Detecting Spillover in Social Networks: Design and Analysis of Multilevel Experiments*

Margaret McConnell
California Institute of Technology

Betsy Sinclair
University of Chicago

Donald P. Green
Yale University

September 21, 2008

*Please do not cite this paper without explicit permission from the authors. This research was funded by a grant from the James Irvine Foundation as part of its California Votes Initiative, a multi-year effort to increase voting rates among infrequent voters - particularly those in low-income and ethnic communities - in California's San Joaquin Valley and targeted areas in Southern California. The Initiative also aims to discern effective approaches by which to increase voter turnout and share those lessons with the civic engagement field. For more information about the Initiative, see http://www.irvine.org/evaluation/program/cvi.shtml. The James Irvine Foundation bears no responsibility for the content of this report. We thank Delia Bailey, participants of the St. Louis Area Methods Meeting, participants of the Networks in Political Science Meeting, participants of the Nuffield Networks meeting and participants of the Annual Meeting of the Society of Political Methodology 2008 for their comments.

1 Abstract

Randomized experiments are seen as the most rigorous methodology for measuring causal effects in the social sciences. A fundamental assumption underlying the analysis of randomized experiments is that of SUTVA (Stable Unit Treatment Value Assumption). SUTVA states that there is no interference between units; the assignment of an individual to treatment should have no effect on outcomes for individuals assigned to control. SUTVA rules out spillovers, effects that are transmitted through social networks, as treated individuals communicate with control individuals. While this assumption is fundamental to analysis, field experiments are rarely attentive to the possibility of spillovers. Multilevel environments provide excellent opportunities for designing experiments which allow for explicit tests of spillovers. This paper presents a social network model that gives a theoretical structure for modeling spillovers under different assumptions about communication in social networks. We then develop a corresponding statistical model in which to analyze how spillovers may introduce bias in the estimation of treatment effects. We analyze actual data from an experiment conducted in California during the November 2006 general election that was designed to measure spillovers at both the precinct and household level. We also provide a recommendation for how to design multilevel experiments in order to test explicitly for spillovers.

2 Introduction

Many randomized field experiments take place within social settings, where the participants of the experiment may communicate with each other. This communication has implications to the analysis of the efficacy of treatment within these settings, in particular if the treatment consists of information. This may result in violations of the stable unit treatment value assumption (SUTVA, as labeled in Rubin 1980), invoked when drawing causal inferences about the efficacy of the treatment. SUTVA states that "The potential outcomes for any unit do not vary with the treatments assigned to any other units, and there are no different versions of the treatment" (Rubin 1986). The first part of SUTVA assumes that there is no interference between units; that is, it assumes that there are no spillover effects.[1] In this paper, we will focus exclusively on the problem of interference between units as a result of treatment, specifically treatment spillovers. Assuming that there is no interference among units is based upon the assumption that the participants in the experiment will not communicate the treatment to each other – that the control group will not be exposed to the treatment from the treatment group, and also that the treatment group will not communicate the treatment to each other. However within social settings, interference among units is possible. In this paper, we explore a methodology to adopt the standard randomized field experiment to these cases.

As is the case in many of the field experiments in political science, if the treatment is a voter mobilization message, it is not only possible but in fact likely that the units do interfere with each other. For example, suppose that as one member of a household heads to a polling place she offers to drive with another household member. Across households, imagine that some fraction of parents are contacted and encouraged to turn out to the next election – it seems likely that they might then discuss the election with the other parents at their neighborhood school. In these cases there would be interference between units, a violation of the SUTVA assumption. In this paper, we will investigate the extent to

[1] The second part of SUTVA assumes that the treatment is the same for each unit. Says Rubin, "SUTVA is violated when, for example there exist unrepresented versions of treatments or interference between units" (Rubin 1986, p. 961).

which we could measure potential spillovers and also design experiments in order to be able to correct for their potential effects. We contend that spillover effects could exist within households, across households or across precincts. We work to estimate these spillover effects and recommend a strategy of precinct-level randomization in conjunction with household-level or individual-level randomization in order to be able to measure them.[2] The two levels of randomization provide for us a control group unfettered by the potential spillover and allow us to measure this effect, similar to the "two control group" recommendation of Holland (1987).

To a large extent, the field experiment literature has ignored the possibility of spillover effects, in part because it is difficult to estimate them. There is a small literature in political science which, via random assignment, can measure spillover effects and also a growing literature in economics and epidemiology on models and estimates of contagion. Empirically scholars have observed within-household interference; in a voter mobilization experiment, Nickerson (2008) finds higher levels of turnout in two-person households when only one of the individuals is contacted to get out to vote via door-to-door canvassing. In this instance, there is interference across within the households. Nickerson finds that 60% of the propensity to vote can be passed onto the other member of the household – a precise measurement of treatment spillover.

While there have been few other measurements of potential spillovers in the political science literature, some of the economics literature has begun to adopt measurements of spillover effects in randomized field experiments, in particular those which incorporate

[2]There is a second more subtle way in which we can observe SUTVA violations as well. It is possible that the interference between units – the spillover of communication within households, across households or across precincts about the mobilization treatment – has the possibility of actually changing the treatment effect itself. Thus any measurement of a treatment effect for the individuals in the study by definition will also include a potential interaction either with other individuals who have received the treatment or with other individuals in the control group. It is possible that either of these interactions can have a result on the individual's behavior and these interactions are outside the scope and control of the experiment. With interaction effects, the individuals in this study are no longer each receiving the same treatment; a violation of the first half of the SUTVA assumption. This makes it extremely difficult to estimate the treatment effect absent such interactions. Here again we recommend a strategy of precinct-level randomization where we believe that there is little likelihood that precincts will interact with each other. This strategy is consistent with other group-level experiments (Rubin, Stuart and Zanutto 2004, Blitstein et al. 2005).

estimates of the effect of individual's social networks on outcomes of interest. This work has not found large spillover effects but are not necessary studies suited for this type of analysis. In a set of experiments on Kenyan agriculture, Duflo, Kremer and Robinson (2006) study the adoption of fertilizer techniques within networks. While this study does find that it is possible to learn to use different fertilizer techniques by observing a demonstration on someone else's farm, they find no effect of diffusion in the adoption of fertilizer techniques within social or geographic groups. In contrast, Munchi (2004) finds evidence of information flow and social learning in the development of different farming techniques for wheat farmers in India and Besley and Case (1994) find similar effects in the adoption of particular seeds within geographic groups for farmers in India. Yet it is possible that these estimates suffer from the reflection problem (Manski 1993) wherein the outcomes of neighbors may be correlated because they face common and unobserved shocks, rather than because they are in fact communicating through network ties.

Spillover effects are more commonly discussed within the observational literature in political science. Huckfeldt (1979) discusses the extent to which political activity occurs within a social context and the relationship between participation and social cues. His work suggests the possibility that individual characteristics, attributes, and personality factors do not entirely determine the extent of individual political activity but that people respond to political events, cues, and opportunities which are specific to a given environment, in particular, external social factors. Other scholars also have emphasized the influence that political discussion within individuals' social networks has on their participation choices. Knoke (1990) finds that individuals with politicized social networks are more likely to engage politically. McClurg (2003) finds the information provided by those networks to also be key. Other researchers find a similar pattern (Bolton 1972; Briet, Klandermans and Kroon 1987; Gerlach and Hine 1970; McAdam and Paulsen 1993; McAdam 1986) This work suggests that social networks influence an individual's subsequent decision to turn out and vote. There has been a small amount of research which has documented the possibility of social transmission of treatment effects via randomized field experiments. Green, Gerber and Nickerson (2003) find within household spillover effects from door-to-door canvassing: an increase of 5.7 percentage points for other household members among households of younger voters. Nickerson (2008) documents within household spillover effects – 60% of the effect from door-to-door canvassing is carried over into the other household member in two member households.

In this paper we provide a strategy for measuring the magnitude of social network effects within a randomized field experiment setting. Randomized experiments are often conducted within a multilevel setting. With limited exceptions, random assignment is typically done at *either* the individual or population level, often precincts.[3] If individuals are assigned to treatment and control groups but treatments are applied at the precinct level, it is possible that within the treated precincts spillovers could occur as individuals who have received treatment transmit to the treatment to others. This could create biases in the estimation of treatment effects. Here we consider the standard setup for a voter mobilization experiment and examine ways an alternative randomization framework provides leverage for estimating these spillover effects.

Here we pursue a strategy of multilevel randomization at the levels where spillovers are likely to happen, based upon our assumptions about individuals' social networks. In this paper, we explore spillovers both the household and precinct level; this enables us to estimate the magnitude of the spillover effects within households and within precincts. A key assumption for the identification of these effects is that we have correctly specified where they are likely to occur. Thus, for example, we assume that there will not be spillovers across voting precincts but instead only within households and across households within precincts. The particular specification of the network spillover model permits us to still invoke SUTVA, although a mis-specified model would still result in a SUTVA violation.

In order to demonstrate how these effects are likely to occur within an individual's

[3] The random assignment may occur at the precinct level while the implementation of the treatment may occur at the individual level. A literature exists already to discuss group-level experiments (see Bloom 2007 for a review of the literature), and the concern with this literature is the relevant unit of observation.

social framework, we have designed two sets of examples. The first set of examples are theoretical network models and are designed to demonstrate how different models of spillovers generate different implications for evaluating the effectiveness of a treatment and the potential biases introduced by spillovers in terms of calculating a treatment effect. The second set of examples are drawn from a field experiment conducted in Los Angeles County in November 2006. Here we randomized at both the household and precinct level and are able to evaluate the extent to which we observe spillovers in this context. The presence of spillovers has the potential to bias treatment effects and thus the inclusion of spillover measurements is important for accurate estimation.

The paper proceeds as follows – we first present our model, where we develop two separate assumptions about the process of mobilization spillover. Spillovers will occur on two levels, the precinct level and the household level. Based upon these assumptions, we present statistical models that allow us to estimate an intent-to-treat effect and a treatment-on-treated effect while also measuring spillovers and will compare these effects to those estimated with the classic experimental design framework. Our aim is to compare these estimates to those obtained using conventional estimators. We analyze the effect of the treatment as well as spillover effects. We conclude with recommendations for both estimation and experimental design which allow for the inclusion of social network communication within an experiment. We find weak spillovers, where the top end of the confidence intervals around the estimates allow for the possibility that spillovers may be a potential source of bias.

3 Model

Each individual i belongs to a precinct p for $p = 1, \ldots, P$. Each individual has a latent probability of voting q_i.

Instead of atomistic individuals, we assume that individuals are part of a social network. An individual's connections are represented by the matrix N. Each individual i's contacts are represented by a row vector of length J such that $N_{ij} = 1$ if individual j is a direct network contact of individual i and $N_{ij} = 0$ if individual j is not a direct contact of individual i. The number of direct contacts of each individual is represented by $n_i = \sum_{j \in N_i}$, where $j \in N_i$ refers to the set of social contacts of individual i. We apply this model to an environment with two specific subnetworks: precincts and households. We define P_i and H_i to be the set of individuals $j \in N_i$ that are members of the precinct or household of individual i. The number of individuals in individual's precinct is represented by $p_i = \sum_{j \in P_i}$, while the number of individuals in individual's household is represented by $h_i = \sum_{j \in H_i}$. By definition $H_i \subset P_i$ and therefore $p_i > h_i$. We focus on two specific network relationships, the relationships between household members and the relationship between precinct residents.[4] Our networks are symmetric networks. This means that $N_{i,j} = 1$ implies $N_{j,i} = 1$.

The matrix of network connections for the entire population is partitioned into submatrices that represent precincts. If N is the number of all individuals $i = 1, \ldots, N$ then this set of individuals can be divided so that all individuals belong to only one of the set of precincts: $p = 1, \ldots, P$. An individual's network N_i in this case consists of all j such that $p_i = p_j$. Individuals are also members assigned to households of size 1, 2, 3 or 4. Individuals are either isolated in the network, or a member of a diad: $N_{i,j} = 1$, triad: $N_{i,j} = 1, N_{j,k} = 1, N_{i,k} = 1$ or network of 4: $N_{i,j} = 1, N_{j,k} = 1, N_{i,k} = 1, N_{i,m} = 1, N_{j,m} = 1, N_{k,m} = 1$. In our setup each individual i will be connected to all other individuals within their precinct, and consequently by default every individual in their household. What differs about the connections for precincts and households is the magnitude of the effect of communicating the mobilization treatment; we expect the type of communication within households to be stronger than the communication within precincts, as household relationships are likely stronger ties.

[4]The structure of our model is general enough to incorporate different network structures. For example, a more continuous definition of neighbors would be the following. Individuals are geographically connected to a certain number of their direct neighbors such that for example, $N_{i,i+1} = 1$ and $N_{i,i-1} = 1, \forall i$. Individuals in this set-up are connected to individuals in other precincts only if they are direct neighbors.

3.1 Base Voting Rates

Two places where heterogeneity will exist in this model are in terms of the ability of the campaign to contact individuals and in the probability of each individual voting. This heterogeneity can be described by each individuals' social connections – this heterogeneity depends on whether these individuals are socially connected to their housemates or their precinct-mates. In this setup we have assumed that if someone is reachable via their household, then they are reachable via their precinct. This is because we assume that the household contacts are a subset of the precinct contacts.

We anticipate that there will be heterogeneity in the ability of a campaign's ability to contact individuals based upon their social networks. Campaigns are typically not able to reach everyone with a mobilization message. In the case where we anticipate positive spillovers, it is possible that the campaign will be able to reach some individuals indirectly, through their social contacts. We therefore write out the probabilities that each individual will receive the message based upon their social connections. We assume that individuals who are directly reachable via a mobilization message have a probability of being contacted p_r, that individuals who are reachable via someone in their precinct have a probability of being contacted p_{pr}, that individuals who are reachable via someone in their household have a probability of being contacted p_{hr}, and that individuals who are not reachable have a probability of being contacted p_{nr}.

Each individual is assigned to one of these categories, so that if an individual is household-reachable then by definition she is not individually-reachable, and similarly if an individual is precinct-reachable then she is neither household-reachable or individually-reachable. Thus each of these probabilities is mutually exclusive. Note that these are not the ex-post outcomes after the mobilization message has been delivered but rather the ex-ante probabilities that an individual will be reachable with the message.

We next describe heterogeneity in voting that is related to the campaign's ability to contact individuals based upon their social networks. We will assume that individuals that are directly reachable via a mobilization message have a probability of voting q_r, while individuals who are reachable through a mobilization message to someone in their precinct have a probability of voting q_{pr} and individuals who are reachable only through a mobilization message to someone in their household have a probability of voting q_{hr}. Some individuals are not reachable in any circumstances, we denote them q_{nr}.

3.2 Mobilization Messages

Within this environment, some individuals receive a treatment that consists of a mobilization message. In each period $t = 1, \ldots, T$, individuals either receive a mobilization visit or do not. The variable $m_{i,t} = 1$ if an individual receives a mobilization message in period t and $m_{i,t} = 0$ otherwise. The effect of the mobilization message $m_{i,t}$ on individuals is indicated by the parameter $\alpha \in [0, 1]$. Individuals may also receive a mobilization message indirectly through their friends.

In each period individuals also communicate with their direct contacts about the mobilization visit. In period t, individuals receive a spillover of the mobilization message received by their direct network contacts in the previous time period, period $t - 1$. The increase in an individual's motivation from mobilization in period t is denoted $M_{i,t}$. The expected value of current period motivation is equal to:

$$q_{i,t} = q_{i,t-1} + M_{i,t}$$

We develop two different models for how mobilization is incorporated. In each period t, let the network contacts j of individuals i be ordered such that individuals $j = 1, 2, \ldots$ receive a mobilization message ($m_{j,t} = 1$) while individuals $j = \ldots, n_i - 1, n_i$ receive no mobilization message ($m_{j,t} = 0$).

We model communication such that the marginal effect of each additional message is decreasing as each message is received; thus for an individual, there is a larger effect of receiving the first mobilization message than there is the second. The decay of messages communicated among close connections (the household network) is represented by $\delta \in [0, 1]$ and the decay of messages communicated in a more distant network (the precinct network) is represented by $\gamma \in [0, 1]$. We assume $\delta > \gamma$. Recall that the network contacts of i are ordered such that the first j network contacts for each individual i have received the mobilization message. We can then define precinct spillovers and household spillovers as

$$S^p_{i,t} = \sum_{j \in P_i} p_r \delta^j 1\{m_{j,t-1} = 1\} \text{ and } S^h_{i,t} = \sum_{j \in H_i} p_r \gamma^j 1\{m_{j,t-1} = 1\}$$

This definition of spillovers provides an intuition for the process by which an individual i is affected by the total number of mobilization messages received within her network. We now compare three different models of spillover mobilization where we make different assumptions about the way in which an individual will update her behavior based upon the mobilization messages she receives both directly and indirectly.

3.2.1 Decaying Mobilization

In the decaying mobilization model, all individuals are affected by spillovers, even those who have been contacted by a campaign themselves. Spillovers are represented by a sum over friends who have been contacted with the size of the effect discounted each time an additional friend brings a mobilization message. The size of the mobilization effect that decays over time is allowed to differ for household and precinct spillovers.

$$M_{i,t} = \alpha m_{i,t} + S^h_{i,t} + S^p_{i,t}$$

3.2.2 Saturation Mobilization

In the saturation model, individuals are affected by spillovers only if they do not receive a mobilization message themselves. An individual becomes mobilized if any of her social contacts are mobilized.

$$M_{i,t} = \alpha m_{i,t} + \alpha(1 - m_{i,t}) 1\{S^h_{i,t} > 0 | S^p_{i,t} > 0\}$$

Reasonable assumptions about voter behavior and communication in networks support either of these models of spillovers. It is possible that once a voter receives the mobilization message, she simply decides to vote or not, and is not affected by any additional messages. In this case, the saturation mobilization model is most appropriate. However, it is also possible that as the voter receives more and more mobilization messages from her social connections, her probability of turning out to vote increases. In this case, the decaying mobilization model is most appropriate. Both models represent reasonable processes of communication among voters and theoretically we have no way to distinguish which model is most appropriate. Thus we will need to rely upon empirical estimation to determine which model is most supported by the data generated from a multilevel randomized experiment, where spillovers can be measured directly. In the calculations below, we evaluate the extent to which spillovers would bias estimation of the treatment effect with a traditional mobilization experiment under each of these models of communication.

4 Estimating ITT and TOT with Spillover

We begin with a group of individuals where each individual i belongs to a precinct p from a set of precincts $p = 1, \ldots, P$ and a household h which may have multiple residents. Our setup in this case will differ from the typical randomized experimental design, in which individuals are randomly assigned to treatment group T or control group C throughout the population. Here, under a multilevel design, individuals will be assigned to control or treatment in each of the three levels of the experiment: precinct, household and individual. Precincts are first assigned to treatment and control. The households living in precincts assigned to treatment are then assigned to treatment or control. The individuals living in households assigned to treatment are then assigned to treatment or control.

We follow the classic GOTV setup presented by Green and Gerber (2000). We do not use the Rubin (1986) model of estimating treatment effects because in a GOTV experiment, there is no problem with inadvertently treating people in the control group. In a typical GOTV experiment, some subset of the experimental group is contacted with a mobilization message. Two treatment effects are estimated – the intent-to-treat effect

(ITT) which describes the difference in turnout between an average individual assigned to the treatment group as opposed to the control group, and the treatment-on-treated effect (TOT), which weights this difference by the proportion of individuals in the treatment group who are actually treated. Estimators of the ITT and TOT are consistent.[5] We change the formulas for calculating ITT and TOT in the multilevel setting because of the different levels of treatment assignments. We explore whether or not the presence of spillovers permits the classic estimators of the treatment effect to remain consistent and also develop mechanisms to estimate these quantities with particular spillover models.

4.1 Assignment to Treatment

With a multilevel design, individuals are assigned to control or treatment in multiple levels of the experiment, each level corresponding to the places wherein which social spillovers of the treatment are likely. Here we describe a multilevel design with three levels of the experiment: precinct, household and individual. In a two stage process, precincts are first assigned to treatment and control. The households living in precincts assigned to treatment are then assigned to treatment or control. The individuals living in households assigned to treatment are then assigned to treatment or control. Figure 1 provides a diagram which illustrates the process of assignment.

Figure 1 Goes Here

[5]In the classic setup, not every individual is reachable with a mobilization message. Let ζ be defined as the fraction of the population that is reachable. Prior to the administration of the mobilization message, let p_{nr} be the probability that a non-reachable person votes and let p_r be the probability that a reachable person votes. After the mobilization message is delivered, let $p_r + \beta$ be the probability that the reachable person votes. Then the expected probability of any member of the control group voting is $P_C = \zeta p_r + (1-\zeta)p_{nr}$ and the expected probability that a member of the treatment group will vote is $P_T = \zeta(p_r + \beta) + (1-\zeta)p_{nr}$. This means that we can derive an expression for the effect of contact in the experiment which is $\beta = \frac{P_T - P_C}{\zeta}$. Then, although we do not observe the population probabilities, we can use both the sample data and the law of large numbers to note that $plim V_T = P_T$ and $plim V_C = P_C$ where V_T is the percentage of the treatment group that votes (and likewise for V_C). Also, $\frac{N_r}{N_T} = \zeta$ where N_r is the number of subjects in the treatment group who were reached and N_T is the number of subjects in the treatment group. We then have a consistent estimator of β, where the intent-to-treat effect is defined as $V_T - V_C$ and the treatment-on-treated effect is defined as $plim \frac{V_T - V_C}{\frac{N_r}{N_T}} = \beta$.

We can therefore assign individuals on three levels, P: precinct, H: household and I: individual - PHI. The two stage assignment of individuals implies that there will ultimately be four randomly assigned groups. There will be individuals assigned to a treatment precinct, a treatment household and assigned to receive an individual mobilization message: TTT, individuals assigned to a treatment precinct, a treatment household and assigned not to receive an individual mobilization message (control): TTC, individuals assigned to a treatment precinct, a control household and assigned not to receive an individual mobilization message (control): TCC and lastly individuals assigned to a control precinct, a control household and assigned not to receive an individual mobilization message (control): CCC. We can then define the average voting rates after mobilization efforts under the two different models about communication of mobilization messages.

5 ITT and TOT with Multilevel Design

5.1 Estimation of Intent-to-Treat Effects

As shown by Gerber and Green (2000), with a simple control and treatment design where there are no spillovers, treatment effects can be consistently estimated by comparing means in treatment and control groups. We will now examine the results of a comparison of treatment and control groups in the presence of household and precinct spillovers. We first analyze the results of a simple comparison of means when random assignment occurs by individual only, as this is the most direct comparison for our strategy of randomization to the standard setup. In this case, there is only one control group and individuals in groups TTC, TCC, CCC are all grouped together into a monolithic group denoted C, as only individuals in group TTT are individually treated. Let p_{TTT} be the share of individuals in group TTT, p_{TTC} be the share of individuals in group TTC, p_{TCC} be the share of individuals in group TCC and $1 - p_{TTT} - p_{TTC} - p_{TCC}$ be the share of individuals in group CCC.

Consistent with the Gerber and Green setup, we first calculate the expected probability of voting within each of these groups of voters in the hypothetical case that that every individual assigned to treatment was successfully contacted. We will then consider the estimation of treatment effects and consider potential biases introduced by the existence of spillovers. We consider voting rates under the two previously defined models of communication:decaying mobilization and saturation mobilization.

5.1.1 Decaying Mobilization

In a decaying mobilization model of communication, average voting rates are as follows:

$$
\begin{aligned}
V_{TTT} &= (q_r + \alpha + S_p + S_h)p_r + (q_{hr} + S_p + S_h)p_{hr} + (q_{pr} + S_p)p_{pr} + (q_{nr})p_{nr} \\
V_{TTC} &= (q_r + S_P + S_h)p_r + (q_{hr} + S_p + S_h)p_{hr} + (q_{pr} + S_p)p_{pr} + (q_{nr})p_{nr} \\
V_{TCC} &= (q_r + S_P)p_r + (q_{hr} + S_p)p_{hr} + (q_{pr} + S_p)p_{pr} + (q_{nr})p_{nr} \\
V_{CCC} &= (q_r)p_r + (q_{hr})p_{hr} + (q_{pr})p_{pr} + (q_{nr})p_{nr}
\end{aligned}
$$

In the case of decaying mobilization, the average voting rate of the control group will equal:

$$
\begin{aligned}
V_C = {} & \tfrac{p_{TTC}}{1-p_{TTT}}((q_r + S_P + S_h)p_r + (q_{hr} + S_p + S_h)p_{hr} + (q_{pr} + S_p)p_{pr} + (q_{nr})p_{nr}) \\
& + \tfrac{p_{TCC}}{1-p_{TTT}}(q_r + S_P)p_r + (q_{hr} + S_p)p_{hr} + (q_{pr} + S_p)p_{pr} + (q_{nr})p_{nr}) \\
& + \tfrac{p_{CCC}}{1-p_{TTT}}((q_r)p_r + (q_{hr})p_{hr} + (q_{pr})p_{pr} + (q_{nr})p_{nr})
\end{aligned}
$$

Assuming that the assignment groups (TTC,TCC,CCC) are all assigned in equal proportion and that the assignment group (TTT) is equal to their sum, we can then write:

$$V_C = p_r(q_r + \tfrac{2}{3}S_p + \tfrac{1}{3}S_h) + p_{hr}(q_{hr} + \tfrac{2}{3}S_p + \tfrac{1}{3}S_h) + p_{pr}(q_{pr} + \tfrac{2}{3}S_p) + p_{nr}(q_{nr})$$

We then calculate the standard $ITT = V_{TTT} - V_C$ as though we were unaware of the spillover effect:

$$V_{TTT} - V_C = \left(p_r(\alpha + \frac{1}{3}S_p + \frac{2}{3}S_h) + p_{hr}(\frac{1}{3}S_p + \frac{2}{3}S_h) + p_{pr}(\frac{1}{3}S_p) \right)$$

Note that this quantity is larger than $\alpha * p_r$, which is the unbiased intent-to-treat effect.

In the case of non-zero spillovers, the intent-to-treat effect will be inflated when there is decaying mobilization without proper accounting for the spillover effect. This goes against the intuition that the existence of spillovers will depress estimates of the treatment effect. Additionally, this indicates that with the presence of spillovers, it is not sufficient to have a single control group or single level of randomization.

5.1.2 Saturation Mobilization

In a saturation mobilization model of communication, average voting rates are as follows:

$$\begin{aligned}
V_{TTT} &= (q_r + \alpha)p_r + (q_{pr} + \alpha)p_{pr} + (q_{hr} + \alpha)p_{hr} + (q_{nr})p_{nr} \\
V_{TTC} &= (q_r)p_r + (q_{pr} + \alpha)p_{pr} + (q_{hr} + \alpha)p_{hr} + (q_{nr})p_{nr} \\
V_{TCC} &= (q_r)p_r + (q_{pr} + \alpha)p_{pr} + (q_{hr})p_{hr} + (q_{nr})p_{nr} \\
V_{CCC} &= (q_r)p_r + (q_{pr})p_{pr} + (q_{hr})p_{hr} + (q_{nr})p_{nr}
\end{aligned}$$

In the case of saturation mobilization, the average voting rate of the control group will equal:

$$V_C = \tfrac{p_{TTC}}{1-p_{TTT}}\left((q_r)p_r + (q_{pr} + \alpha)p_{pr} + (q_{hr} + \alpha)p_{hr} + (q_{nr})p_{nr}\right)$$
$$+ \tfrac{p_{TCC}}{1-p_{TTT}}\left((q_r)p_r + (q_{pr} + \alpha)p_{pr} + (q_{hr})p_{hr} + (q_{nr})p_{nr}\right)$$
$$+ \tfrac{p_{CCC}}{1-p_{TTT}}\left((q_r)p_r + (q_{pr})p_{pr} + (q_{hr})p_{hr} + (q_{nr})p_{nr}\right)$$

Assuming that the assignment groups (TTC,TCC,CCC) are all assigned in equal proportion and that the assignment group (TTT) is equal to their sum, we can then write V_C as:

$$V_C = \left(q_r p_r + q_{pr}p_{pr} + \tfrac{2}{3}\alpha p_{pr} + q_{pr}p_{pr} + \tfrac{1}{3}q_{hr}p_{hr} + q_{nr}p_{nr}\right)$$

Then we calculate $ITT = V_{TTT} - V_C$ as:

$$V_{TTT} - V_C = \alpha p_r + \frac{1}{3}(\alpha)p_{pr} + \frac{2}{3}(\alpha)p_{hr}$$

Note that this quantity is larger than $\alpha * p_r$, which is the unbiased intent-to-treat effect.

In the case of non-zero spillovers, the treatment effect will also be inflated when there is saturation mobilization. The intuition behind these findings is that within the treatment precincts, the voters who are not contacted by the campaign can be indirectly reached by others within their household or across households within their precinct. Yet at no point are individuals in the control precincts reachable; as a consequence, it appears that the treated precincts are voting at higher levels where the causal mechanism is due to the treatment, whereas instead the treated precincts are voting at higher levels because of the combination of treatment and spillover.

These two findings demonstrate that the existence of spillovers can bias the treatment effect estimates and that without a precise understanding of the model by which mobilization information is communicated within households or precincts, the multilevel randomization allows us to analyze experimental spillover effects. In particular, however, our interest is not simply that of correcting the potential bias in these estimates. We are particularly interested in estimating the degree to which individuals are communicating about politics to others within their social networks.

6 Spillover Effects: Southwest Voter Experiment, California General Election 2006

The Southwest Voter Registration and Education Project (SVREP), a grassroots political organization in Los Angeles, organized a Get Out the Vote campaign for the 2006 midterm elections targeting first time Latino voters. This study used a two stage random assignment process similar to the one described in the design of the multilevel experiments above. In the first stage, all precincts are randomly assigned to either treatment or control. In the second stage, individuals in the treatment group are randomly assigned to treatment and control groups. The random assignment of individuals to treatment and control group creates a random assignment of household composition that is equivalent to the random assignment of households to treatment and control after conditioning on household size. We propose this study as an example for how to conduct randomization and demonstrate potential mechanisms by which to measure spillover effects. In this experiment, Latino voters were identified from voter rolls by their surname. Voters were eligible for SVREP's campaign if they had registered after August 1, 2004 and if they had not voted in any primary or general election since 1998. The study is composed of five districts in Los Angeles that are heavily populated with Latino voters; the average precinct in the study contains 48% Latino voters. SVREP's GOTV activities consisted of mailings, multiple phone calls and visits by canvassers. The campaign is described in detail in Michelson et al (2008). In the analysis that follows, we will consider the effect of campaign contact, broadly defined, on voter turnout. In an average precinct assigned to treatment, the contact rate was 24%. The vast majority of these contacts were made by phone by motivated volunteers, who were given a script to read (as seen in the Appendix) and who were given permission to depart from their script and speak more broadly about the need to participate in the election. Each volunteer asked to speak to a specific member of the household when making the phone call and after the conversation, manually recorded whether or not each individual was successfully contacted. Volunteers were primarily residents of the community, often recruited from youth centers. Volunteers were bilingual and many of the conversations took place in Spanish. Voter turnout was determined from voting records provided the Los Angeles County Registrar Recorder's office.

6.1 Experimental Design

The experimental design consisted of the random assignment of precincts to treatment and control. Treatment precincts were then randomized at the individual levels. A total of 478 precincts were divided up into 432 treatment precincts consisting of 27,928 individuals and 46 control precincts consisting of 2,945 individuals. The 27,928 individuals in the treated precincts were randomized into a treatment and control group at the level of individual voters. Within the treated precincts, there were 23,794 voters in the individual voter treatment group and 4,134 voters in the individual voter control group. The breakdown of treatment assignments is presented in Table 1. Note that here we assign individuals to treatment and control at both the precinct and the individual level.

Table 1 Goes Here

The subject pool is restricted to individuals who live in households of fewer than five registered voters. In Table 2 breaks down the subject pool by household size. The largest number of individuals in the experiment are in one-person households – but we have a sizable number in two-person households as well. Although the randomization occurred by individual and by precinct, and not by household, if we control for household size then we are able to estimate treatment results to examine household spillovers – that is, to compare turnout of individuals from the control group where at least one person in the household was assigned to treatment to individuals where no one in the household was assigned to treatment.

Table 2 Goes Here

We can then look across these groups and see which subset of individuals were most likely to turn out to vote - in our case, individuals who had been assigned to treatment at both the precinct and individual level, as demonstrated in Table 3. These numbers not only suggest an effect of the mobilization campaign but also these voting rates are sorted exactly as we would expect from spillover effects within precincts. The average voting rate of individuals who were assigned to the control group but who live within treated precincts are voting at slightly higher rates than the individuals who were assigned to the control group but who live within controlled precincts. This difference (.62 percent) suggests that in fact individuals who are assigned to the treatment group are communicating about voting to others within their precinct. Our subsequent analysis will determine whether or not this difference is significant at traditional levels and also look for evidence of household spillovers.

Table 3 Goes Here

In order to illustrate that the randomization achieves balance in the covariates Table 4 presents the means of the individual covariates by randomization assignment. Slightly less than half of the voters in this sample are female. The average age across all groups is relatively young (about 32), which is consistent with the selection criteria of being a recently registered voter. Slightly more than half of the individuals in this study are registered Democratic, and almost 30% are registered as ”decline-to-state”. Across all groups, it has been approximately 650 days since they registered. There are a very small number of voters who had participated in a previous election – vote history refers to the number of previous elections that voters have participated in, and is extremely small across all groups (the average is below 6/10ths of an election).[6] None of these differences are statistically significant.

Table 4 Goes Here

Thus we are assured that the randomization was successful in achieving balance across the covariates.[7] We next proceed to analyze both the effect of the treatment as well as to measure the spillover effects within households and across households within precincts.

6.2 Results

In the tables which follow, we estimate three quantities. First, we estimate the effect of the treatment using the traditional statistical model and standard experimental design. This demonstrates that this experiment did have a large treatment effect, which makes this particular experiment a good candidate for which to look at spillovers. We next estimate two models which allow for the incorporation of spillover effects in the treatment effect analysis: we do this first where we look for precinct spillovers and next where we look for household spillovers. We find suggestive results but no statistically significant spillovers, giving additional validity to our initial calculation.

In the first table, we present the treatment-on-treated effect (TOT) , whereas in the tables which follow, we present the intent-to-treat effect (ITT). The intent-to-treat effect is defined as the observed difference in turnout between those assigned to the treatment and control groups. If the contact rate is 100%, the intent-to-treat effect is identical to the treatment-on-treated effect. Generally this is not the case, however, and to calculate the treatment-on-treated effect we must adjust for the contact rate. TOT is defined as the

[6] These are cases where some people got through the algorithm used to determine "new voters", but fortunately it is not only small, it is also very similar across assignment groups.

[7] Additionally we present these covariates when aggregating up to the level of precinct assignment in Table 5 and confirm that we have balance across these groups as well. Note that none of these covariates have statistically significant differences.

intent-to-treat effect divided by the contact rate. Treatment is defined as any campaign contact. We recognize that there is likely correlation between an individual's ability to be contacted and their unobservable characteristics, so we estimate TOT in a regression context with an instrumental variables approach where we use the random assignment variable as an instrument.

Formally, let Z be an indicator which defines whether or not the individual is assigned to the treatment or control. Let x denote whether the individual actually receives the treatment. Let y indicate whether or not the individual voted. Then the intent-to-treat effect is defined as: $ITT = E(y|Z = 1) - E(y|Z = 0)$. Following from this, then $TOT = \frac{ITT}{E(x|Z=1)-E(x|Z=0)}$. We can estimate TOT by regressing y on x and using Z as an instrument.[8]

We begin by calculating the TOT instrumental variables estimate without a measurement of spillover. We do this so that it is possible to compare the magnitude of this coefficient with the effects expected from the field experiment voter mobilization literature. We use the random assignment at the individual level as an instrument and we cluster the standard errors based upon the random assignment to a treated or control precinct or individual. Here we calculate this estimate both with and without covariates, and find that the effect of receiving a campaign contact has a large TOT effect of 8.9 percentage points. When we account for the voter covariates, we observe a TOT effect of 8.4 percentage points. Both of these estimates are statistically significant at traditional levels. That this campaign was so successful at mobilizing voters is surprising given that the campaign contacts were almost completely phone calls, which are not typically considered so efficacious. However, this may be attributable to the particular population contacted in this experiment – first-time Latino voters. As we begin to account for spillover effects, we also find that this particular population may affect our results.

Table 6 Goes Here

We can improve our treatment effect estimates by incorporating a measurement for

[8]In the context of Imbens and Angrist's (1994) definitions, this is the local average treatment effect (LATE) given an inducement to treatment.

spillovers directly, which we do in the next section, we we hope to both measure the treatment effect without bias and to measure the magnitude of any (possible) spillover effects. Whether or not we find statistically significant evidence for spillovers in the next calculations will determine to what extent the treatment effect in Table 6 is correct.

6.3 From Models to Estimates

The section above calculated the treatment effect as though this were a traditional mobilization experiment with a single treatment group. When weighted by the contact rate, this corresponds to the ITT which we described as an overestimate when the saturation or decaying models of voter mobilization were true because it fails to subtract the effect of V_{TC} from the estimate. To get an accurate measurement for the effect of the treatment it is necessary to account for the spillover present in this estimate. Unlike those calculations, however, the size of our assignment groups is not equal. Suppose that we were only interested in accounting for precinct spillovers. To calculate ITT we subtract the spillover effect and weight these quantities by their size. We will need to account for the spillover effect in our empirical analysis. To do so, we incorporate an additional variable into our estimate of the intent-to-treat effect which controls for the individuals who have been randomly assigned to the cases where we believe there is likely to be spillover – in this case, we control for the assignment of individuals to treament-control when calculating the intent-to-treat effect. This will give us an unbiased measurement of the treatment effect.

6.4 Tests of Spillovers

The first spillover exercise is to simultaneously estimate the turnout effect of 1) having been assigned to be a treatment individual in a precinct assigned to treatment and 2) having been assigned to be a control individual in a precinct assigned to treatment. If we observe a statistically significant and positive coefficient for the assignment indicator which describes the control individual in the treatment precinct, then in the intent-to-treat analysis resulting from this estimation reveals a precinct spillover effect. Table 7 presents these estimates.

Table 7 Goes Here

We first estimate the intent-to-treat effect without a control for the spillover effect. Here we find an ITT of 2.1 percentage points. We then compare this to an estimate of the ITT with the inclusion of the spillover control in the second column of Table 7. We observe a coefficient for the individual assigned to actually receive treatment that is positive and significant; this value represents the marginal increase in turnout from increasing the number of individuals randomly assigned to the treat-treat group relative to the control-control group and corresponds to an ITT of 2.5%. We are particularly interested, however, in the coefficient for the assignment of individual control and precinct treatment. This coefficient, with a value of .006, is also positive, but unfortunately is not statistically significant at tradition levels. We repeat the analysis in the second column and find a similar pattern of results. Thus we have no statistically significant precinct spillovers. Note, however, that with the inclusion of the spillover effect we did see a larger estimate for the ITT – this effect suggests that spillover would produce an underestimate of the treatment effect and that the saturation model may be most appropriate.

We next examine evidence for household spillovers. Here, because we conducted the randomization at the individual and precinct levels, we must account for the other members of the households to capture the conditional random assignment. Thus we produce an indicator for those individuals who were assigned to be in the control group within a household where at least one individual was assigned to the treatment group and indicate this in the table which follows as "Control Individual in Treatment Household". We also have an indicator for when an individual is assigned to receive the mobilization treatment which is labeled "Treatment Individual". Then, because again we did not conduct the randomization at the household level, we include a series of additional control variables for if this was a 3 person household and if this was a 4 person household.

We calculate the intent-to-treat effect incorporating our new covariates. These coefficients are presented in Table 8.

Table 8 Goes Here

Here again we observe a coefficient for the individual assigned to actually receive treatment that is positive and significant; this value is the marginal increase in turnout from increasing the number of individuals randomly assigned to the treat-treat group relative to all individuals assigned to the control group. This corresponds to an ITT of 2.4%. Additionally we include the coefficient for individuals who are also residents within the household but are part of the control group; this value is the marginal increase in turnout from increasing the number of individuals assigned to be co-residents of households with a treatment individual relative to the number of individuals assigned to the control group. This value is positive but not statistically significant at traditional levels.

Thus for both households and precincts, while we see suggestive evidence of spillovers in this experiment, the results are not statistically significant.

Finally, we use Hansen's J-statistic to determine if there is overidentification. This is a useful test – because we have conducted the randomization at two levels, we have two instruments, and thus should be able to identify two causal effects – the individual treatment effect and the within-precinct spillover effect. Hansen's J-statistic will test whether or not we are able to estimate the individual-level treatment effects without the second instrument – and thus this will reveal whether or not we do have spillovers. We test the null that both individual and precinct-level assigment variables are valid instruments. The J-statistic is 0.242, corresponding to a p-value of 0.62. We do not have enough evidence to reject the null that both individual and precinct level assignment are valid instruments. This indicates that we can continue to estimate treatment effects without the incorporation of the precinct-level instrument in this particular case, where there is no evidence of statistically significant spillover effects. Yet for alternative experimental setups, this would not necessarily be true. In experiments where there is social interaction, these tests are useful to detect spillover and to correct estimates of the treatment effect.

7 Conclusion

This experiment presents a research strategy which allows for the measurement of spillover within individuals' social networks via a multilevel randomized field experiment. Within this setting, it is possible to design an experiment to both estimate the effect of a stimuli and additionally to measure the degree to which this stimuli will spread throughout an individual's social setting. In this paper we explore this possibility with a mobilization message where voters are located within households and precincts. We demonstrate that in the presence of spillovers SUTVA is violated, but that with a properly specified model and a multilevel randomized experiment it is possible to measure treatment effects without the SUTVA violation, and hence without the introduction of bias into the estimates. We follow our own recommendations with a multilevel randomized field experiment where we examine spillovers within households and across households within precincts. In our experiment, we do not find evidence of spillovers that is statistically significant at traditional levels, although our results are suggestive that there is a possibility that people are in fact to some small extent communicating about the mobilization message. One important note in our findings is that due to the small size of our control group – the individuals who are not contacted within precincts assigned to treatment – our standard errors are particularly large. We note that the spillover effect from the seminal study in this field – Nickerson (2008) which finds that 60% of the mobilization message will be shared to other household members – is included within the 95% confidence interval of our estimates.

There are several other possibilities for the lack statistical significant spillovers in our findings – the first is that this may be a setting within which people are unlikely to be communicating about the message. Our sample population consists of first time voters who are unlikely to have established networks within these communities. Additionally, in this experiment, we randomized at both the individual and precinct level. Had we extended that randomization to the household level, might have been able to more precisely adopt our empirical findings to our theoretical model.

There are many situations in which the social spillover of the treatment is likely. We recommend a strategy of randomization at multiple levels such that it is possible to test hypotheses about the presence of spillovers. Multilevel randomization enables estimation strategies which identify treatment effects under fewer assumptions. Multilevel randomization enables a test of one aspect of SUTVA – interference among units where there is reason to believe that interference is likely to take place. Additionally, multilevel randomization continues to allow for the clustering of standard errors and does not require the assumption of homoskedasticity that a meta-analysis would require to use precinct fixed-effects if ex-post, if the experiment had not been randomized at the precinct level. Applying a strategy in the design of a multilevel randomized field experiment also has the added benefit of exploring more instances where social scientists believe that social interactions are theorized to exist, and to draw valid causal inferences about the impact of those interactions on political behavior.

8 Works Cited

Guido Imbens J. Angrist. 1994. ”Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, Vol. 62, No. 2, pp. 467-476.

Besley, Timothy and Anne Case. 1993. “Modeling Technology Adoption in Developing Countries,” *American Economic Review*, 83 (2), pp. 396-402.

Blitstein, Jonathan L. , David M. Murray, Peter J. Hannan and William R. Shadish. 2005. “Increasing the Degrees of Freedom in Future Group Randomized Trials: The df* Approach,” *Evaluation Review*, 29, 269.

Bloom, Howard S. 2007. “The Core Analytics of Randomized Experiments in Social Science Research,” *The SAGE Handbook of Social Science Methodology*, Editors William Outhwaite and Stephen P. Turner. Sage Publications Ltd.

Bolton, Charles. 1972. “Alienation and Action: A Study of Peace Group Members.” *American Journal of Sociology* 78: 537-61.

Briet, Martien, Bert Klandermans and Frederike Kroon. 1987. “How Women Became Involved in the Women’s Movement of the Netherlands.” In Mary Katzenstein and Carol Mueller, eds., *The Women’s Movements of the United States and Western Europe: Consciousness, Political Opportunities, and Public Policy*. Philadelphia: Temple University Press.

Duflo, Esther, Michael Kremer and Jonathan Robinson. 2006. “Understanding Technology Adoption: Fertilizer in Western Kenya.” Working Paper.

Gerlach, Luther and Virginia Hine. 1970. *People, Power and Change: Movements of Social Transformation*. Indianapolis: Bobbs-Merrill.

Gerber, Alan S. and Donald P. Green. 2000. “The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment.” *American Political Science Review*, Vol. 94, No. 3. pp. 653-663.

Green, Donald P., Alan S. Gerber, and David W. Nickerson. 2003. "Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments," *Journal of Politics*, Vol. 65:1083-96.

Holland, Paul W. 1987. "The Role of a Second Control Group in an Observational Study: Comment." *Statistical Science*, Vol. 2, No. 3 (Aug.), pp. 306-308.

Huckfeldt, R. Robert. 1979. "Political Participation and the Neighborhood Social Context." *American Journal of Political Science* 23, 3 (Aug.): 579-592.

Knoke, David. 1990. "Networks of Political Action: Toward Theory Construction." *Social Forces* 68: 1041-1063.

Rubin, Donald B. 1980. "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment." *Journal of the American Statistical Association*, Vol. 57, No. 371 (Sep.), pp. 591-593.

Manski, Charles. 1993. "Identification of Exogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, pp. 531-542.

McAdam, Doug. 1986. "Recruitment to High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92: 64-90.

McAdam, Doug and Paulsen, Ronnelle. 1993. "Specifying the Relationship Between Social Ties and Activism." *American Journal of Sociology* 99, 3: 640-667.

McClurg, Scott D. 2003. "Social Networks and Political Participation: The Role of Social Interaction in Explaining Political Participation." *Political Research Quarterly* 56, 4 (Dec.): 449-464.

Munshi, Kaivan. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73(1), pp. 185-213.

Nickerson, David. 2008. "Is Voting Contagious? Evidence from Two Field Experiments," *American Political Science Review*, Forthcoming.

Rubin, D. B., E. A. Stuart and E. L. Zanutto. 2004. "A Potential outcomes view of value-added assessment in education," *Journal of Educational and Behavioral Statistics*, 29: 103-116.

9 Appendix A: Excerpt from Phone Bank Script

I'm calling on behalf of the Southwest Voter Registration Education Project. We'd like to remind you to vote this election and we'd like to urge you to vote by mail or on Tuesday, November 7. It's important to vote, many issues will affect your community. Our call is part of the March Hoy, Vota Manana Campaign to turnout Latino voters.

You will be voting for a candidate of your choice for the United States Senate and Congress, Governor, State Legislators and initiatives that set policy that will impact you and your family.

For the United States Senate and Congress, they will be making decisions or the Iraq war, immigration, gas prices, the economy and health care. The Governor and State Legislators will be setting policies that impact your children's education and the environment. And, you will be voting on several initiative that affect our roads, schools, housing water, parks, dams as well as elections and energy sources.

10 Tables and Figures

Table 1: Treatment Assignment

	Individual Control	Individual Treatment	Total
Precinct Control	2,945	0	2,945
Precinct Treatment	4,134	23,794	27,928
Total	7,079	23,794	30,873

Note that individuals and not households are the unit of assignment to treatment and control.

Table 2: Assignment to Treatment and Control by Household Size

Precinct	Individual	1 person	2 person	3 person	4 person
Treatment	Treatment	17769	4820	1047	158
Treatment	Control	3066	850	191	27
Control	Control	2271	566	96	12

Table 3: Overall Voting Rates

Assignment (Precinct)	Assignment (Individual)	Voted
Treatment	Treatment	36.72
Treatment	Control	34.71
Control	Control	34.09

Table 4: Covariate Balance by Randomization Assignment

Variable	Treatment Precinct Treatment Individual	Treatment Precinct Control Individual	Control Precinct Control Individual
Female	0.425	0.419	0.430
Age	32.421	32.162	32.614
Democratic Party	0.541	0.550	0.529
Republican Party	0.122	0.114	0.134
Decline to State	0.294	0.292	0.292
Days since Reg.	652.105	662.136	652.820
Vote History	0.051	0.054	0.058

* Indicates that the differences in means between treatment and control are different at statistically significant levels

Table 5: Covariate Balance by Treatment and Control Precinct Assignment

Variable	Treatment Precinct Mean	Control Precinct Mean
Female	0.44	0.45
	(.50)	(.50)
Percent Age 30 or Younger	0.59	0.58
	(.49)	(.49)
Percent Older than 60	0.07	0.07
	(.25)	(.26)
Percent Republican	0.12	0.13
	(.32)	(.34)
Days since Registration	818.40	817.95
	(256.54)	(259.64)
Voting History	0.05	0.06
	(.42)	(.49)
N	25134	2688

Standard errors are included below the means in parenthesis

None of the differences in means between treatment and control are different at statistically significant levels

Table 6: IV Estimation: Effect of Contact on Vote

Dependent Variable, Turnout Instrument, Individual Random Assignment		
Campaign Contact	0.089**	0.084**
	(.029)	(.029)
Female		0.036**
		(.005)
Age 30 or younger		-0.175**
		(.006)
Older than 60		0.075**
		(.012)
Days Since		-0.000**
Registration		(.000)
Republican Party		-0.069**
		(.008)
Participation in		0.044**
Participation in Past Elections		(.006)
Intercept	0.346**	0.576**
	(.006)	(.013)
N	30873	30873
R^2	0.015	0.064

** 1%, * 5%

Robust standard errors in parenthesis.

Standard errors are clustered by the unit of assignment to treatment or control (precinct or individual)

Table 7: Least Squares ITT Analysis of Precinct Spillovers

Dependent Variable, Turnout	**W/out Covariates Vote**	**W/out Covariates Vote**	**With Covariates Vote**
Treat Prec & Control Ind Assignment		0.006	0.011
		(0.013)	(0.035)
Treat Prec & Treat Ind Assignment	0.021**	0.025**	0.027*
	(0.007)	(0.013)	(0.012)
Female			0.036***
			(0.005)
Age 30 or Younger			-0.177***
			(0.006)
Older than 60			0.081***
			(0.012)
Days since Registration			-0.000***
			(0.000)
Republican			-0.071***
			(0.008)
Voting History			0.043***
			(0.006)
Intercept	0.346***	0.342***	0.575***
	(0.006)	(0.011)	(0.015)
N	30873	30873	30873
F	$F_{(1,27973))} = 9.11$	$F_{(2,27973)} = 4.56$	$F_{(8,27973)} = 220.34$
Clusters	27872	27872	27872

Robust standard errors in parentheses. Standard errors are clustered by level of observation where the assignment to treatment or control occurred.
*** Significant at 1%, Significant at ** 5%, Significant at *10%

Table 8: Least Squares ITT Analysis of Household Spillovers

Dependent Variable, Turnout	
Control in Treat Household	0.022
	(0.018)
Individual Treatment Assignment	0.024***
	(0.008)
2 Person Household	0.029***
	(0.007)
3 Person Household	0.020
	(0.014)
4 Person Household	-0.101***
	(0.033)
Intercept	0.337***
	(0.007)
N	30873
F	$F_{(5,27973)} = 8.38$

Restricted to the sample of households under 5 people.
Robust standard errors in parentheses. Standard errors are clustered by level of observation where the assignment to treatment or control occurred.
*** Significant at 1%, Significant at ** 5%, Significant at *10%

Figure 1: Assignment to Treatment