*0.000*

Note: All regressions additionally include controls for year, probability weighting for sample inclusion, robust standard errors, and clustering of standard errors at the state-cbsa level. Coefficients are included with t-statistics in square brackets. The Hansen J and the Cragg Donald F-statistics refer to the ivreg linearized model of the regression rather than the ivprobit model. * indicates significance at the 5% level, ** indicates significance at the 1% level.