Political Networks Draft: Political Mobilization

D. Alex Hughes

June 20, 2016

Abstract

Recent scholarship has examined the role of individuals' social networks in shaping political behavior. Although this theory is hardly new, recent advances in measurment and computation have made it possible to adjudicate the role of social networks in politics. In this paper, I present evidence that those who are better socially connected are better able to mobilize others toward political action. To do so, I present novel social network and political data of more than 4,000 individuals living in Honduras. With these individuals, l ran a field experiment in which individuals were randomly assigned to receive an incentive to mobilize other to attend a political event. The evience suggests not only a clear causal effect that those who are more socially connected are uniquely capable to spur political action, but also provides for the first time a possible mechanism for this effectiveness.

Contents

1 Political Mobilization Through Social Pathways

1.1 Introduction

Why do some citizens choose to participate in political and civic activities while others do not? Most previous explanations have focused either on individual characteristics like education and past voting history, or on large-scale, societal characteristics like civic-mindedness. In this article, I test a prediction from the theory of *social information*: political mobilization occurs through actors' social networks. Even more, because of the costly nature of political mobilization – the individual who is mobilized to take action is paying concentrated costs for what are likely to be diffuse benefits – a theory of social information predicts that mobilization is most likely to occur in local networks.

I make the following contributions in this article. First, I build on recent work concerning drawing causal inference in the presence of spillover (Bowers, Fredrickson and Panagopoulos, 2013; Sinclair et al., 2014; Baird et al., 2014; Hudgens and Halloran, 2008). Whereas previous scholars have largely developed methods to cope with spillover as a nuisance parameter by estimating average differences beween clusters not subject to spillover (e.g. Sinclair et al., 2014; Baird et al., 2014), I demonstrate that with spillover-pathway data (e.g. social networks), it is possible to recover causal estimates, at least in some cases. Second, I argue that previous two-stage randomization procedures advocated for inference in the presence of spillover are not a *necessary* precondition for inference, especially if within-cluster spillover can be evaluated. Third, and finally I make several substantive contributions to the understanding of political mobilization drawn from a unique dataset and intervention fielded in Honduras. Using between-village variation in the connectedness of randomly-assigned political mobilizers, I identify that targeting well-connected mobilizers causes a marked increase in town-level political activity. Then, I examine the two plausible mechanisms leading to this effect: either well-connected mobilizers hold sway across the entire network, or well-connected mobilizers hold sway in their (relatively larger) local networks. On the whole, I find that the evidence supports the second of these propositions.

1.2 Existing Explanations

Existing explanations for political mobilization behavior fall into four classes of explanations. The most common explanations in the current literature are individual characteristics, typified by the work of Rosenstone and Hansen (1993) and Brady, Verba and Schlozman (1995). The other four explanations that I identify are supra-individual explanations; these explanations, while recognized in the current literature, play a relatively minor roll compared to the individual characteristics explanations. I describe, in detail, these explanations in subsequent subsections, but briefly they are society-wide characteristics, social group characteristics, and social network explanations.

Perhaps the best, most succinct characterization of these classes of explanations is due to Brady, Verba and Schlozman (1995), who opine:

> "Why do citizens participate in political life? One way to think about this puzzle is to invert he question and ask why people don't take part in politics. Three answers immediately suggest themselves: *Because they can't, because they don't want to, or because nobody asked*," (Brady, Verba and Schlozman, 1995, p. 271).

The features identified in this subsection comport with the categories of motivation identified by Brady, Verba and Schlozman. However, the mapping is not one-to-one. Specifically, individual characteristics may affect both capacity and willingness; in the same way, supra-individual characteristics may affect any of Brady, Verba and Schlozman's categories.

1.2.1 Individual Characteristics

Existing, large-sample observational studies have identified a range of characteristics that are predictive of individual political behavior. These features can be grouped into the broad classes of time, money, and civic skills (Brady, Verba and Schlozman, 1995). While consensus seems to have emerged around these predictive features, for quite some time, this participatory question was at the heart of both American and Comparative behavioral political science.

1.2.2 Society Wide Characteristics

1.2.3 Previous Explanations Identify Social Groups

Participation and mobilization are distinct concepts. While participation is a stable, long-term, equilibrium outcome that has typically been explained using equally long-term independent variables. Accordingly, many scholars explain variance in political participation with characteristic variables like race, level of education, feelings of personal efficacy, and years in residing in community which are unlikely to shift over small periods of time and as such are fixed in the short-run (Rosenstone and Hansen, 1993). Brady, Verba and Schlozman (1995) expand the range of explanatory variables, but maintain a focus on long-term variables of civic skills and resources. For instance, an individual may accrue resources over time, and those may precipitously increase or decrease from exogenous shocks, but typically there is little movement in these *characteristics* between any two consecutive elections.

Conversely, mobilization is an attempt to push an individual's action off of the equilibrium determined by the aforementioned fixed participation variables; mobilization is an external push off an equilibrium. The "calculus of voting" model Riker and Ordeshook (1968) suggests that voters undertake a simple cost benefit analysis: if the probability weighted benefit (pB) of voting outweighs the structural and cognitive costs (C) of casting a vote, the voter turns out and casts a ballot (Downs, 1957; Riker and Ordeshook, 1968; Gerber, Green and Larimer, 2008). Although Riker's model is useful for analyzing equilibrium comparative statics, the model is not trying to accurately predict any particular voter's decision. Downs (1957) and then Riker and Ordeshook (1968) attach an unmeasured valence (D). The term functions identically even though the scholars have different name for the valence: Downs (1957) believes this term is a voters' preference for the long-run function of democracy, while Riker and Ordeshook (1968). The calculus of voting is represented as

$$pB + D > C \rightarrow VOTE;$$

and thus, attempts to mobilize are an attempt to alter one of the short-term parameters of the calculus.

Mobilization is distinct from exogenous shocks to participation variables in that mobilization requires strategic action taken by an external third party. For example, if an actor wins the lottery, she may be more likely to participate because of the increase in resources that stem from the lottery. In this case winning the lottery would of course be an exogenous shock. If however there is a change in electoral registration laws which cause short-term changes in the participation of voters, then the legal changes could be exogenous socks but it would depend on whether the officials who changed the laws did so with the intent to influence participation rates. For mobilization to occur, the citizen must be intentionally moved off their equilibrium behavior and as such, after a mobilization effort ceases, the likelihood an individual participates should return to the pre-mobilization likelihood unless having participated increases one's likelihood of further participation. In which case, a new equilibrium would be established. Finally, mobilizers are often opinion leaders who increase the salience of a political issue with the intent to alter the likelihood of participating for a particular group (Bartels, 2006, 2008). %[1]

1.3 Explanations for Participation

In the canonical argument for the effects of social economic status on participation, Rosenstone and Hansen find that income, education, unemployment, internal and external political efficacy, Party Identification, church attendance, and mass-elite strategic mobilization form a model that correctly predicts the decision to turn-out to to vote in 75% of cases (Rosenstone and Hansen, 1993, p. 275). Brady, Verba and Schlozman (1995) updates Rosenstone and Hanson's argument to include a broader conception of SES, and concludes that people do not participate for three reasons: they are not able to, do not want to, or are not asked.

Gerber, Green and Larimer (2008) and Gerber and Green (2000) investigate two central unexplored questions from Brady, Verba and Schlozman (1995); *if* voters are asked, does it matter *who* asks them and *how* they are asked? Although these are very broad questions, the authors contend that yes it matters who asks and how citizens are asked because different attempts to influence will activate different social norms. First, they find that the choice of media used to mobilize matters. Calls from a phone bank, mass mailings and house-calls by campaign activists do hot have the same effect on participation (Gerber and Green, 2000). Second, the authors find social pressure applied from a neighbor has a stronger effect in changing the likelihood of an experimental subject than social pressure applied from an unknown researcher and even state "Exposing a person's voting record to his or her neighbors turns out to be an order of magnitude more effective than conventional pieces of partisan or nonpartisan direct mail," (Gerber, Green and Larimer, 2008, pp. 34).

[1] See Also: Brooks and Manza 1997; Manza and Brooks 1999; Leege et al. 2002; Frank 2004; Shor, Bafumi, Park, and Cortina 2008.

1.3.1 Network Explanations

Among others, Fowler and Christakis have linked obesity, smoking, and cooperation to social networks (Christakis and Fowler, 2007, 2008; Fowler and Christakis, 2010). While this research, together with work by Sinclair (2012), Rolfe (2012), Siegel (2013), Klofstad, Sokhey and McClurg (2013) has brought the question to the attention of political scientists in the modern epoch, the thinking is by no means unique to political science, nor this era. Earlier work has shown that medical innovation adoption (Marsden and Podolny, 1990), contraceptive choices (Valente et al., 1997), and adolescent smoking (Bearman, Jones and Udry, 2000; Alexander et al., 2001) can all be correlated with the behavior of peers. Social influence is also an important determinant in the decision to participate in politics. This is a non-controversial claim, but to date little empirical has studied the phenomenon (see e.g. Gerber, Green and Larimer (2008)). This paper demonstrates two distinct, but related, ways that social networks shape political participation: First, political information spreads through interpersonal relationships, and more importantly some agents in social networks are more influential than others because of their structural position within the network. Taken together, this research aims to fill a lacuna in the political mobilization literature noted by Brady, Verba and Schlozman: "*There are three reasons that individuals fail to participate in politics; they are unable to, they are uninterested, or nobody asked*." This paper answers the final possibility; are some people better at asking than others?

To date, reliable inference for questions about social network influences have proven difficult. "Given the number of confounding factors and some of the data requirements, it may be prohibitively difficult to substantiate the role of social networks. . . through survey methods alone" (Valente, 2005). Indeed, due to the difficulty gathering whole-network data, most empirical research has either focused on using either egocentric data from sampled individuals – data that samples individuals at random from a population and queries about social alters, possibly following up with these alters in a snowball sampling fashion – or repurposed, found, sociocentric network data (e.g. Fowler and Christakis, 2010).

An additional problem that arises in the evaluation of social network explanations of behavior is that the nature of social networks, indeed, the very hypothesis being tested, is that behavior contains a component that is socially influenced. In the case of observational data, this social relationship leads to an exceedingly difficult problem to adjudicate: is the similarity in behavior between socially proximate individuals a result of pre-observation selection into groups of like-minded individuals[2], or instad are observed correlations in behavior a consequence of the causal factor under study. Increasingly, scholars are advocating the use of experimental intervention on whole-network populations as a means of separating, and identifying specific causal quantities (e.g. Valente, 2005).

1.4 Social Information

How, specifically does a theory of social information fit into an explanation of political behavior? What are the mechanisms that might lead either an *information* mechanism or an *influence* mechanism to really take hold? This is a remarkably subtle question to adjudicate with data, since developing a test that makes differential predictions between the two among real-world networks challenging.

Consider, for example, the case that a scholar can field only a single experiment. Further, presume that the scholar has an interest in making a causal statement, have measured a large number of social connections, and has the resources to run an intervention. Furthermore, presume that the specific content of the outcome measure and the intervention material are largely immaterial – the researcher is solely interested in designing a test capable of clearly identifying that one mechanism is responsible and not another. Even in this fanciful case, a finding of a causal effect that well connected individuals are more effective, on *some* dimension, would struggle to identify which of the mechanisms was at play.

Indeed, the best designed tests to distinguish between these two effects would run several trials with different costs on system "alters". What does this mean? In one trial, the researcher would imbue system egos with stimulus that is highly contagious; that is, a stimulus with relatively low-cost to spillover. If it is the case that the primary network effect in all circumstances is an *influence* effect, then even in the case of this highly-contagious stimulus that (a) highly connected individuals are the most effective, and (b) that the bulk of the network effect occurs in the highly connected individuals' local social networks. If it is the case that the primary network effect in all circumstances is an *information* effect, then even in the case of a low-contagious stimulus we

[2]DEFINITION NOT FOUND.

would expect highly connected individuals to be the most effective mobilizers, but we would expect that the catchment of uptake would be more broadly represented in the across all the actors of the network.

1.4.1 Information

1.4.2 Influence

For Gerber, Green and Larimer (2008), when a neighbor is *asking*, it increases the pressure for conforming to social norms. When a stranger asks the subject to participate, the same social norms are present, but the pressure is nonexistent. What differentiates the neighbor from the stranger is that the neighbor has a shorter social distance to the subject. In the future, the subject will interact with the neighbor again, and that neighbor may communicate with other neighbors about the subject's willingness or refusal to conform. Whereas the subject will never see the stranger again. Therefore, the neighbor has a greater potential to increase the cost of defecting on and the rewards of conforming to social norms. If this social cost-benefit analysis is part of the decision to participate, then the next logical question is which neighbors are capable of exerting more social influence than others?

Social influence begins with the premise that people all people are seated in a social environment. This environment structures group pressures and the socialization of party identification (Berelson, Lazarsfeld and McPhee, 1954; Campbell et al., 1960). However, distinct from characteristics, without activation social influence does not necessarily hold a direct influence on behavior (Dasgupta and Serageldin, 2001; Lin, 2001; Smith, 2005). Similarly to mobilization, social influence requires intentional activation to alter behavior. However, unlike mobilization efforts, the effect of social influence need not attenuate through repeated use. Indeed, social influence may manifest positive feedback loops that build toward increased effectiveness through repeated use (Milgram, 1974).

Moreover, the results from Gerber, Green and Larimer (2008) results probably underestimate the magnitude of social influence. The use of neighbor-ness is an intuitive proxy for social influence precisely because it differentiates the set of people with whom a person has iterated contact from the set of people with whom a person will never see again. While insightful, neighbor-ness captures a minute portion of the complex relationships that are embedded in the larger social environment. "The social in social cognition research is largely missing" (Kuklinski, Luskin and Bolland, 1991). The perceiver in this literature is a "passive onlooker, who... doesn't *do* anything – doesn't mix it up with the folks he's watching, never tests his judgment in action or inaction. He just watches and judges," (Neisser, 1980, pp. 603-604, emphasis in original). But, we are social beings, and theories of social cognition must, eventually take account of that fact (Krauss, 1981). Accordingly, this study intervenes within empirical communities and the experimental design allows precise operationalization and measurement of influence, and also allows for causal effects to be estimated. Although this study does not capture all of the complex relationships within the social environment, it should represent a significant step forward.

2 A Theory of Social Information

Popkin (1994) builds on the framework of Simon (1955) and Downs (1957) to argue that voters use easily deployed criteria to evaluate candidates. This framework, rationalized by positive political theorists, but whose core insight is drawn from the earlier Columbia studies, argues that "Voters are not always aware of what the government is or could be doing," (p.13 Popkin, 1994, citing Downs (1957)), and yet the government continues to function in a more or less successful way. The core contribution of the Popkin (1994) work is to identify the information market within which voters exist. Recall, the core analogy that Popkin uses is that of the voter as an investor. Rather than a voter purchasing a good as if in an insurance market, instead the voter is trying to make a long-term bet on an mechanism that has some uncertainty about the return. Popkin argues that if voters can obtain two pieces of information that equally allow them to discriminate between predictions about future performance the voter is likely to prefer the less costly signal.

Less developed in Popkin (1994) is the dichotomous feature of mass political decision making. Popkin's theory of a reasoning voter is clearly located in the context of national two-party elections in the US. Thus, a tacit assumption of Popkin's theory is that at every point voters are making a single, dichotomous choice. In the primary phase, voters can choose to support the frontrunner of the party, or not; in the general phase, voters can choose to turnout or not; if they turnout to vote, voters can choose to support the Republican, the Democrat. At each step, for those voters who use a heuristic, that heuristic operates in a decision that is constrained to be between two choice. Many elections in other contexts share this same dichotomous feature. The lack of breadth of this choice, together with the forced election-day timeline, means that it is the role of campaigns to stack diverse factions onto a single, energized and action-ready dimension.

For the voter, the forced timeline and dichotomous choice mean that for those relying on heuristic decision making, very little information can sway a vote. As an example, Popkin identifies Presidential hopeful Gerald Ford's admonition to, "Always shuck your tamales." In the run-up to the 1976 Republican presidential nomination, Ford was touring through San Antonio, TX. When food was passed at the gathering, Ford took a tamale from a plate and took a bite, husk and all. Harmless as it was, the implication of this gaffe for Hispanics, Latinos and Mexican-Americans is that someone who was unfamiliar with how to eat *tamal* must know so little about the preferences of the group of voters that he must have been the wrong candidate.

While existing formal institutions clear the field in many cases, a considerable amount of political decision making occurs that is not simply a dichotomous choice. Indeed, many decisions where the output is black and white are argued in shades of grey by political sophisticates and broadcast to relative novices in dichotomous form (Zaller and Feldman, 1992). These cases might be about decisions for funding levels for a block grant, numbers of troop to deploy to a peacekeeping mission, time to allocate to a community project. Minimally, at least, there are location in the political system where coordination needs to occur, but the set of decisions has not been cleared by some institution that facilitates coordination. For example, consider the voting example in Norton Shores: in the first round of selection, council members needed to complete the difficult task of coordinating on one of a large slate of options with little information. The second round of selection presented a less difficult task because of the field-clearning and signaling in the first round.

2.1 Previous Iterations of Social Information

An important strain of political science research identifies the role of the social environment to shape political outcomes. Berelson, Lazarsfeld and McPhee (1954) anchor the sociological political science tradition and squarely rest their account of political life in Elmira on the *social milieu* – the set of relationships and norms voters hold. “To a large extent, political discussion follows the composition of friendship groups," (Berelson, Lazarsfeld and McPhee, 1954, p.103). This work laid the foundation for later scholars to measure how civic-mindedness affects political participation (Verba and Nie, 1987), assess how discussion networks shape identity formation (Huckfeldt, Johnson and Sprague, 2004; Huckfeldt and Sprague, 1995), and examine how shared norms (Rosenstone and Hansen, 1993; Sinclair, 2012) and conditional cooperation (Rolfe, 2012) explain variation in individuals' levels of political activity.

Although political sciences' sociological tradition identifies a natural place for social context to influence citizens' behaviors, many past empirical works in the paradigm have not measured important parts of this social context or have relied on imprecise approximations or somewhat clunky aggregations. Two influential books published in the last year develop extensive theory about how social context shapes Americans' participation in electoral politics (Sinclair, 2012; Rolfe, 2012). Sinclair (2012) argues that a preference for conflict avoidance leads voters to form preferences that match those of political discussants. Rolfe (2012) develops a different theory that comports well with Sinclair; she argues that voters take conditional action based on the actions of those political actors with whom they interact. This conditional behavior immediately calls to mind the work of Gerry Mackey and behavioral norms. However, because data on the social relationships their theories implicate does not exist, both scholars instad rely on imprecise geographic proxies reasoning that living in the same Congressional district increases the likelihood that two voters share social ties. As a result, in both Sinclair and Rolfe present results that are suggestive of a role for social context, though most of the hypothesized relationships were not borne out by the data at traditional levels (e.g. Sinclair, 2012, p. 68).

Previous theories have argued that social relationships serve as an important conduit for information (Almond and Verba, 1965; Robinson, 1976; Huckfeldt, 1979; Huckfeldt and Sprague, 1987; Kenny, 1992; Rosenstone and Hansen, 1993; Walsh, 2004; McClurg, 2004, 2006; Klofstad, Sokhey and McClurg, 2013) and form the basis of norm based behavior (Wolfinger and Rosenstone, 1980; Spitzer et al., 2007; Sinclair, 2012; Goette, Huffman and Meier, 2006). Rolfe (2012) notes, "Citizens who are asked to vote are more likely to do so"; in 1920s, multi-ethnic Chicago, Gosnell (1927) found that women's social-groups affected women's probability of turning out to vote. In a nationally representative sample, Rosenstone and Hansen (1993) find higher turnout among members of social organizations, and Verba, Schlozman and Brady (1995) argue that "congregational churches" and other civic organizations train their congregants and members in the civic skills necessary to be effectively engaged in politics. McClurg (2004) finds that party contact of political discussant pairs can spill-over from one discussant to another, changing the substance of what these pairs discuss. These theories hold that information largely flows along social channels, and that understanding the contours of these channels might help us to better understand political and behavioral outcomes (Mutz, 2002; Friedkin, 1998).

Much of this literature studies the flow of information between actors, and so focuses on the transmission and diffusion of information explicitly related to electoral politics (Huckfeldt and Sprague, 1987, 1995; Walsh, 2004). The most recent work in the area has focused on the related notions that social networks engender norm-based (Sinclair, 2012), conditional (Rolfe, 2012) responses to stimuli. and has contributed to our understanding of how political opinions and identities are formred in mass politics.

Sinclair and Rolfe (independently) argue that voters' decisions to turnout to vote or participate in local-level political activism are shaped by individual predispositions of voters and how those predispositions change when voters interact with one another. Sinclair argues that broad political outcomes can be well described by a conflict-avoidance mechanism whereby voters choose positions that are in line with the positions of their social connections. In this model, because one selects social alters that are likely to share tastes, and beliefs, and, indeed genetic material (Christakis and Fowler, 2014), taking the position of social alters both does a good job of taking a position near one's own ideal point, but also minimizes the conflict between social alters who take different positions.

Rolfe arrives at a similar conclusion, but argues that political action is the result of conditional-cooperation filtering through a socially-connected world. This view takes a tit-for-tat, game theoretic, view of political action, and argues that there are a small number of actors in society that will take costly political action independent of the actions of others, but that others' actions are conditional on being forced into action by political active members of society. While Rolfe identifies conflict avoidance as one of several mechanisms driving this social effect, she is not interested in specifically engaging the mechanism. Rather, Rolfe's modeling focuses on the effects of broader social-structure rather than interpersonal influences.

The recent acknowledgement of the connections between actors have helped bring about the rapid development of both theoretical and empirical work.

However, with a few notable exceptions, many of these scholars have conceptualized network effects in one of two ways that are, at their core, unnetworked. The first type of simplifying assumption is that of a group-level examination (e.g. Putnam, 1993, 1995; Sinclair, 2012); the second type of simplifying assumptions are those of peer-effects (e.g. Sacerdote, 2001; Fowler and Christakis, 2008; Fafchamps, Vaz and Vicente, 2013). Both the group-effects and peer-effects conceptualizations fail to utilize the core insight of the social network approach, and, as a consequence, neither approach is able to make sufficiently specific predictions about political behavior as a result of the actions of social actors' political behavior. In this section, I present these concerns through the lense of political mobilization and vote choice.

2.2 Social Information and Political Mobilization

First, consider the case of political mobilization. A cohort of scholars have identified group-level evidence for social considerations in the choice to take political action. Fowler (2005) argues that a model with minimal social mimicry and close social connections may lead to group-level externalities for a single individual choosing to vote. Cox, Rosenbluth and Thies (1998) argues that rational mobilizers (campaigns) "should target supporters plugged into wider and more tightly knit social networks, in the hope of producing favorable secondary mobilization," (Cox, Rosenbluth and Thies, 1998, p. 448). They find that rates of turnout correlate with district-level measurements termed "social capital." Gerber, Green and Larimer (2008) find that the most effective mobilization primes social pressure from a group of neighbors; high levels of turnout among elderly Japanese women can be partially attributed to the *social obligation*. Fowler (2005) predicts that individual decisions to turn should have group-level consequences, at least when there is some incentive for individuals to behave like others in the social network, and the social network is sufficiently small that it is possible to observe others' behavior. Sinclair (2012) finds that congressional districts with large numbers of donors—districts argued to have high levels of social capital—also donate more to political campaigns than districts with lower levels of social capital.

Collectively, these scholars surround a position that increased numbers of social connections are associated with an increased likelihood of individual political mobilization. However, two limitations hamstring each of these studies: the authors are neither able to make a strong claim that increasing social connections causes the increase in mobilization; nor can any of this set of authors make a specific prediction about individual level behaviors. For example, consider the most specific prediction possible under the empirical results of Sinclair (2012). A district identified have a large number of coincident voters—that is, one identified as possessing high social capital—is predicted to raise more donation money than one with lower levels of political capital. But these predictions exist only at the group level. Under these results, knowing some specific individual has donated to a campaign is uninformative about whether another is likely to also participate.

The lack of specificity in prediction of mobilization outcomes continues among conceptualizations of social networks as collections of peers. This conceptualization, rather than measuring group-level outcomes, instead, focuses on the specific relationships held between individuals. A leading example of this work, which also demonstrates the limitations of the programme, is that of Bond et al. (2012). In this work, the scholars use an excellent experiment that manipulates the information environment of voters, randomly exposing voters to information that social contacts had taken political action. In contrast to the group-level programme, studies of this class can make limited predictions about behavior at an individual level, however, the specificity of these predictions are limited only to average predicted changes as a result of the myopic evaluation of alters' one degree separated. More specifically to the point, these studies make the strong assumption that the effect of *one particular* social alter are same as the effects of *any possible* social alter. This simplifying assumption does not jibe with the empirical reality. There are both social and non-social reasons to suspect that the actions of one social contact convey a different meaning than the same actions of another social contact.

The limitations faced by Bond et al. (2012) as a result of the dyad-collection view of a social network is also faced by others. Nickerson (2008), in an inventive field experiment, demonstrates the spillover of door-to-door canvassing. However, in this study the design collapses all types of cohabiting adults into a single category. This type-imprecision might be solved with a straightforward model that classes social alters; however, the dyad-collection view essentially pushes off the reductionist, rational actor assumption by one layer. Rather than assuming the core unit of analysis is the individual, the peer-effects conceptualization acknowledges the existence of extra-individual considerations, but it does not actually model the broader sets of connections. In effect, the peer effects framework fails to fully address the problem that it sets out to

Table 1: Social information hypotheses are listed for both mobilization and election tasks. Distinct between the two tasks are the cost of subjects' participation.

Task	**Hypothesis**
Mobilization	1. More social connections $\rightarrow$ more alters mobilized
	2. Costly nature of a attendance $\rightarrow$ socially proximate alters mobilized

solve; although the peer-effects scholars acknowledge the role of system level influences, the models in this paradigm do not meet the requirements of these system-level considerations.

2.3 Hypotheses

From a theory of social information, it is possible to develop a series of discrete, observable tests of individual behavior. In these paragraphs, I present a number of these hypotheses. While the hypotheses presented herein hue toward those that will actually be tested in this work, included as well are other hypotheses that might be tested in future work. In particular, I present a series of hypotheses about distributions of goods that might be tested in future analysis using data already collected as a primary part of this dissertation research. Finally, where it is possible, I highlight the difference between the predictions that results from a theory of social information and more standard theories of rational and behavioral political actors.

Two main empirical tests are brought forward in this work. One test is a test of political mobilization wherein a random sample of individuals are chosen from within the residents of villages and are provided an incentive to recruit as many individuals to take costly, public political action. The second test is of political choice: in these same villages, a separate assignment procedure randomly assigned individuals to stand for representation in a village council. Both tasks require the transmission of information from the seeds assigned to hold it to the others in the village. However, distinct between the two tasks is the particular set of incentives for the task outcome. This difference — coordination in the election experiment and influence in the mobilization experiment — provide the distinct context to evaluate the functioning of social information.

3 Methods

To examine the causal quantities, I build on the modeling framework most closely similar to Bowers, Fredrickson and Panagopoulos (2013). I also utilize the very useful quantities proposed by Baird et al. (2014), and the conceptual contributions of Sinclair (2012).

3.1 Statement of Causal Model

Following Bowers, Fredrickson and Panagopoulos (2013) consider a $n \times n$ adjacency matrix labeled $\mathbf{S}$ that records network relationships between individuals. $\mathbf{S}_{i,j}$ contains a value of 1 in the i, j position of there is a social network connection between individuals i and j and 0 if there is no relationship between i and j. For undirected networks—networks where a connection from one individual to another implies a reciprocal connection—the adjacency matrix $\mathbf{S}$ is symmetric across the diagonal. In this article, I propose to use only undirected networks.

Further, let treatment, when assigned at an individual level, be recorded in a vector $\mathbf{Z}$. Conventionally, before assignment, $\mathbf{Z}$ is a random variable that can obtain one realization of the sample space Ω defined in $\mathbf{Z}$. The sample space, Ω has size k^n, where k is the number of treatment levels and n is the number of subjects possibly treated. After assignment, let $\mathbf{z}$ be the particular realization, and $z_i = 1$ if subject i is assigned to treatment, and $z_i = 0$ if i is assigned to control. Let $\dot{\mathbf{z}}_i$ be the vector of treatment assignments for all individual not i; for compactness in writing, whenever i is explicitly indexed in a quantity, let $\dot{\mathbf{z}}_i \equiv \dot{\mathbf{z}}$. For example, in the statement $Y_{z_i=1,\dot{\mathbf{z}}_i} \equiv Y_{z_i=1,\dot{\mathbf{z}}}$. Finally, use the prime notation ($'$) to define a distinct vector, meaning that $\mathbf{z}'$ indicates that for at least one z_i indexed in $\mathbf{z}'$, $z_i \neq z_i'$. Similarly, $\dot{\mathbf{z}}_i'$ indicates that for some $z_{j \neq i}$ in $\dot{\mathbf{z}}_i'$, $z_j \neq z_j'$.

I use the potential outcomes framework to formalize the notion of a particular treatment causing an individual change (Rubin, 1974). Under this framework, then for every individual i, prior to the assignment of treatment there are two potentially observable outcomes: the individual's outcome if she receives treatment and her outcome if she receives control. In this notational scheme, let $Y_{i,\mathbf{z}}$ represent the potential outcome observed for individual i under treatment regime $\mathbf{z}$, and $Y_{\mathbf{z}}$ represent the potential outcomes to treatment $\mathbf{Z}$ for all subjects.

Standard reasoning—which assumes non-interference—would at this point define an intent to treat effect for individual i as the difference between the potential outcomes when i is assigned to treatment and when i is assigned to control,

$$ITT \equiv E[Y_{i,z_i=1}] - E[Y_{i,z_i=0}].$$

In the presence of full compliance, this ITT is an unbiased estimate of the average treatment effect (ATE), the average of all the causal effects of treatment. However, in the presence of interference – alternatively called spillover in the literature – the schedule of potential outcomes for each individual is defined not only through the assignment status of the referent individual, i, but also the assignment statuses of other individuals. Hence, the necessity of the $\mathbf{z}$ and $\dot{\mathbf{z}}$ notational system. It is useful to note at this point that there is no *ex ante* prediction for which direction interference will bias the two group difference from the true causal estimand. The relationship depends on the particular relationship of the potential outcomes to treatment and control, the nature of spillover, and the possibility of displacement of subjects.

Baird et al. (2014) use this conceptual system to define causal estimand in the presence of interference. Particularly useful are identifying the quantities of the **Intent to Treat** (ITT), **Spillover on the Non-Treated** (SNT), and the **Total Causal Effect** (TCE) in the face of spillover.

Let the **Intent to Treat** (ITT) be the difference between expected potential outcomes of individuals when they are assigned to treatment compared to the case when they are assigned to control, holding all other assignments constant.

$$ITT = E[Y_i(z_i = 1, \dot{\mathbf{z}})] - E[Y_i(z_i = 0, \dot{\mathbf{z}})]$$

This indicates that the ITT effect for individual i is the result of setting her treatment status while holdling the vector of other treatment assignments constant. This might mean changing the assignment of z_i from control to treatment while holding all other assignments at control ($\dot{\mathbf{z}}_i = 0$), or treatment, ($\dot{\mathbf{z}}_i = 1$), or some other vector value ($\mathbf{z} = \{1, 0, 1, 1, 0, ...\}$).

Let the **Spillover on the Non-Treated** (SNT) be the difference between the expected potential outcomes for individuals assigned to control as the result of distinct treatment assignments.

$$SNT = E[Y_i(z_i = 0, \dot{\mathbf{z}})] - E[Y_i(z_i = 0, \dot{\mathbf{z}}_i')]$$

Finally, the **Total Causal Effect** (TCE) is the difference in overall expected potential outcomes, unconditional on identifying the treatment status of a particular individual i. Then, the TCE is

$$TCE = E[Y(\mathbf{z})] - E[Y(\mathbf{z}')]$$

Note, here, that because there is no individual-level indexing in the statement of TCE, any comparison in TCE from a comparison of $\mathbf{z}$ and $\dot{\mathbf{z}}$ must occur at the whole network level.

Two classes of $\dot{\mathbf{z}}_i'$ vectors hold particular interest. The first is the vector corresponding to all individauls are assigned to control, $\dot{\mathbf{z}}_i = 0$, for all z_j. The second is a $\dot{\mathbf{z}}_i'$ vector with known network characteristics. Examination of the second class of vectors highlights that there is a potentially very large set of interesting potential outcomes comparisons to be made. For example, two particularly interesting examples of network characteristics are, (a) all one degree alters, $\mathbf{S}_{i,} = 1$, being assigned to treatment; and (b) the five individuals with the largest number of social connections are assigned to treatment.

3.2 Estimation of Causal Effects

In practice, while estimation of causal quantities could proceed by group-means estimation, in the case that there exists some right-hand-side, pretreatment covariate $\mathbf{X}$ such that the sum of the covariance between the potential outcomes and the outcome variable is greater than the variance of $\mathbf{X}$, then either rescaling outcomes or adjusting by covariate adjustment using $\mathbf{X}$ will yield a more efficient estimate of the causal effect than an estimate that fails to use that covariate. That is, if

$$Cov(Y_i(0), X_i) + Cov(Y_i(1), X_i) > Var(X_i),$$

implying,

$$\frac{Cov(Y_i(0), X_i)}{Var(X_i)} + \frac{Cov(Y_i(1), X_i)}{Var(X_i)} > 1$$

then, including that $\mathbf{X}$ indicator will improve estimation. Colloquially, a pretreatment variable that supplies more predictive signal about the outcome variable than noise encoded in the pre-treatment variable will usefully improve the efficiency of causal estimand. This opens regression analysis as a straightforward estimation technique, not only because it is familiar to all empirical scholars, but also because under the design considerations highlighted above, and some reasonable regression assumptions, regression is a maximum efficiency, unbiased estimator of the treatment effect.

3.3 Assumptions about the form of interference

Common in the analysis of spillovers is to place some bound beyond which spillover is assumed to be zero. Termed *Stratified Interference* by Hudgens and Halloran (2008), the typical assumption is that within some identified cluster of subjects—a media market, a city, a classroom—interference exists, but that between these clusters there exists no interference. These assumptions are often quite easily met in the data.

Important under the *Stratified Interference* assumption is an assumption concerning the form of spillover within spillover clusters. Baird et al. (2014) assume a random effects form of interference. Under this random effects interference, for any treatment saturation—the number of individuals assigned to receive treatment from the fixed size population—the particular **z** vector of individuals assigned to treatment does not effect any individual Y_i potential outcomes. That is, for some saturation π, built of treatment assignment vector **z**,

$$E[Y_i(\pi, \mathbf{z})] = E[Y_i(\pi, \mathbf{z}')]$$

A consequence of this assumption is that within spillover clusters, exactly *who* is treated does not shape outcomes. Perhaps, this assumption is tenable in a spillover cluster that consists of co-habitant, intimate partners: whether one partner or another is treated, the causal effect is unchanged. However, in many other types of relevant spillover networks, this assumption does not meet well with substantive understanding. For example, within a network of staffers for political office, it would seem to be highly relevant if a treatment is assigned to the legislative intern or the director of staff.

4 Data

I obtained consent and followed data collection practices according to the guidelines of the Human Subjects review board at the University of California, San

Diego.

Honduras is a relatively poor nation by comparative Latin American standards, and La Union is comparatively poor within Honduras. Greater than 60% of the residents of La Union are in the lowest quintile of Honduran Income distribution. The residents of La Union have less access to potable water, more residents without access to latrines, lower television penetration, and fewer residents who continue to secondary education. Despite these characteristics—or perhaps as a result of them—the region is home to a vibrant set of social relationships. Both Christian and traditional holidays are celebrated, farmers are organized into collections of cooperatives and aggregators, school-aged children participate in local club sports. Religious life, while for many in the area a core part of social events, is not a monolithic social force. Indeed, in most every village both Catholic and Evangelical adherents participate together in daily life, and a non-trivial number of residents report that they do not participate in any religious services. Therefore, in many ways, at a high-level the residents of this region might be typical of those in other parts of Honduras, Latin America, and rural residents broadly.

The sample for the study was drawn from a census-penetration sampling frame in the thirty-two villages that surround La Union. The set of villages selected to be a part of the sampling frame represent a choice of thirty-two drawn from more than sixty in the area. The choice of village was non-random, as villages that were more readily accessible from the county seat were chosen. These villages range in population size from as few as twenty to as large as six-hundred fifty. For example, on village, *Los Perdomos*, literally "The Perdomos", is a town in which every resident is a part of an the organized Perdomo family group. In contrast with Los Perdomos, San Bartolo is a village that houses a large coffee production center and is home to more than five hundred residents. Across the entire set of villages included in the study, nearly ninety percent of the residents of the village were surveyed.

Data collection in the towns surrounding La Union proceeded in the following four steps. First, the researcher and research assistants met with residents of the village and together canvassed the village to generate a map of the physical locations of all buildings in the village. Although there was little ambiguity about whether a particular building was a residence or a storage unit, it was somewhat more difficult to ascertain whether a residence was presently occupied or of the residents of the building has moved. The more difficult case to identify as a residence was when a building was being constructed and was partially completed. In the case that the research assistants thought the building was potentially inhabited, the research assistants marked the location on a map and scheduled it for a return visit. The reasoning for this enumeration follows that of the US census and other population representative samples: everyone living in a town needs either a permanent or temporary residence, and so by enumerating the buildings, all individuals could be located. In addition, in each village we spoke with residents in the town to query about whether there existed any residents of the town who lived without a permanent structure.

Second, after generating a map of the buildings in each town, the re- search team returned to each structure to speak with the individuals who lived at the structure or on the property. In this stage of the enumeration, we spoke with an adult living on the property and asked the adult to list, by name, the individuals who lived at the building, or on the property. Because there is a concern that subjects with limited recall and large families might not accurately recall each individual living on the property, we employed an age mnemonic to aid respondents; we asked respondents to start with the eldest individual and work toward the youngest individual. For each listed individual, we also queried whether that individual had a partner who lived in the same location. In this manner, by completing the same procedure at every building in town, the research team was able to create a list of every person living in town.

Third, and concurrently with the enumeration of the individuals living at each residence, we snapped a photograph of each individual to be keyed onto that respondentâĂŹs database record. To take this photograph and facilitate the inclusion on the database, we asked each resident to hold a small white board on which we noted the "address" of the building and a record for the individual.[3] The reasoning behind taking this photograph merits some further description.

A key component of this data collection, and one that is described in detail later in this article, is the ability to measure high-fidelity social network relationships in our dataset. In preliminary interviews in the area, the research team identified two features which were problematic for the successful completion of this task. First, because of naming conventions in the area, there exist a relatively small number of given names and family names. Second, and relatedly, individuals living in the area developed a series of "called-names"

[3]DEFINITION NOT FOUND.

to circumvent this problem. The difficulty with this called-name, or nickname convention is that nicknames themselves were neither unique to a single individual, nor was a single nickname always sufficient to identify someone. That is, two individuals in a single town might be called "Chuck", while at the same time, one individual might be known as "Chuck" with one group of friends while being known as "Charlie" to a separate group of friends. This convention is problematic for a researcher with the aim of measuring ties between individuals. By taking a photograph and pairing this with a high-recall search, the research team was able to quickly, and effectively link town-residents with database records.

The fourth, and final step in the primary data collection was returning to administer the survey to each resident enumerated on the population list. This process was a multi-day process, and involved survey enumerators spending considerable time in the town. Because the ambition of this data collection was to elicit social network information from respondents, very high response rates were required. As such, rather than setting a stopping condition beyond which a targeted respondent is considered unreachable, as is commonly used in population representative sampling, in this data collection the enumerators continued to work in a town until they reached ninety percent completion. As a result, those individual who were enumerated in the initial canvassing of the town, but are not included in the primary data collection, are a non-random set of town members.

While this data collection task was not able to gather direct evidence about the individuals who were not surveyed, semi-structured interviews with these missing individuals' social contact suggest that typically these individuals were more likely to be male adults working at locations far from home. Importantly, it was sometimes the case that when we returned to conduct primary data collection, that we would ask when *Potential Respondent A* might return home for the enumerators to speak with them that the residents of the house would respond that the individual was living in the capital city and would not return for several months. In the event that an individual had not stayed at the home she was identified as living in for the month prior, and was not anticipated to return in the coming month, we removed this individual from the population list. If the individual was to be present in the village within a month prior or post the data collection, the individual was left in the population list. This decision is consequential, as individuals who are identified as non-residents from the town are unable to be nominated as a social connection by other residents of the town.

4.1 Assignment

The empirical distribution of social connections is plotted in Figure 1. Consistent with theory (Williamson, 1975; Sidanius and Pratto, 2001) and past measurements (e.g. Apicella et al., 2012), the distribution of social connections in this population is heavily right-skewed. Because estimation of the effect of social connectedness relies on building experimental variation in the connectedness measure, and because a simple random assignment mechanism would assign a relatively large number of poorly-connected mobilizers, care was taken to build a criteria-stratified assignment mechanism. In particular, to ensure adequate variation in the hypothesized causal variable, I employ a stratified random sample with an intuitively derived stratification heuristic—distance from the center of town. In each town, the enumerator assigned approximately 40 percent of the mobilizers from the 20 percent most centrally located homes. Distance from the center of town was expected, *ex ante* to serve as a relevant instrument for social connectedness; this instrument was necessary, given the design, because in some villages assignment occurred prior to the completion of primary data-collection. In this case, individuals who had not been surveyed would be ineligible for assignment into treatment.

In each village, an enumerator was tasked with assigning individuals to a mobilization role. This enumerator was provided a laptop computer loaded with a population list, a randomizer, and a map. The enumerator would select a button in the survey software and a unique identifier was presented. The unique identifier contained information about the randomly selected subjects' house number; this house number was matched against the map of the town made by surveyors. A decision about how many mobilizers to assign was made by the researcher prior to the enumerator starting his task, and was based on the size of town. The smallest towns were assigned either three or four mobilizers, while the largest towns were assigned to receive between four and eight mobilizers.

At this point, the criteria-stratification was evaluated. Specifically, enumerators were instructed to select between one and two mobilizers from the houses located "near" the center of town. In the event that the criteria stratification had not been met, the enumerator would seek out the house and the indvidual inform the subject that he or she had been selected as a mobilizer. The mo-

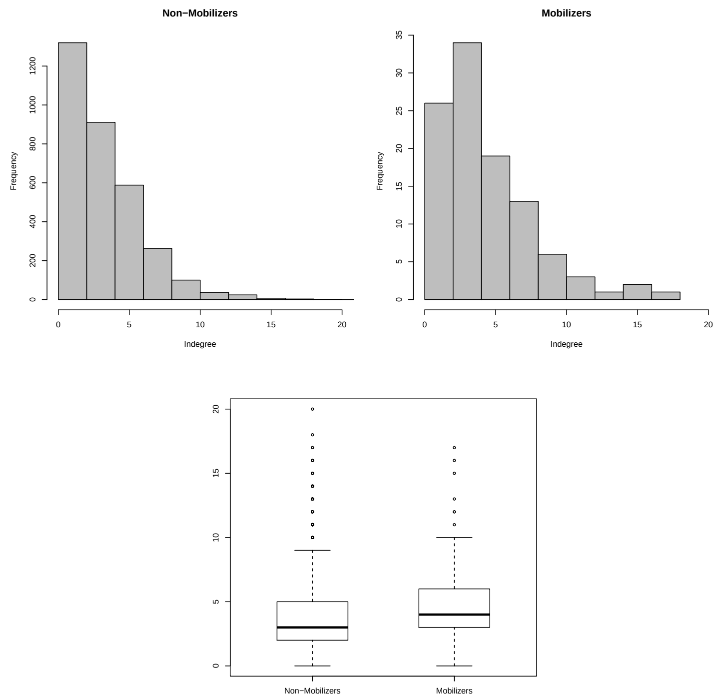

Figure 1: Histogram of the distribution of indegree connections. The distribution is right-skewed. A large number of individuals have a relatively small number of social connections while a small number of individuals have a relatively large number of social connections.

bilization script is included in the appendix to this article. In the event that the criteria stratification had been met – two or more mobilizers had already been selected from near the center of town, then upon drawing a home within the center of town, the enumerator would simply re-select another individual. Likewise, in the event that the individual could not be identified—he or she was at work, visiting another town, or otherwise unable to be located—the enumerator would select another random draw using the town software. Care was taken to attempt assignment after the men and women had returned home from work to mitigate the threat that reached mobilizers were systematically different from unreached mobilizers; the initial recruitment was typically successful, with the enumerator estimating that he contacted approximately eighty-five percent of the primary targets.

An examination of the covariates of those selected to be mobilizers suggests

Table 2: Covariate balance check. Logistic regression predicting assignment as mobilizer.

	Assigned as Mobilizer	Assigned as Mobilizer
	(1)	(2)
Age		0.002
		(0.002)
Married		0.422
		(0.462)
Years Education		-0.001
		(0.005)
Female		-0.223
		(0.300)
Voted		0.539
		(0.342)
Connectedness		0.147^{***}
		(0.037)
P(Assign)		16.790^{***}
		(3.880)
Constant	-3.480^{***}	-5.887^{***}
	(0.115)	(0.534)
N	2,609	2,576
Log Likelihood	-350.600	-226.340
AIC	703.210	468.690

*p $<$ .1; **p $<$ .05; ***p $<$ .01

Logistic regression predicting assignment of individuals to be a mobilizer. Connectedness is the total number of social nominations from other individuals. 'P(Assign)' is the exact probability of assignment according to the targeting procedure.

that the mobilizer corps is largely simliar to the population not assigned to be mobilizers. These results are reported in Table 2. Consistent with the targeting strategy, mobilizers were better connected than non-mobilizers. In addition, mobilizers were more likely to carry covariate profiles predicted by being better connected – they are more likely to be married, and more likely to have taken part in civic activity in the past. Importantly, there are no difference between mobilizers and non-mobilizers on covariates that are not predicted by being well connected – namely age and gender.

Recall, however, that the core causal claim under examination in this paper is not that assigning individuals to be a mobilizer makes those individuals more likely to take political action. Indeed, the targeting used in the assignment mechanism explicitly unbalances the group of individuals assigned as mobilizers from those not assigned. Instead, the core question under examination is whether the assignment of a social alter as a mobilizer increases a subject's likelihood of taking political action. Critical, then, is establishing that, of those not assigned to be mobilizers, there are no pre-treatment differences between individuals assigned to have a social alter serve as a mobilizer and those assigned to have no social alters assigned as mobilizers. To examine this question, I compare mean values, and test for a difference of means on observable characteristics. The mean and sem of this comparison are reported in Table 3.

4.2 Treatment Package

Individuals assigned to be mobilizers were informed of two facts. First, mobilizers were informed that the study team planned to hold a town meeting two-days in the future where the team would thank members of the community, explain our research, and provide information about a microfinance organization that intended to start work in the region in the coming months. Second, mobilizers were instructed that they had been selected at random help the study team bring individuals to this meeting, and that we were going to provide a form of compensation for their help. Specifically, we informed each assigned mobilizer that at the subsequent meeting we would hold a raffle; for each individual who the mobilizer convinced to attend the meeting, we would enter a ticket in that raffle to win 100 Lempira.[4] Our prior experience in the region informed us that this concept of a raffle was well understood, a fact confirmed in our enumerators' conversations with each mobilizer. The specific language read to mobilizers is included in **??**.

Among those who were assigned to receive the treatment, seventy percent attended the meeting. The because better connected individuals were explicitly targeted through the center-of-town heuristic, a two group comparison between those individuals assigned to receive treatment and those assigned not to receive treatment does not hold a causal a design-based causal interpretation. Indeed, as reported in Table 2, those assigned as mobilizers are married at greater rates, have previously been more politically active, and, by design, are hold a greater number of social connections. Among the set of subjects as-

[4]DEFINITION NOT FOUND.

Table 3: Social Balance Table. Comparing covariate values of those assigned to have a social alter at distance *dist=k* assigned as a mobilizer.

	Mean.d1	SE.d1	Mean.d2	SE.d2	Mean.d3	SE.d3	Mean.d4	SE.d4
Age	40.30	1.55	42.03	1.60	38.18	2.07	34.20	4.11
P(Female)	0.48	0.02	0.52	0.01	0.60	0.02	0.65	0.03
P(Vote)	0.70	0.02	0.63	0.01	0.52	0.02	0.38	0.03
P(Married)	0.88	0.01	0.83	0.01	0.73	0.01	0.65	0.03
P(Assign)	0.04	0.001	0.03	0.001	0.03	0.001	0.02	0.001
N Obs.	594		1,357		958		218	

Notes: *Comparison of means and standard errors for distances from assigned mobilizers. 'd1' are subjects who are one degree from a mobilizer, 'd2' are subjects who are two degrees from a mobilizer, and so on. Distance is calculated as shortest-path distnace along all social network ties.*

Table 4: Ordinary Least Squares Model Predicting Mobilizer Turnout at Political Meeting

	Dependent variable:	
	Mobilizer Attended Meeting	
	(1)	(2)
Age	0.001	0.02
	(0.02)	(0.02)
Married	1.47*	−0.42
	(0.77)	(1.27)
Female	0.58	0.94
	(0.50)	(0.69)
Voted	−0.66	−1.29
	(0.61)	(0.87)
Probability of Assignment	9.82	−13.02
	(8.04)	(19.60)
Intercept	−0.66	20.51
	(1.24)	(5,315.00)
Town FE	No	Yes
Observations	99	99
Log Likelihood	−54.92	−34.99
Akaike Inf. Crit.	121.84	127.99
Note:	*p<0.1; **p<0.05; ***p<0.01	

signed as mobilizers, however, little predicts the mobilizers' eventual turnout at the meeting. Table 4 presents these results, and shows that there is little evidence of a systematic relationship between covariates and meeting attendance.

5 Results

5.1 Estimating Within Cluster Controls

To facilitate comparisons between treated and untreated individuals in the pursuit of causal estimands, it is necessary to identify the units designated as control. In a design where spillover of assignment status is possible, it furthermore necessary to identify those individuals who maintain the apples-to-apples comparison of balance on potential outcomes, but who also receive no spillover from individuals assigned to treatment. One promising avenue for this estimation, though it is not pursued in this design, is the creation of individuals to receive "ghost-treatment", or effectively individuals randomized into the treatment group but whom *by design* do not receive treatment (Johnson, Lewis and Nubbemeyer, 2015). Developed in the context of poorly-defined comparison groups in the online-advertising context, the concept may also be usefully applied to estimating causal estimands in the context of interference. In essence, once a model for interference has been identified, it is possible to use the individuals who receive *ghost-treatment* and the social alters of these *ghosts* to form a comparison group for those who do, in fact, receive treatment. The comparison then, would be the difference in outcomes between those indviduals spilled over from a treatment individual and the outcomes from those individuals (not) spillved over from a ghost-treatment individual.

A first alternative approach, and the one pursued in this work, attempts to identify non-spilled over individuals through estimation rather than design. One such method would identify subjects within a treatment cluster who are social isolates – individuals who have no social alters – and therefore are unlikely to have received spillover from treated individuals. These individuals can be used to identify the baseline "hum" or "buzz" about treatment that cannot be attributed to a systematic (e.g. peer-to-peer, or institution-based) spillover mechanism. Once concern with this approach is that those individuals do not hold social alters are unlikely to be broadly representative of the individuals randomized into either treatment or possible interference conditions. In the case that the researcher has *ex ante* expectations that the systematic differences will cause a two-group comparison to underestimate (biased downward, attenuation biased) the true causal effect, this *social isolate* strategy may still be profitably used to estimate a minimum-causal effect, though this strategy clearly comes at the expense of potentially failing to reject the null hypothesis of no-interference when in fact interference does exist.

A second alternative approach would use *ex post* outcome data to estimate the systematic components of spillover, and then use those identified as unlikely to be spilled-over-to as control individuals. The procedure might take the following form:

1. Randomize individuals into treatment;
2. State a model for systematic spillover, e.g. "spillover occurs as a function of social distance";
3. Estimate the degree of spillover, e.g. "evidence is found for spillover at one and two degree relationships, but not three-or-more degree relationships";
4. Compare individuals identified as being unlikely to receive spillover against those who receive treatment (identifying the ITT), or against those likely to receive spillover (SNT).

5.2 Characterization of Mobilizers in Towns

The key, causal variable in this analysis the connectedness of the mobilizer corps in each town. Recall this mobilizer corps is the result of a single assignment vector, $\mathbf{z}$. Figure 2 characterizes the distribution of social connectedness realized by the realized assignment vector in blue, and plots this realization against a simulation of 10,000 alternative draws. The use of the location-based caliper appears to have successfully targeted slightly more well-connected mobilizers than would be expected by chance. Indeed, in the realized mobilizer corps, there are fewer very-poorly connected mobilizers, and slightly more well-connected mobilizers. However, it is important to note that there is good coverage across the entire range of mobilizer connectedness.

Figure 3 further examines the distribution of mobilizer connectedness by breaking down the distribution of connectedness based on the size of the town. Towns are binned into population bins of size one hundred, and the connectedness measure reported is the per-mobilizer total number of connections. In small towns, with population less than one hundred, the median number of connections held by a mobilizer is just fewer than four. In larger towns, with populations greater than one hundred, the median number of connections is between 5.5 and 6.5 ties per mobilizer. A correlation test does not find evidence that mobilizers in larger towns are better connected (Spearman correlation test, $\rho = 0.28$, $p = 0.13$).

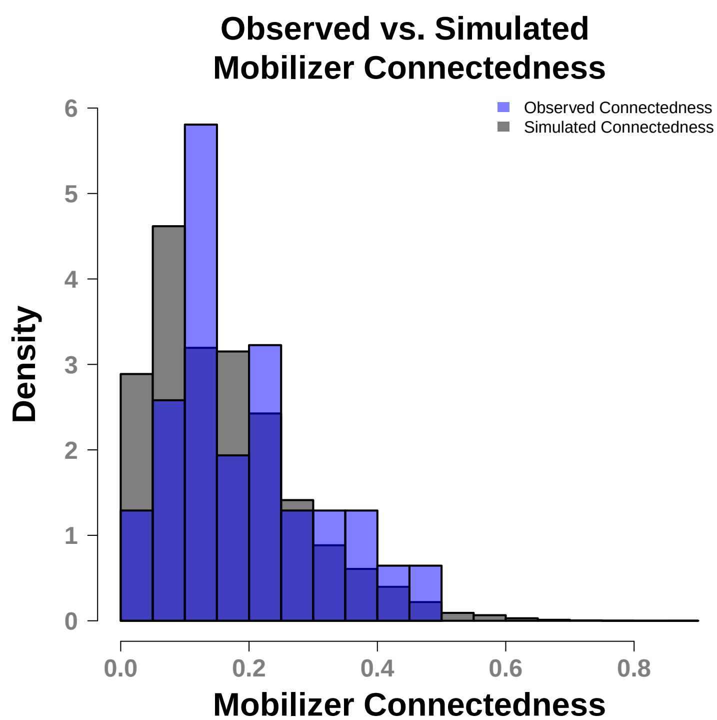

Figure 2: Mobilizer connectedness, measured as each mobilizers' *indegree* scaled by the total population in the village. The blue histogram plots the observed values of mobilizer connectedness, and the grey histogram plots the histogram of connectedness from 10,000 simulated draws of mobilizers within towns. The assignment mechanism performed well at covering the range of possible assignments while targeting assignment of subjects to be mobilizers who were slightly better connected.

5.3 Town Level

I begin the analysis of this mobilization experiment by estimating the **Total Causal Effect** of a well-connected mobilizer corps. To do so, I use the exogenously generated, between-town variation in the connectedness of mobilizers. Recall the TCE causal model presented in subsection 3.1,

$$TCE = E[Y(\mathbf{z})] - E[Y(\mathbf{z}')].$$

As previously noted, the potential outcomes for the TCE are not indexed in i, and estimation comes as a result of between-cluster comparisons of treatment

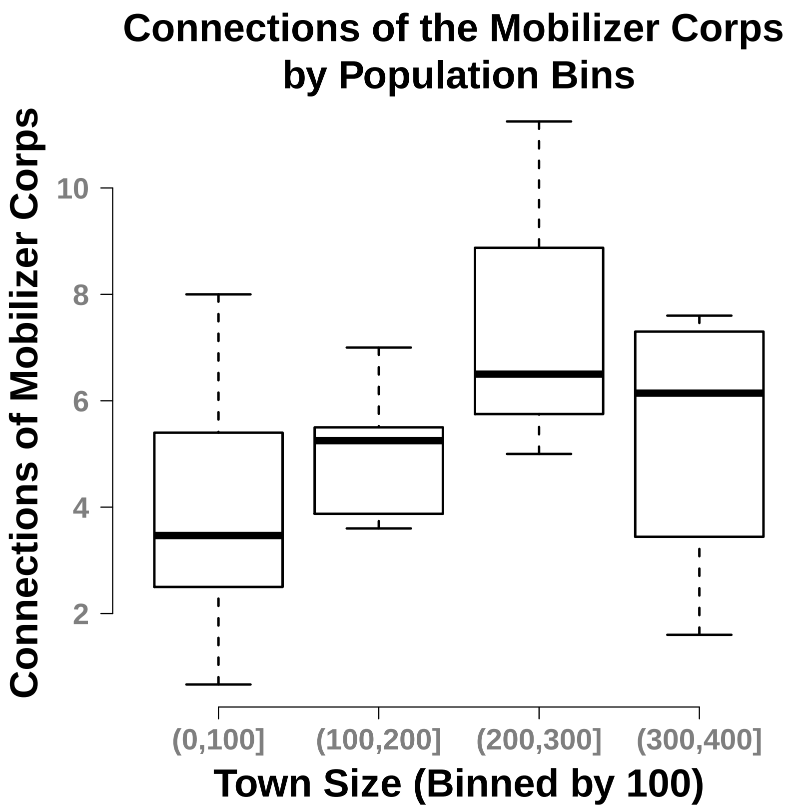

Figure 3: Mobilizer connectedness plotted against mobilizers' town size. The y-axis plots the total number of connections held by the mobilizer corps, divided by the number of mobilizers. Mobilizers living in larger towns have slightly more connections, though there is no difference in between any town with at least 100 residents.

regimes.

A useful reference for $\mathbf{z}'$ would be the vector where no subjects are assigned to treatment: the pure control case. Then, the average causal effect could be estimated as

$$\begin{aligned} E[TCE] &= E[Y(\mathbf{z})] - E[Y(\mathbf{z}(\mathbf{0}))] \\ &= E[Y(\mathbf{z})] - 0, \end{aligned}$$

Although useful to assess the turnout that would arise as a result of residents' curiosities and then estimate the turnout in treatment towns compared against this baseline, in the study fielded for the RSNS, assigning clusters (villages) to pure control could not be completed for logistical reasons. Creating pure control clusters might best be achieved through a stepped-wedge design where some proportion of data receives treatment, an observation is made of all data, and then those units that did not receive treatment in the first intervention stage receive treatment in the second intervention stage. Because the group of enumerators recruited for this were students — either college students in Michigan or highschool students in Honduras — there was insufficient time in the summer break between sessions to allow two intervention periods. Additionally, and even more to the point, the concept of a "pure control" in this particular case is somewhat misguided. Indeed, to create a set of village as pure control would have meant surveying a entire village for baseline demographic characteristics and then holding a political activity without informing any of the village members such an activity was scheduled.

Rather than pursue this pure control strategy, as an alternative inferential strategy I estimate the marginal changes in turnout resulting from variation arising in the connectedness of the mobilizer corps. Provided the assignment mechanism of mobilizers is not correlated with features of the town units, this strategy retains the causal interpretation warranted by the previous estimate of the TCE, but with the benefit of retaining individuals in treatment in all units.

To estimate this relationship, I estimate the following model:

$$Y_v = \alpha + \tau_{TCE} * D_v + Z_v\beta + \epsilon_c,$$

The dependent variable in this model, Y_v, is a count of the number of individuals who attend the village meeting in village v; τ_{TCE} is the marginal increase in attendance at the village meeting as a result of a marginal increase in the exogenously assigned mobilizer connectedness; β is a vector of non-causal estimates of the relationships between predictors, Z_v that potentially improve the efficiency of the causal estimand; α is an estimate of the turnout when a mobilizer corps has zero social connections (and other covariates are zero); and, finally ϵ_v is the residual between the count of village meeting attendance and the fit model.

The results from a linear probability model (OLS regression) for the relationship between mobilizer connectedness and town-level turnout are presented in Table 5. The results from a poisson regression are presented in Table 6.

To test for the possibility that the relationships between mobilizer connectedness operates differently in towns of varying size, Table 6 presents the regression estimates for the total causal effect broken down by town population. Table 6, column one presents the estimates for towns of all size. Here, as reported in the previous paragraph, for every extra social connection in the mobilizer corps approximately 1.5 more individuals attended the village meeting. In the hypothetical case that the mobilizer corps held *no* social connections with others living in the village, this model predicts that 4.5 individuals would turn out to vote. While in general interpreting the intercept of a regression holds little substantive meaning, in this case, the predicted number of attendees at the village meeting when mobilizers hold no social connections is remarkably similar to the 4.17 mobilizers assigned on average in each town.

Table 6, columns two through five examine the possibility that the relationship between mobilizers' social connectedness and village-level attendance at the meeting are different by subsetting the data on village size and running separate regressions. While there is *very* little data in any one of these regressions, in each, the relationship between mobilizer connectedness and turnout is positive, and remarkably stable around the whole-sample estimate. As a formal test for differences in the relationship conditional on town size, Table 6, column six interacts the causal variable with town size.

While there is little data for a model of this complexity to fit the results presented in Table 6 are notably stable, and suggest a positive relationship between mobilizers' connectedness in all size towns. Indeed, fitting a model on the full dataset suggests that a two standard deviation change in mobilizer connectedness[5] causes the predicted attendance at the village meeting to more than double, from about twenty-two residents in attendance to about forty-seven (95% prediction $CI_{low} = [19.23, 23.91]$; 95% prediction $CI_{high} = [43.18 - 50.28]$). These data also suggest that the causal effect of better connected mobilizers is more than twice the magnitude in this sample's larger towns than the smaller towns. To form this comparison, compare the estimated interaction coefficients in Table 6, column (6). In every case, the interaction term for treatment in each of the larger town indicators is larger in magnitude than the baseline effect estimated in the smallest towns. However, it bears

[5]DEFINITION NOT FOUND.

Table 5: Linear Probability Model. Total town political activity (meeting attendance) regressed on mobilizers connections and town size.

	Meeting Attendance		
	(1)	(2)	(3)
Mobilizers Connectedness	1.43^{***}		1.30^{***}
	(0.24)		(0.45)
Town Population		0.09^{*}	0.03
		(0.05)	(0.05)
Constant	6.03	22.90^{***}	5.00
	(6.01)	(7.03)	(6.63)
N	24	24	24
R^2	0.50	0.21	0.51
Adjusted R^2	0.48	0.17	0.47
F Statistic	22.11^{***} (df = 1; 22)	5.69^{**} (df = 1; 22)	11.11^{***} (df = 2; 21)

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Notes: *OLS regression of total number of individuals attending village political meeting on total indegree connectedness of mobilizers. Column (1) reports the regression with only mobilizer indegree; column (2) reports the regression with only town population; column (3) reports the regression with mobilizer connectedness and town population; and, column (4) tests for a differential relationship conditional on town size.*

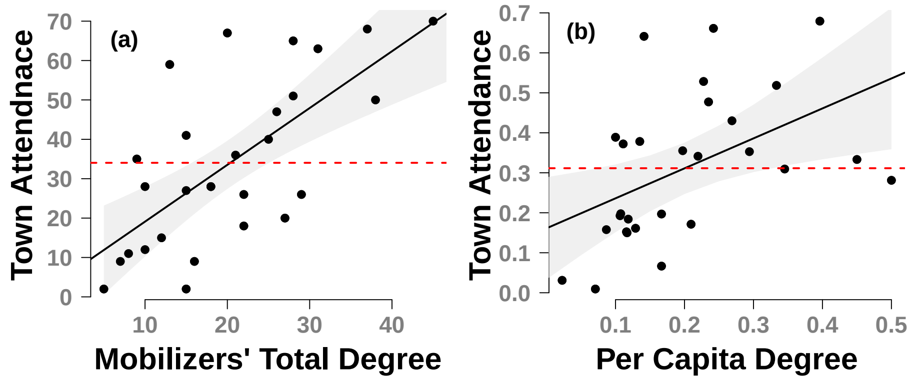

Figure 4: Meeting attendance by mobilizer connectedness. Each plot displays the causal variable, mobilizer connectedness, on the x-axis and the dependent variable, town meeting attendance, on the y-axis. Data realizations are plotted as points. The solid black line is the best fit OLS regression line, and the dashed red line is the best fitting null (intercept only) model. The grey region is the 95% confidence interval for the estimated best fit OLS regression line. Panel (a) presents these results without rescaling by town population; while panel (b) rescales both variables by town population.

mention that these heterogeneous effects are both fundamentally non-causal and also the product of estimating a flexible model on a very small dataset.

Figure 4 plots the bivariate relationship between the dependent variables, meeting attendance at the village level, and the causal variable, mobilizer corps connectedness. Window (a) plots this relationship without rescaling by the size of the town; figure (b) rescales these raw numbers by total town population. Importantly, there is little change in the causal relationship as a result of this rescaling. For every two additional connections held by the mobilizer corps, nearly three more village residents turned out to participate in the village meeting ($\beta = 1.44$, $SE_\beta = 0.30$, $p < 0.001$).

5.3.1 Individual Level

How does the mobilization cue spillover through political actors' social networks? In the previous section I presented evidence about the positive relationship between mobilizers' social network characteristics and turnout at a village-wide political event. The more social connections held by a cohort of

Table 6: Poisson regression. Total town political activity (meeting attendance) regressed on mobilizers connections and town size.

	Meeting Attendance						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mobilizer Connectedness	0.037***	0.025***	0.054***	0.021**	0.048***	0.035***	0.025***
	(0.003)	(0.008)	(0.013)	(0.010)	(0.009)	(0.005)	(0.008)
Town Pop: 100-200						0.296***	−0.338
						(0.090)	(0.354)
Town Pop: 200-300						0.111	0.465
						(0.146)	(0.407)
Town Pop: 300-400						0.095	−0.516
						(0.115)	(0.330)
Connectedness * Town Pop: 100-200							0.029*
							(0.015)
Connectedness * Town Pop: 200-300							−0.004
							(0.013)
Connectedness * Town Pop: 300-400							0.024**
							(0.012)
Intercept	2.731***	2.840***	2.501***	3.305***	2.323***	2.669***	2.840***
	(0.086)	(0.139)	(0.325)	(0.382)	(0.299)	(0.096)	(0.139)
Data Subset	All	0-100	100-200	200-300	300-400	All	All
N	24	12	6	2	4	24	24
Log Likelihood	−149.250	−79.294	−38.779	−5.890	−15.096	−143.260	−139.060
AIC	302.490	162.590	81.558	15.780	34.192	296.520	294.120

*p < .1; **p < .05; ***p < .01

Notes: *Poisson regression of total number of indiviudals attending village political meeting on total indegree connectedness of mobilizers. Column (1) reports the regression with the whole dataset. Columns (2) - (6) examine the possibility that the relationship is conditionally different based on town population. Columns (2) - (5) subset the data by town population, and column (6) interacts a town-size indicator with the causal variable. Having well connected mobilizers in small towns increases turnout; having well connected mobilizers in larger towns has roughtly double the effect.*

mobilizers, the more individuals turned out to participate in the political activity.

In this section, I continue to examine the features leading to the political turnout, paying particular attention to difference in turnout attributable to spillover from treated to non-treated units. A theory of social information predicts that in the case of costly action, individuals will be most effective at spreading information and influence among alters who are socially proximate. In Figure 5 I present evidence that those subjects with a greater number of socially proximate alters assigned as mobilizers are more likely to take political action, and in Table 8 I present complementary, model based evidence for this causal effect. I expand upon each in turn.

Table 7 presents that results of a non-causal regression predicting the number of individuals a mobilizer is responsible for bringing to the political meeting as a function of mobilizer characteristics. Because mobilizers are assigned at random, the number of individuals turned out by a mobilizer holds a limited causal interpretation; indeed, this quantity holds a clear relationship with the spillover on the non-treated causal quantity previously described. However, because the *traits* of mobilizers are not randomly assigned, it is not possible within this framework to identify which trait of the mobilizer is responsible for increasing the mobilizers' effectiveness. Stated another way, it is possible to identify causal effects in this setup, but it is not possible to establish, with certainty, causal mechanisms. Even with this caveat, examining the features of mobilizers that are associated with differential performance in the mobilization task can lend useful circumstantial evidence to the argument for social information.

Table 7 presents the results of a poisson regression of the number of individuals mobilized to take political action on the characteristics of each mobilizer. In this model, observations are individuals assigned as mobilizers, meaning there are 104 observations (99 that hold full covariate data) in the regressions. Consistent with the predictions of a theory of social information, and specifically with a prediction of increased effectiveness for costly action among individuals who are socially proximate, mobilizers who hold a larger set of social connections are more successful at bringing individuals to attend the political activity. Indeed, a two standard deviation change in the number of social connections held by a mobilizer predicts a doubling of the number of alters the mobilizer is predicted to bring to the political activity, from a predicted value of 4.8 to a predicted value if 9.8. These results are similar to those presented in

Figure 4 which aggregated this relationship up to the town-level. Additional features that predict a mobilizer being more able to perform the mobilization task are living in the center of town, ($\beta_{M(4)} = 0.121$, cluster robust $SE = 0.03$) and being older ($\beta_{M(4)} = 0.012$, cluster robust $SE = 0.006$). There is little evidence in this data mobilizers who have previously taken political action, or are married have greater mobilization capacity. One possibility for this null finding is that because connectedness is not randomly assigned, these other effects are operating through other channels (e.g. *Connectedness* and *Probability of Assignment*.

Recall Table 2 and Table 3 which presented estimates of covariate balance between mobilizers and subjects who were assigned to have a mobilizer at varying degrees of social distance, respectively. With the exception of a difference in the probability of being married, subjects assigned to have a mobilizer at a social distance of one[6] were indistinguishable from subjects assigned to have a mobilizer at a distance of two.[7] These individuals one and two degrees from a mobilizer represent more than sixty percent of the observed data; those one, two, or three degrees from a mobilizer represent more than 95% of the observed data.

To estimate both the direct causal effects of being assigned as a mobilizer, as well as the spillover effects from mobilizers to social alters, I estimate a model of the following form:

$$Y_i = \alpha + \tau D_i + \phi E_i + \delta \sum_{j \neq i} (D_{j,S}) + \gamma \sum_{j \neq i} (E_{j,S}) + \beta X_i + \mu T + v_i$$

In this model, the dependent variable, Y_i, is the binary outcome for whether individual i attended the town meeting. α is the baseline turnout rate when all other variable are zero, D_i is the assignment of individual i to serve as a mobilizer, and τ is the difference in the rates of turnout between those who were assigned to be mobilizers and those who were not assigned to be mobilizers; ϕ estimates the difference in the probability of each being assigned as a mobilizer, under the stratified random assignment regime. The term $\sum_{j \neq i}(D_{j,S})$ is the sum of the number of alters, j, assigned as mobiliers under an assumed spillover model, S. In Table 8 the spillover model examined is spillover across social networks, with the flexibility to estimate differential spillover across levels of social distance. Then, δ is the causal effect of having social alters as-

[6]DEFINITION NOT FOUND.
[7]DEFINITION NOT FOUND.

Table 7: Poisson Regression. Mobilizer effectiveness as a function of social connections.

	Number of People Turned Out				
	(1)	(2)	(3)	(4)	(5)
Connectedness	0.119^{***}	0.120^{***}	0.127^{***}	0.121^{**}	0.121^{***}
	(0.025)	(0.026)	(0.026)	(0.048)	(0.032)
P(Assign)		0.843	2.909	8.123	8.123
		(3.294)	(3.305)	(7.270)	(6.245)
Age			0.011^{**}	0.012	0.012^{*}
			(0.006)	(0.009)	(0.006)
Voted			0.296	0.197	0.197
			(0.281)	(0.348)	(0.305)
Female			−0.112	−0.112	−0.112
			(0.238)	(0.365)	(0.260)
Married			0.858	0.600	0.600
			(0.639)	(1.003)	(0.769)
Intercept	1.452^{***}	1.406^{***}	−0.205	−0.784	−0.784
	(0.156)	(0.240)	(0.782)	(1.416)	(1.044)
Town FE	No	No	No	Yes	Yes
N	104	104	99	99	99
Log Likelihood	−531	−531	−475	−402	−402
AIC	1,067	1,069	965	865	865

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Notes: *Poisson regression predicting count of individuals turned out to meeting. DV is number of people at meeting identifying a mobilizer as responsible for their attendance. Connectedness is the total number of social connections of a mobilizer. Models (1-4) use Huber-White heteroskedastic-consistent errors. Model (5) uses Huber-White heteroskedastic-consistent errors clustered at the town level.*

signed to be mobilizers, the SNT. The term $\sum_{j \neq i}(E_{j,S})$ estimates the effect of the expected number of social alters of individual i assigned as mobilizers; this expectation is the product of two features, the social location of subject i, specifically the number of social alters he or she holds, and the stratified randomization scheme. Then, γ serves as a control variable bringing into balance social connections between individuals having social connections according to S_i.[8] X_i is a vector of individual covariates that may improve model fit and contextualize estimated causal relationships. β then is a vector of coefficients associated with non-experimentally assigned individual-level covariates; T is a town-fixed effect to remove unmeasurable town-level difference, and v_i is the vector of residuals from the fit regression. This model is intentionally similar to the model estimated in Bond et al. (2012), which estimated a simliar effect on a network of internet-based, social network users. Importantly, Bond et al. (2012) simulated the properties of this estimator; these MCMC simulations found little bias in estimates, and found Type-I and Type-II error rates in line with specified rates (Bond et al., 2012, SI p. 12).

Across all fit models reported in Table 8 the effect of being assigned as a mobilizer ranges between a four and five times increase in the probability of taking political action if assigned as a mobilizer ($OR = [4.05, 5.47]$, lowest 95% $CI = 2.69$, highest 95% $CI = 10$). Turning to examine the effect of assigning an individuals social alters as a mobilizer, there is clear evidence that assigning both "friends" and "friends of friends" as mobilizers increases the likelihood an individual turns out to take political action. Indeed, Table 8, Model (6) estimates that for each distance one social alter assigned as a mobilizer, an individual is approximately one and a half times more likely to take political action ($OR = 1.56$, 95% $CI = [1.18, 2.07]$). At estimated values then, the marginal effect of assigning an individual as a mobilizer is the same as assigning 2.5 friends as mobilizers. Even more, there is evidence across all models for the effect of assigning an alter a social distance two; the odds ratio for a single alter is 1.38 (95% $CI = [1.21, 1.52]$).

Figure 5 presents a non-model based representation of the meeting attendance rates of individuals. Individuals with neither a distance one or distance two mobilizer turn out at the meeting at a rate of 21%. Individuals with a single distance two mobilizer (but no distance one mobilizers) attend at slightly higher rates, about 35%; a single distance one mobilizer (and no distance two

[8] DEFINITION NOT FOUND.

Table 8: Core causal model. Logistic regression of turnout on experimental and non-experimental factors

	Individual Attends Meeting		
	(1)	(2)	(3)
Mobilizer	1.516***	1.696***	1.705***
	(0.248)	(0.259)	(0.314)
Distance 1 Mobilizer Alters	0.261***	0.466***	0.451***
	(0.086)	(0.122)	(0.151)
Distance 2 Mobilizer Alters		0.323***	0.306**
		(0.096)	(0.125)
Distance 3 Mobilizer Alters		0.158*	0.108
		(0.091)	(0.113)
Distance 4 Mobilizer Alters		0.083	0.077
		(0.081)	(0.097)
Age			−0.001
			(0.002)
Married?			0.338**
			(0.133)
Years Edu			0.002*
			(0.001)
Female?			0.775***
			(0.102)
Voted Last Election?			0.385***
			(0.107)
Connectedness			0.003
			(0.027)
Intercept	−0.334	−0.564*	−1.709***
	(0.225)	(0.299)	(0.371)
Town FE	Yes	Yes	Yes
Propensity Score Mobilizer?	Yes	Yes	Yes
N	3,346	3,346	2,604
Log Likelihood	-1,856	-1,846	-1,369
AIC	3,768	3,760	2,818

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Notes: *Logistic Regression of individual turnout to political activity. The outcome variable is measured turnout. "Distance 1 Mobs" are the number of individuals assigned as a mobilizer at a social distance of 1.The same is true for distance of 2, 3, and 4.* P(Assign as Mob) *is the probability an individual was assigned as a mobilizer. All models include a term for the expected number of mobilizers given the randomization scheme. All models use Huber-White HC errors.*

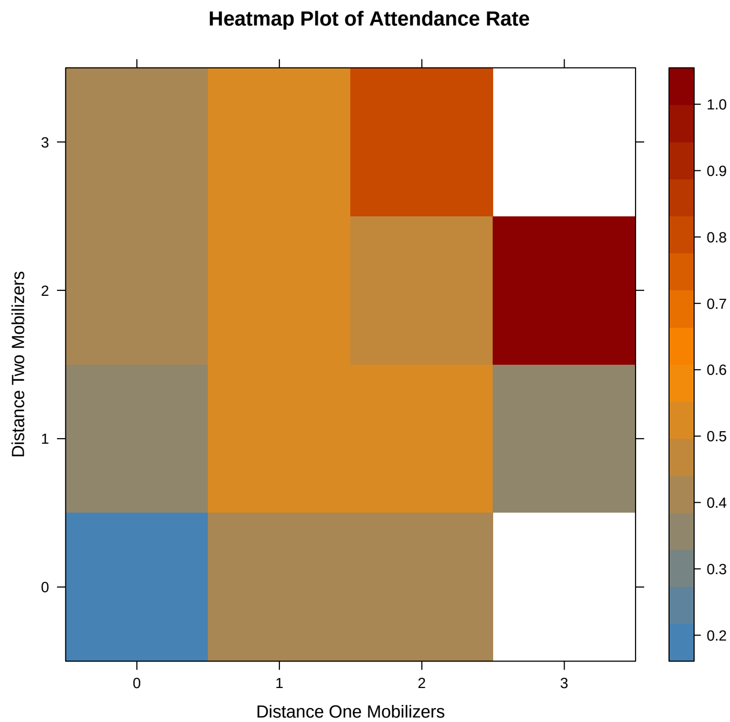

Figure 5: Heatmap plot of attendance at political activity. Blue colors are the lowest likelihood of attending the political activity, red are the highest probability. Probability is a data-based (i.e. not model based) simple probability that an individual with Distance 1 and Distance 2 mobilizers of a particular number attended the political activity.

mobilizers) attend at a rate of 41%. Those with a single one degree mobilizer and a single two degree mobilizer turn out at the meeting at a 50% rate. Figure 5 presents the remainder of the comparisons.

5.4 Limitations

Despite the strength of the random assignment procedure and the inferential strategy there continue exist some limitations in this study. The primary inferential limitation in this study is drawn from the form of political activity being monitored. That is, the chief limitation is the lack of designed pure-control villages. This was a decision evaluated in the design phase, where it was decided that costly data collection in an entire village, while not applying treatment in

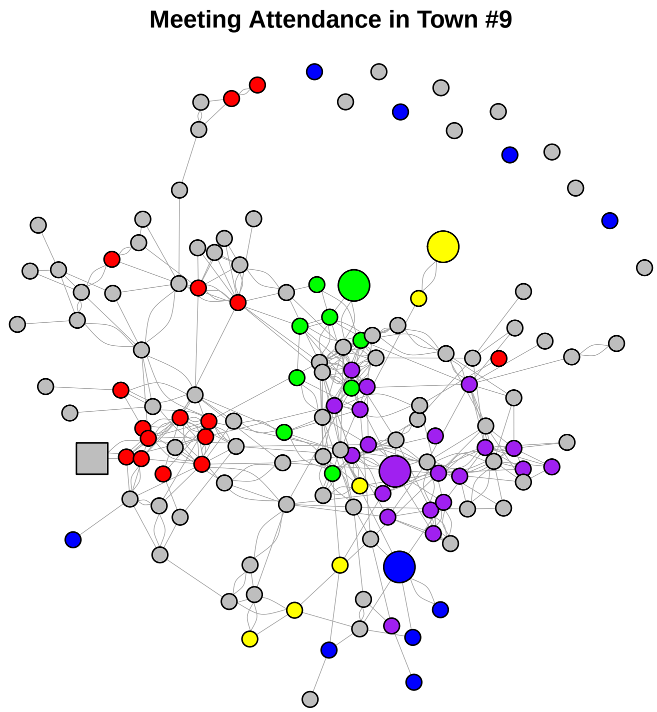

Figure 6: Social network plot for village 9. Nodes in this plot represent individuals; lines represent social connections. The large nodes are those individuals who were assigned as mobilizers, and the small nodes are all other residents. Grey nodes did not attend the meeting. Colored nodes attende the meeting, and are colored to match the color of the mobilizer who brought them to the meeting. Note that the large grey square was assigned as a mobilizer (and would have been colored red), but did not attend the meeting.

that village was infeasible. Despite this limitation, this design is able to estimate rates of change in turnout that have a causal interpretation, even if this estimate is somewhat less straightforward on first impression.

The second limitation concerns the how generally these results might be mapped. Indeed, this study was both undertaken in a unique global setting—rural Honduras—and also explicitly incentivized subjects to spillover behavior to non-treated alters. To be certain, the first of these limitations poses considerable challenge; should one expect that a similar social spillover mechanism might exist among political activists in the United States, or among political violence reduction campaigns in east Africa, or in maternal and child health outcomes in Latin America? These are valid concerns, though the concerns do not uniquely apply to the research in this volume, but instead apply to laboratory, quasi-laboratory, and even field experiments at any time they are run. To the extent that what has been demonstrated in this article—that individuals seek to influence those with whom they are socially proximate, and that individuals are, in fact, subject to this influence—is a general trait of humans as a social species, then these results hold broad import into other domain. As noted in the theory development in an earlier article in this dissertation, the major scoping that is theorized to increase or decrease the importance of social ties are the costs borne by up-taking subjects.

Finally, the design-choice was made to specifically incentivize subjects to spill the behavior over to untreated subjects. Indeed, this choice almost certainly ensures that the magnitude of the spillover observed in this data is greater than the magnitude that might be expected due only to social observation of behaviors; the magnitude of spillover for a voter education campaign, a decision to support a controversial candidate, or an emerging policy-position will very likely be more subtle than the results presented here. Yet, a large part of political action involves creating consensus among a coalition of actors. When subjects have incentives to form as large a group as possible for some campaign, then these results suggest that they may be most successful among those with whom they hold social connections.

6 Conclusion

In this article I have argued that when a researcher holds data about the pathways of spillover it is possible to estimate causal quantities in a field experimental setting. I built upon the causal framework presented in Baird et al. (2014) to measure the Intent to Treat Effect, Spillover on the Non Treated, and Total Causal Effects of a randomly assigned incentive structure that was presented to three thousand residents of a rural, Honduran system of towns. As in Baird et al. (2014), statements of causal models influence the specification of regression-based models for the purposes of drawing inference. However, unlike the previous findings in (Baird et al., 2014), in this intervention where subjects assigned to receive treatment were specifically incentivized to spillover treatment, I find clear and robust effects of spillover on the non-treated. In the finding for spillover to the non-treated, the results in this article are most similar to those of Bond et al. (2012) and Fafchamps, Vaz and Vicente (2013). However, several features distinguish these results from previous findings. First, in this article I find evidence that is robust to the level of inquiry that when political action is seeded with individuals who are highly connected within their social networks that program uptake or spillover are greater than if the political action had been seeded with individuals who are not highly connected within their social networks.

There are several implications for this work, both theoretic and practical. I take each in turn. The findings in this article closely comport with the theory of social information presented in the theory article of this dissertation. The theory of social information predicts that the content of messages between individuals is conditioned by both the person who is sending the message and also by the person who is receiving the message. In particular, in this case when the sender of the message about political activity was better connected, the message was more readily taken up by social alters. Additionally, social alters were more likely to take up a message if the sender of that message was socially proximate.

Even with the acknowledged limitations for generalizability identified in the previous section, these findings hold the possibility for considerable change to be made.

References

Alexander, C, M Piazza, D Mekos and T W Valente. 2001. "Peer networks and adolescent cigarette smoking: An analysis of the national longitudianl study of adolescent health." *Journal of Adolescent Health* 29:22–30.

Almond, Gabriel A and Sidney Verba. 1965. *The Civic Culture*. Boston: Little, Brown and Company.

Apicella, C L, F W Marlowe, J H Fowler and N A Christakis. 2012. "Social networks and cooperation in hunter-gatherers." *Nature* 481(7382):497–501.

Baird, Sarah, Aislinn Bohren, Craig McIntosh and Berk Ozler. 2014. "Designing Experiments to Measure Spillover Effects." *World Bank Policy Research Working Paper* 6824(March).

Bartels, L M. 2008. "The Study of Electoral Behavior.".
URL: *https://www.princeton.edu/ bartels/electoralbehavior.pdf*

Bartels, Larry M. 2006. Priming and Persuasion in Presidential Campaigns. In *Capturing campaign effects*, ed. Henrey E Brady and Richard Johnston. Ann Arbor: University of Michigan Press p. 78.

Bearman, P S, J Jones and J R Udry. 2000. "The National Longitudinal Study of Adolescent Health: Research Design.".
URL: *http://www.cpcp.unc.edu/projects/addhealth/design.html*

Berelson, Bernard R, Paul F Lazarsfeld and William N McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. The University of Chicago Press.

Bond, Robert M, Christopher J Fariss, Jason J Jones, Adam D I Kramer, Cmeron Marlow, Jaime E Settle and James H Fowler. 2012. "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature* 489:295–298.

Bowers, Jake, Mark M. Fredrickson and Costas Panagopoulos. 2013. "Reasoning about interference between units: A general framework." *Political Analysis* 21(1):97–124.

Brady, H E, S Verba and K L Schlozman. 1995. "Beyond SES: A resource model of political participation." *American Political Science Review* 89(2):271–294.

Campbell, Angus, Philip E Converse, Warren E Miller and Donald E Stokes. 1960. *The American Voter*. Survey Research Center, University of Michigan.

Christakis, N. a. and J. H. Fowler. 2014. "Friendship and natural selection." *Proceedings of the National Academy of Sciences* 2014:1–6.
URL: *http://www.pnas.org/cgi/doi/10.1073/pnas.1400825111*

Christakis, Nicholas and James H Fowler. 2007. "The spread of obesity in a large social network over 32 years." *New England Journal of Medicine* 357(4):370–379.

Christakis, Nicholas and James H Fowler. 2008. "The collective dynamics of smoking in a large social network." *New England journal of medicine* 358(21):2249–2258.

Cox, Gary W., Frances M. Rosenbluth and Michael F. Thies. 1998. "Mobilization, Social Networks, and Turnout: Evidence from Japan." *World Politics* 50(3):447–474.

Dasgupta, Partha and Ismail Serageldin. 2001. *Social capital: A Multifaceted Perspective*. World Bank Publications.

Downs, Anthony. 1957. "An economic theory of political action in a democracy." *The Journal of Political Economy* 65(2):135–150.

Fafchamps, Marcel, Ana Vaz and Pedro C Vicente. 2013. "Voting and Peer Effects : Experimental Evidence from Mozambique." (0).

Fowler, J H and N A Christakis. 2010. "Cooperative behavior cascades in human social networks." *Proceedings of the National Academy of Sciences* 107(12):5334.

Fowler, James. 2005. "Turnout in a small world." *Temple University Press, Philadelphia* pp. 269–287.

Fowler, James H and Nicholas a Christakis. 2008. "Estimating Peer Effects on Health in Social Networks." *Journal of Health Economics* 27(5):1400–1405.
URL: *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1162213*

Friedkin, Noah E. 1998. *A structural theory of social influence. Structural analysis in the social sciences*. Cambridge: Cambridge University Press.

Gerber, A S and D P Green. 2000. "The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment." *American Political Science Review* 94(3):653–663.

Gerber, A S, D P Green and C W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* 102(01):33–48.

Goette, L, D Huffman and S Meier. 2006. "The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups." *The American economic review* 96(2):212–216.

Gosnell, Harold F. 1927. *Getting Out the Vote: An Experiment in the Stimulation of Voting*. Chicago: University of Chicago Press.

Huckfeldt, R and J Sprague. 1995. *Citizens, politics, and social communication: Information and influence in an election campaign*. Cambridge University Press New York.

Huckfeldt, R R, P E Johnson and J D Sprague. 2004. *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge University Press.

Huckfeldt, Robert. 1979. "Political Participation and the Neighborhood Social Context." *American Joural of Political Science* 23:579–592.

Huckfeldt, Robert and John Sprague. 1987. "Networks in Political Context: The Social Flow of Political Information." *American Political Science Review* 81:1198–1216.

Hudgens, Michael G and M Elizabeth Halloran. 2008. "Toward Causal Inference With Interference." *Journal of the American Statistical Association* 103(482):832–842.

Johnson, Garrett A, Randall A Lewis and Elmar I Nubbemeyer. 2015. "Ghost Ads: A Revolution in Measuring Ad Effectiveness." pp. 1–36.

Kenny, Christopher B. 1992. "Political participation and effects from the social environment." *American Journal of Political Science* pp. 259–267.

Klofstad, Casey A, Anand Edward Sokhey and Scott D McClurg. 2013. "Disagreeing about disagreement: How conflict in social networks affects political behavior." *American Journal of Political Science* 57(1):120–134.

Krauss, Robert. 1981. Impression Formation, Impression Management, and Non-Verbal Behaviors. In *Social Cognition*, ed. E Tory Higgins, C Peter Herman and Mark P Zanna. Vol. 1 Erlbaum: Hillsdale.

Kuklinski, J H, R C Luskin and J Bolland. 1991. "Where is the schema? Going beyong the" s" word in political psychology." *The American Political Science Review* 85(4):1341–1356.

Lin, Nan. 2001. "Social Capital:: A Theory of Social Structure and Action.".

Marsden, P V and J Podolny. 1990. Dynamic Analysis of Netowrk Diffusion PRocesses. In *Social Networks Through Time*, ed. J Weesie and H Flap. The Netherlands: ISOR: Utrect.

McClurg, Scott D. 2004. "Indirect Mobilization: The Social Consequences of Party Contacts in an Election Campaign." *American Politics Research* 32(4):406–443.

McClurg, Scott D. 2006. "The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation." *American Journal of Political Science* 50(3):737–754.

Milgram, S. 1974. *Obedience to authority: An experimental view*. Taylor & Francis.

Mutz, Diana C. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46(4):838–855.

Neisser, Ulric. 1980. "On" Social Knowing"." *Personality and Social Psychology Bulletin* 6(4):601.

Nickerson, David W. 2008. "Is voting contagious? Evidence from two field experiments." *American Political Science Review* 102(1):49.

Popkin, Samuel. 1994. *The Reasoning Voter*. University of Chicago Press.

Putnam, R D. 1995. "Bowling alone: America's declining social capital." *Journal of democracy* 6:65–78.

Putnam, Rober. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. University of Michigan Press.

Riker, W H and P C Ordeshook. 1968. "A Theory of the Calculus of Voting." *The American Political Science Review* 62(1):25–42.

Robinson, John P. 1976. "Interpersonal Influence in Election Campaigns: Two Step-Flow Hypotheses." *The Public Opinion Quarterly* 40(3):304–319.
URL: *http://www.jstor.org/stable/2748249*

Rolfe, Meredith. 2012. *Voter Turnout*. New York: Cambridge University Press.

Rosenstone, S J and J M Hansen. 1993. *Mobilization, participation, and democracy in America*. New York: Macmillan Puublishing Company.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatment In Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5):688–701.

Sacerdote, Bruce. 2001. "Peer effects with random assignment: results for dartmouth roommates." *The Quarterly Journal of Economics* 116(2):681–704.

Sidanius, Jim and Felicia Pratto. 2001. *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Siegel, David a. 2013. "Social Networks and the Mass Media." *American Political Science Review* 107(04):786–805.
URL: *http://www.journals.cambridge.org/abstract_S0003055413000452*

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *The Quarterly Journal of Economics* 69(1):99–118.

Sinclair, Betsey. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago: The University of Chicago Press.

Sinclair, Betsy, Margaret Mcconnell, Donald P Green, Betsy Sinclair and Donald P Green. 2014. "All use subject to JSTOR Terms and Conditions Detecting Spillover Effects : Design Multilevel Experiments and Analysis of of Chicago." 56(4):1055–1069.

Smith, S S. 2005. ""Don't put my name on it": Social Capital Activation and Job-Finding Assistance among the Black Urban Poor." *ajs* 111(1):1–57.

Spitzer, M, U Fischbacher, B Herrnberger, G Gr\\"on and E Fehr. 2007. “The neural signature of social norm compliance.” *Neuron* 56(1):185–196.

Valente, T W, S Watkins, M N Jato, A Van Der Straten and L Tsitol. 1997. “Social Netowrk Association with Contraceptiove Use among Cameroonian Women in Voluntary Associations.” *Social Science and Medicine* 45:677–687.

Valente, Thomas W. 2005. Notwork Models and Methods for Studying the Diffusion of Innovations. In *Models and Methods in Social Network Analysis*, ed. Perer J Carrington, John Scott and Stanley Wasserman. Cambridge University Press chapter 6, pp. 98–116.

Verba, S and N H Nie. 1987. *Participation in America: Political democracy and social equality*. University of Chicago Press.

Verba, Sidney, Kay Schlozman and Henry Brady. 1995. *Civic Participation*. Cambridge, MA: Harvard University Press.

Walsh, Katherine Cramer. 2004. *Talking about politics: Informal groups and social identity in American life*. University of Chicago Press.

Williamson, Oliver. 1975. *Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization*. Vol. 46 Free Press New York.

Wolfinger, R E and S J Rosenstone. 1980. *Who votes?* Yale Univ Pr.

Zaller, John and Stanley Feldman. 1992. “A simple theory of the survey response: Answering questions versus revealing preferences.” *American Journal of Political Science* 36(3):579–616.