Social Networks, Social Context, and Political Participation: Evidence from Uganda

Nicholas Eubank,*
Guy Grossman,†
Melina Platas,‡
Jonathan Rodden§

April 9, 2018

Click here for most recent version

Abstract

Citizens often mirror the political behavior of family, friends, co-workers and neighbors, but our understanding of the dynamics of peer influence is limited. For example, under what conditions does an individual's decision to vote cascade through a community, producing high voter turnout? Despite substantial theoretical inroads into this question, direct empirical tests remain scarce due to data limitations. Using unique data on complete social networks of 15 villages in rural Uganda, this paper develops theoretical predictions about expected cross-village variation in turnout based on the network structure of each village, and demonstrates that these predictions are tightly linked with actual turnout in low-salience local elections with limited media attention, but not in high-salience presidential elections. These results provide the first direct empirical validation of "social context" theory, and introduce a finding of importance for future empirical network research: the salience of social networks may be conditional on the information environment.

Key words: Turnout, networks, social context, Africa

*Center for the Study of Democratic Institutions, Vanderbilt University, nick@nickeubank.com
†Department of Political Science, University of Pennsylvania and EGAP, ggros@sas.upenn.edu
‡Department of Political Science, New York University Abu Dhabi, mplatas@nyu.edu
§Department of Political Science, Stanford University, jrodden@stanford.edu

Individual characteristics are often inadequate predictors of behavior, including many forms of political participation (Huckfeldt and Sprague, 1995; Kenny, 1992; Leighley, 1990; Mcclurg, 2003). Rather, individuals' actions can often be traced to the behavior of their peers (Ioannides, 2013). While our social *ties* define who is likely to directly influence our behavior, scholars have suggested that it is the *structure* of whole networks that determines how changes in behavior in small groups may propagate to the broader population (Siegel, 2009; Sinclair, 2012; Rolfe, 2012; Fowler, 2005; Larson et al., 2017). Drawing on simulations, scholars have shown that variations in network structure can substantially impact the "contagiousness" of political participation that starts from a few individuals, holding other factors constant. However, real-world network data to test this proposition—that the structure of social networks shapes levels of participation—has been scarce.

In this paper we provide a rare empirical test of the "social context" model of social influence, which posits that an individual's likelihood of political participation is determined by two components: a personal disposition to participate and the level of participation among one's peers (Fowler, 2005; Siegel, 2009; Rolfe, 2012). While some people will always be inclined to participate politically (individuals labeled "unconditional" decision makers by Rolfe (2012) and "rabble-rousers" by Siegel (2009)), others will only participate if they observe sufficient levels of participation among their peers.

The influence of social context has been documented in psychology experiments (Ross and Nisbett, 2011). In some cases, mirroring behavior may be the result of Bayesian updating by rational agents about the desirability of a behavior or strategic social conformity (Goyal, 2012), but research also suggests this dynamic may not be fully conscious (Cialdini, 2015). While this mechanism of influence has been well-documented among individuals, the dynamics of diffusion to larger populations has received less empirical attention, primarily because it is difficult to corroborate the results of theory-driven simulations due to the paucity of real-world network data.

We fill this gap in the literature by estimating the theoretically-predicted level of political participation (TPP) based only on the structure of village networks using the theoretical insights of Siegel (2009) and Rolfe (2012). These values of TPP are then correlated with actual turnout for two types of elections that took place in Uganda in 2016. The highest levels of turnout are indeed found in villages with network structures that, according to theory, should support the highest levels of participation. We also find these same villages have substantially higher attendance at village meetings, and supportive evidence of higher contributions to village projects. Using lab-in-the-field behavioral games, we rule out the possibility that differences in village pro-sociality are driving both differences in network structure and political participation. Consistent with theory, we find that social networks appear to be substantially more important in the lower-salience elections for district chairperson than in the high-salience presidential election.

Data

We collect data from 16 villages in Uganda.[1] Surveying all village households, we construct 16 independent 'whole' networks by using a simple name generator technique (Knoke and Yang, 2008). We collect information on respondents' ties to family, friends, money lenders, and local problem solvers. All networks are unweighted, and undirected (i.e., do not require reciprocity of ties). Results are presented for the union of these four networks (termed the Union network) in the body of the paper, and results for each sub-network are presented in Appendix G.[2] Summary statistics for final empirical networks are available in Appendix B. We are unaware of any political science study that has collected data on as large a number of independent whole, real-world networks.

Data on turnout comes from precinct-level electoral returns and the official voter register as compiled by the Electoral Commission of Uganda. Because polling stations do not correspond precisely to Ugandan villages, we extrapolate turnout using the voter registration data, which provides information on the polling station at which residents of each village are registered, a procedure detailed in Appendix C.

[1] The number of villages was determined by resource constraints. Additional details on our survey and sampling are available in Appendix A.

[2] Throughout the analysis, our sample is restricted to 15 of the 16 villages, because one village is substantially smaller than any other village. Our core findings are generally robust to the inclusion of the 16th village, as shown in Appendix H.

Average turnout for the Presidential election in our data was 60% (compared to 68% nationally); average turnout for the district chairperson was 25% (compared to 31% nationally). Turnout across the elections is correlated at 0.61%, suggesting the elections are distinct but that villages with high turnout tend to have high turnout independent of election type.

Simulating Influence Dynamics on a Network

Social context theory is premised on the assumption that individuals are more likely to participate politically if their peers do so. To understand the dynamics of how this assumption shapes behavior on different types of networks, we use a slightly modified version of the simulation model of Siegel (2009) (which is substantively analogous to Rolfe (2012). Details of our small technical modification to Siegel (2009) can be found in Appendix D).

The starting point of the simulation is that vertices $v \in V$ in a network choose whether or not to participate in a political activity. Initially, the simulation begins with all vertices endowed with some *individual* proclivity to participate. All vertices are assumed to begin in a state of non-participation at $t = 0$, but in the first stage, vertices (or nodes) with very high individual proclivities begin to participate. In each state of the simulation, vertices observe the behavior of *only* their peers and then decide whether to participate. A vertex decision to participate is increasing in the share of her peers that are participating. The simulation then continues this cycle of vertices observing their peers, updating their own behavior, then observing their peers once more until the network converges to a stable configuration in which behavior no longer changes between stages. More specifically, the simulation proceeds as follows:

Stage 1: Model Initialization

- Vertices are randomly assigned an individual propensity to participate $\beta_v \sim Normal(\beta_{mean}, \beta_{sd})$. Once assigned, these values are fixed for the duration of the simulation.
- All vertices begin in a state of non-participation ($participation_{v,0} = 0 \; \forall v \in V$)

Stage 2: Social Influence Simulation

- At each step of the simulation $t \in T$, each vertex $v \in V$ calculates a *local participation* *rate* ($lpr_{v,t}$) which is equal to the share of the people connected to v at time t in the network who also plan to participate. These two factors can then be combined using a very simple decision rule: $participation_{v,t+1} = 1$ if $\beta_v - (1 - lpr_{v,t}) > 0$.

- Calculate overall "Theoretically Predicted Participation" $TPP_t = \frac{\sum_v^V participation_v}{|V|}$.
- Repeat Stage 2 until the value of TPP converges.

Several aspects of this framework are worth noting. First, individuals with high values β (specifically, $\beta > 1$) will participate politically *even if none of their immediate neighbors plan to participate.* Similarly, individuals with very low values of β ($\beta < 0$) will never participation, even if all of their peers are participating. For anyone with a value of $\beta \in (0, 1)$, there is a threshold level of peer participation that will induce those individuals to participate. For example, if $\beta_v = 0.5$, then v will participate if and only if at least 1/2 of her friends participate.

The second aspect of this model is that it is dynamic. We begin in a state of non-participation at time $t = 0$, then in the first period only people with $\beta > 1$ will participate. But as people with $\beta > 1$ announce they are participating, that changes the value of lpr for everyone connected to one of these rabble-rousers, potentially leading them to plan to participate as well. These spillovers may—but also may not, *depending on network structure*—cascade for a period of time before eventually the network stabilizes into an equilibrium level of political participation, which may occur at any level between no one participating and everyone participating.

The focus of our analysis is on the average level of political participation to which this model converges for given values of β_{mean} and β_{sd} – what we term Theoretically Predicted Participation (TPP). TPP calculated by simulating this process of influence repeatedly on the network of each village until the simulation converges, then calculating the average level of participation at these convergent states. In other words, TPP is a network structure property. For a given pair of parameters β_{mean} and β_{sd}, villages whose networks converge to higher levels of *simulated* participation (higher TPP) should also have higher levels of *actual* (observed) voter turnout.

Importantly, the use of simulations is motivated by the fact that basic network properties, like average number of connections or even full degree distribution, do not map cleanly onto the ability of a network to diffuse social influence. This is not to say that different network properties may not be highly correlated. Indeed, in our data, the correlation between average degree and index of simulated equilibrium participation is 0.96 for the Union network. Nevertheless, this will not always be the case, and cannot be taken for granted.

Following Siegel (2009), we focus on parameter values of $\beta_{mean} \in \{0.5, 0.6, 0.7\}$ and $\beta_{sd} \in \{0.25, 0.5\}$ – the range of parameter values that give rise to interesting dynamics. For more on parameter choices, see Appendix E. Average TPP scores across study area villages for different parameter values and network specifications are presented in Table 1.

Table 1: Average Theoretically-Predicted Participation (TPP)

β_{mean}	β_{sd}	Mean, Union	Mean, Family	Mean, Friends	Mean, Lender	Mean, Solver
0.50	0.50	0.42	0.39	0.36	0.35	0.36
0.60	0.50	0.67	0.59	0.54	0.51	0.53
0.60	0.25	0.45	0.38	0.30	0.24	0.23
0.70	0.50	0.83	0.77	0.71	0.67	0.70
0.70	0.25	0.98	0.96	0.89	0.80	0.74

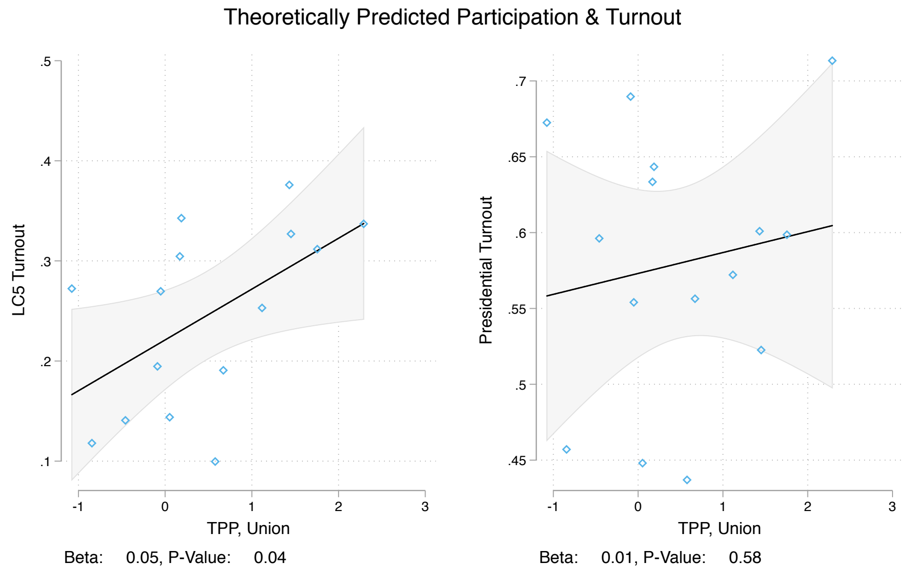
Moreover, the inter-village correlation in TPP scores across this parameter space is quite high, as shown in Tables 3-7 in Appendix F. The overall average correlation across parameters for the Union network is 0.68, and so for ease of exposition (and to reduce the number of regressions we run on the same data), most of the following results will be presented using an index constructed as the first principle component of these statistics for each network type.

Social Context and Turnout

Figure 1 presents the bivariate correlation between the TPP and actual turnout in the Presidential and district (LC5) elections for the Union network. Results for separate network types can be found in Appendix G.

First, in both LC5 specifications, TPP is positively correlated with turnout. Second, despite our relatively small sample size, these results are significant. Moreover, the correlation between these factors appears relatively uniform — results are not being driven by outliers, as is the

Figure 1



TPP in standard deviations of PCA index across parameter values, turnout in shares.

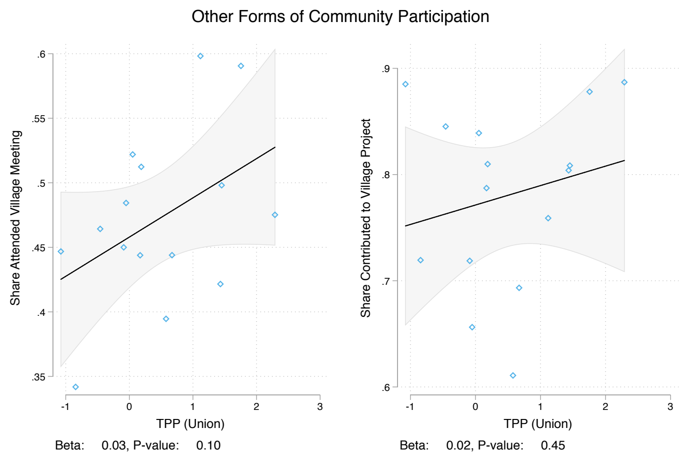
risk in small-N studies — and the relative consistency across specifications provides further evidence of a genuine relationship.

These correlations are also robust to the introduction of different controls and sample restrictions. The results are consistent when controlling for ethno-linguistic fractionalization or the share of each village that has completed primary school, when restricting the sample to the set of villages for which our estimates of turnout are likely most accurate, and are not substantially weakened by the inclusion of the exceptionally small 16th village (Appendix H). Results are also robust to consideration of reciprocated ties (Appendix I).

We interpret the correlation in the LC5 elections as support for network structure theories of social influence. In Uganda, the presidential election is a much higher-salience election that garners substantially greater media attention and entails far more campaign efforts. It seems likely that voters are exposed to information about the likelihood that peers and non-peers will turn out from many non-network channels, such as election rallies. As a result, the specific topology of village networks should matter less for shaping the social contexts that influence voter turnout decisions (though the correlation is still in the predicted direction). In the lower-salience LC5 election, by contrast, a larger share of the information voters receive about anticipated participation may come through their day-to-day interactions and conversations, which are largely dictated by their social networks, increasing the observed correlation between network structure and turnout.

While the focus of much research on social influence and networks has been on voter turnout, social context theory is generally agnostic about the specific form of political participation being fostered. With that in mind, we further examine the relationship between TPP and self-reported information on participation in village governance. In particular, we find that TPP correlates with (a) the share of villagers reporting having attended a village meeting and (b) the share who report having contributed (in time or labor) to a village project. These results are presented in Figure 2. Consistent with theory, we find that this correlation holds up across network specifications and outcomes. In addition, in the case of meeting attendance, the correlation is quite strong and significant despite the relative small sample size.

Figure 2



We briefly address some alternative explanations for the observed correlation between TPP and voter turnout.

The most common concern raised in any network study with observational data is the possibility that a third factor – for example, pro-social norms – drives higher turnout and also leads coincidentally to networks with high TPP. We test for this directly using behavioral games conducted as part of a "lab-in-the-field" component of the survey from which this network data is drawn. In particular, we test whether villages with higher TPP are also villages in which participants are more generous in a divide-the-dollar dictator game. If pro-sociality is driving both network structure and turnout, generosity in the divide-the-dollar game should be positively correlated with TPP. We find that there is a *negative* correlation between pro-sociality among lab subjects and TPP (Appendix J).

In addition, we also find that higher turnout is correlated with high TPP when we look only at the network formed by family connections (Appendix G). As family connections are less likely to have been forged in response to an unobserved third factor, we take this as additional evidence that it is network structure that is driving this relationship.

Are differences between the networks in our study are driven by the presence of central actors (Rojo, Jha and Wibbels, 2014; Cruz, Labonne and Querubin, 2017)? To test this, we drop the five people with the highest eigenvector centrality from each network. If driven by "brokers," the results should be weakened by dropping the most well-connected individuals, but instead they are strengthened, indicating that the results are driven by general network connectedness rather than a small number of facilitators (Appendix K).

A final concern is that networks that give rise to higher TPP may also be networks that better support the efficient diffusion of information, leading to greater turnout. We run diffusion simulations to test which village networks might better support the efficient spread of information. We operationalize "diffusion efficiency" as the average share of each village, reached within a given number of steps of a diffusion simulation (see Appendix L). We correlate simulated values with actual turnout and find no support for the information channel.

Conclusions

This study provides, to the best of our knowledge, the first direct, empirical test of the theoretical predictions of social context theory, which hitherto was only substantiated using simulated (as compared to real-world) data. It finds that at least among the 15 Ugandan villages examined herein, the theoretical predictions of Siegel (2009) and Rolfe (2012) on how network structure may impact political participation are borne out.

In addition to providing the first empirical (i.e., not simulation-based) test of these theories, our exercise offers several important lessons for network scholars. Most importantly, it shows that the importance of networks may be contingent on the environment being studied. When individuals are exposed to extensive messaging by extra-network mediums, the influence of network dynamics may be diminished. Future research might explore how increasing dependence on network-based social media rather than mass media for all forms of political information may influence political behavior.

This paper also hints at resolutions to remaining puzzles in the study of social context and turnout. For instance, large differences in turnout between urban and rural voters may have to do with geographic variation in the structure social networks. Members of racial minority groups may be more likely to participate when living amongst other minorities because of the social networks in which they are embedded, and declines in turnout associated with residential moves might have to do with disruptions in the social networks that sustain political participation.

References

Cialdini, Robert B. 2015. "Basic Social Influence Is Underestimated." *Psychological Inquiry* 16(4):158–161.

Cruz, Cesi, Julien Labonne and Pablo Querubin. 2017. "Village Social Network Structures and Electoral Competition." *Working Paper* .

Fowler, James H. 2005. Turnout in a Small World. In *Social Logic of Politics*. Philadelphia: Temple University Press pp. 269–287.

Goyal, Sanjeev. 2012. *Connections: an introduction to the economics of networks*. Princeton University Press.

Huckfeldt, Robert and John Sprague. 1995. *Citizens, Politics and Social Communication: Information and Influence in an Election Campaign*. Cambridge University Press.

Ioannides, Yannis Menelaos. 2013. *From neighborhoods to nations: The economics of social interactions*. Princeton University Press.

Kenny, Christopher B. 1992. "Political Participation and Effects from the Social Environment." *American Journal of Political Science* 36(1):259.

Knoke, David and Song Yang. 2008. *Social Network Analysis*. Los Angeles: Sage Publications.

Larson, Jennifer M and Janet I Lewis. 2017. "Ethnic Networks." *Forthcoming in American Journal of Political Science* .

Larson, Jennifer M, Jonathan Nagler, Jonathan Ronen and Joshua A Tucker. 2017. "Social Networks and Protest Participation: Evidence from 130 Million Twitter Users ." *Working Paper* .

Leighley, Jan E. 1990. "Social Interaction and Contextual Influences On Political Participation." *American Politics Quarterly* 18(4):459–475.

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

Newman, Mark. N.d. *17 Epidemics on networks*. Networks: An Introduction Oxford University Press.

Rojo, Guadalupe, Subhash Jha and Erik Wibbels. 2014. "Political Networks, Clientelism and Public Goods: Evidence from Slums in Udaipur, India.".
URL: *https://sites.duke.edu/wibbels/files/2014/10/Rojo_Jha_Wibbels.pdf*

Rolfe, Meredith. 2012. *Voter Turnout: A Social Theory of Political Participation*. Cambridge: Cambridge University Press.

Ross, Lee and Richard E Nisbett. 2011. *The person and the situation: Perspectives of social psychology*. Pinter & Martin Publishers.

Siegel, David A. 2009. "Social Networks and Collective Action." *American Journal of Political Science* 53(1):122–138.

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

Online Appendix

A Network Survey

We collect data from a set of Ugandan villages that took part in a multi-year program called Governance, Accountability, Participation, and Performance (GAPP), which was implemented by RTI International and funded by the United States Agency for International Development (USAID) in Arua district, Uganda. 16 villages were selected from a set of over 131 villages that were part of the U-Bridge program, the maximum number that could be enumerated considering budget constraints. Half of the villages had a relatively high level of adoption of the U-Bridge program given village characteristics, and half of which had low levels of adoption. the process of selecting the highest and lowest performers was as follows. We regressed village level adoption of the U-Bridge technology on village-level predictors, and generate a set of predicted values for the dependent variable. We then calculated the difference between the predicted value and the actual value of U-Bridge adoption. Using these residuals, we selected the 8 highest performing and the 8 lowest performing villages with respect to U-Bridge adoption.

We conducted a census in each village in order to collect complete network data, interviewing every available adult who was a resident in the village. This involved a village listing prior to enumeration, the purpose of which was to create a written record of all of the names and household locations for all adults in the villages. To do this the enumerator met with the village leader (in Uganda, called an LC1) and two other village leaders (usually a member of the Village Health Team and an additional village elder or community leader). Together they drew a map of the village identifying every location of a household and major geographical features (e.g., rivers, churches, etc.). Then the group created a list—using their shared knowledge of those households—to identify and name every adult in the village. The enumerator entered this information into a tablet along with other key identifying information such as quadrant (an arbitrary division of the village created by the enumerator—designed to divide the village into four equal portions), age and gender of the potential respondent. These names were then used in the network section of the survey, where respondents were asked about four types of social ties. The exact question wording for the social ties is as follows:

"In each of the following questions, we will ask you to think about people in your community and their relationships to you."

- **Family**: "Think about up to five family members in this village not living in your household with whom you most frequently spend time. For instance, you might visit one another, eat meals together, or attend events together."
- **Friends**: "Think about up to five of your best friends in this village. By friends I mean someone who will help you when you have a problem or who spends much of his or her free time with you. If there are less than five, that is okay too."
- **Lender**: "Think about up to five people in this village that you would ask to borrow a significant amount of money if you had a personal emergency."
- **Problem solver**: "Imagine there is a problem with public services in this village. For example, you might imagine that a teacher has not come to school for several days or that a borehole in your village needs to be repaired. Think about up to five people in this village whom you would be most likely to approach to help solve these kinds of problems."

B Network Summary Statistics

Table 2: Network Summary Statistics

	Union	Friends	Family	Lender	Solver
Average Size	210.3	210.3	210.3	210.3	210.3
Average Num Connections	1,693.9	520.4	810.9	403.3	450.2
Average Degree	15.9	4.9	7.7	3.8	4.2
Min Size	160.0	160.0	160.0	160.0	160.0
Max Size	283.0	283.0	283.0	283.0	283.0

C Turnout Estimation

As previously noted, this analysis assumes that votes at a precinct (polling place) are evenly distributed among voters registered at that precinct from different villages.

To illustrate, assume that at Precinct 1, 200 votes were cast. If 75% of the voters registered at Precinct 1 come from Village A and 25% come from Village B, we assume that Village A contributed 150 votes and Village B contributed 50 votes. Turnout for each village is then calculated as the sum of votes we infer to have been case by its residents at all Precincts. By assuming that at a given precinct all villages have the same turnout, this estimation should bias our analysis in favor of *not* finding differences in turnout across villages.

One constraint of this measures is that the accuracy of these estimates will be related to the correspondence of villages to precincts. If each village sends all residents to its own polling place (i.e. that polling place is only attended by residents of one village), inferences about village voting will be perfect. If, by contrast, all villages send their voters to a single polling place, effectively no information can be learned about how individual villages voted.

This correspondence can be summarized using a *concentration* statistic, where higher values mean the mapping from precincts to villages is more precise. For a village $v \in V$ and a polling place (precinct) $p \in P$, let $voters_v$ be the set of voters who live in v, let $voters_p$ be the set of voters who vote at polling place p, and let $voters_{v,p}$ be the set of voters who reside in v and voted at p. Then:

$$concentration_{v,p} \equiv \frac{\#voters_{v,p}}{\#voters_p} \tag{1}$$

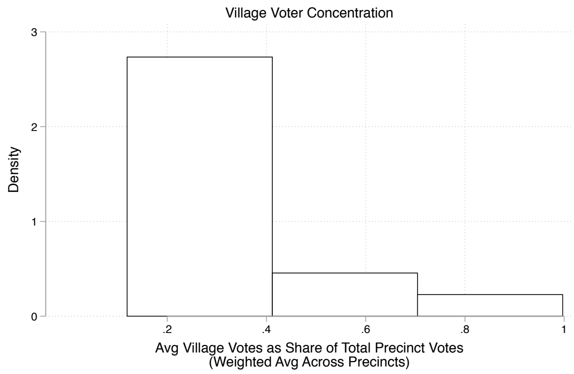
In other words, $concentration_{v,p}$ is the share of voters at a precinct from village v.

Since each village sends people to multiple polling places, however, we must then calculate a weighted average of $concentration_{v,p}$ across polling places where the weight for each polling place is the share of voters from each village going to that polling place. Formally:

$$concentration_v \equiv \sum_{p \in P} concentration_{v,p} * \frac{\#voters_{v,p}}{\#voters_v} \tag{2}$$

The distribution of this statistic are presented in Figure 3:

Figure 3



D TPP Simulation Notes

D.1 Participation Updating

Updating of the $participation_{v,t+1}$ is accomplished by iterating through all vertices in the network in random order and having each vertex update its value of lpr and its participation $participation_{v,t+1}$ *sequentially* rather than simulatenously. This is the one departure from Siegel (2009). When all vertices update lpr simultaneously, it is possible to converge to a "flashing" state in which at time t a portion of the network is planning to vote while another portion is not planning to vote, at time $t+1$, these two groups flip inclinations, and at time $t+2$ they return to their initial state. This is caused by knife-edge simultaneity of updating, which seems unrealistic, since real updating is almost certainly sequential. Thus simulation uses sequential updating.

D.2 Convergence

The simulation is run until no more than 1% percent of the vertices in the network change participation status for at least 20 consecutive periods. Results below are averaged across 2,500 runs for each set of parameter values.

E Parameter Choices

These parameter values are chosen because they effectively cover the range of values that give rise to interesting dynamics. Significantly higher values of β_{mean} tend to result in convergence to full participation, while substantially lower values lead to non-dynamic simulations (those with values of $\beta > 1$ participate, but they are rare and others tend to have very low proclivities to participate, as a result of which almost no vertices flip from non-participation to participation). Similarly, larger values of β_{sd} increase the share of individuals whose behavior is unaffected by the behavior of other so much that the simulations tend not to be dynamic. In these non-dynamic settings, all networks are essentially comparable, as participation ends up being roughly equal to the share of nodes with $\beta_{mean} > 1$, which is the same for all networks in expectation.

Note we exclude one parameter pair from those sets ($\beta_{mean} = 0.5, \beta_{sd} = 0.25$), as it generates almost no unconditional participators, and thus no dynamics.

F Social Context Simulation Validity

Table 3: Correlations across Parameter Values, Union Network

Variables	Mean 0.5, SD 0.5	Mean 0.6, SD 0.5	Mean 0.6, SD 0.25	Mean 0.7, SD 0.5	Mean 0.7, SD 0.25
Mean 0.5, SD 0.5	1.00				
Mean 0.6, SD 0.5	0.86	1.00			
Mean 0.6, SD 0.25	0.56	0.72	1.00		
Mean 0.7, SD 0.5	0.89	0.96	0.67	1.00	
Mean 0.7, SD 0.25	0.32	0.59	0.78	0.45	1.00

Table 4: Correlations across Parameter Values, Family Network

Variables	Mean 0.5, SD 0.5	Mean 0.6, SD 0.5	Mean 0.6, SD 0.25	Mean 0.7, SD 0.5	Mean 0.7, SD 0.25
Mean 0.5, SD 0.5	1.00				
Mean 0.6, SD 0.5	0.95	1.00			
Mean 0.6, SD 0.25	0.80	0.86	1.00		
Mean 0.7, SD 0.5	0.95	0.98	0.88	1.00	
Mean 0.7, SD 0.25	0.83	0.87	0.85	0.91	1.00

Table 5: Correlations across Parameter Values, Friends Network

Variables	Mean 0.5, SD 0.5	Mean 0.6, SD 0.5	Mean 0.6, SD 0.25	Mean 0.7, SD 0.5	Mean 0.7, SD 0.25
Mean 0.5, SD 0.5	1.00				
Mean 0.6, SD 0.5	0.99	1.00			
Mean 0.6, SD 0.25	0.97	0.99	1.00		
Mean 0.7, SD 0.5	0.98	0.99	0.98	1.00	
Mean 0.7, SD 0.25	0.97	0.98	0.97	0.99	1.00

Table 6: Correlations across Parameter Values, Lender Network

Variables	Mean 0.5, SD 0.5	Mean 0.6, SD 0.5	Mean 0.6, SD 0.25	Mean 0.7, SD 0.5	Mean 0.7, SD 0.25
Mean 0.5, SD 0.5	1.00				
Mean 0.6, SD 0.5	0.99	1.00			
Mean 0.6, SD 0.25	0.98	0.98	1.00		
Mean 0.7, SD 0.5	0.99	0.99	0.98	1.00	
Mean 0.7, SD 0.25	0.97	0.97	0.97	0.98	1.00

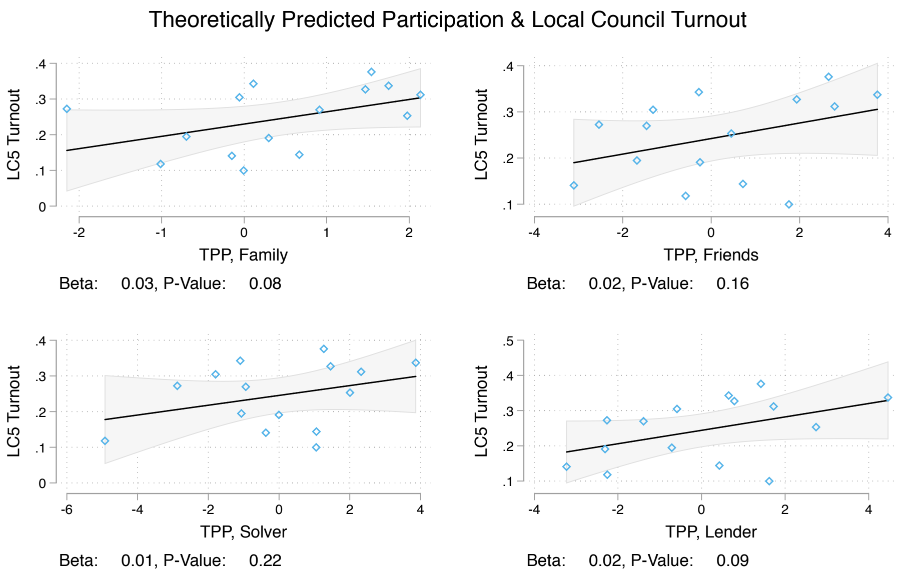
Table 7: Correlations across Parameter Values, Solver Network

Variables	Mean 0.5, SD 0.5	Mean 0.6, SD 0.5	Mean 0.6, SD 0.25	Mean 0.7, SD 0.5	Mean 0.7, SD 0.25
Mean 0.5, SD 0.5	1.00				
Mean 0.6, SD 0.5	0.96	1.00			
Mean 0.6, SD 0.25	0.91	0.95	1.00		
Mean 0.7, SD 0.5	0.96	0.99	0.95	1.00	
Mean 0.7, SD 0.25	0.84	0.89	0.90	0.91	1.00

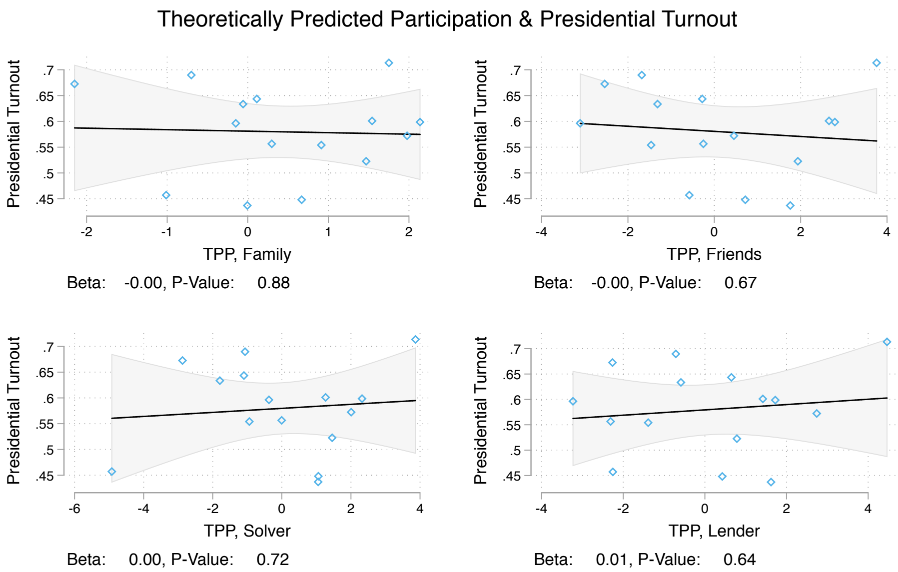
G Turnout and TPP by Network Type

Figure 4 below shows the relationship between turnout and TPP simulated separately on the four different types of networks that contribute to the Union network presented in the main paper.

Figure 4



TPP in standard deviations of PCA index across parameter values, turnout in shares.



TPP in standard deviations of PCA index across parameter values, turnout in shares.

H Robustness Regression Tables

The following tables show robustness of our primary result when controlling for ELF (Column 2), conrolling for education (Column 3), subsetting on the half of villages with the best polling-place-village correspondence (Column 4), and when including the exceptionally small 16th village surveyed (Column 5).

Table 8: Robustness for LC5 Election

	(1) Basic	(2) Basic	(3) ELF	(4) ELF	(5) Educ
TPP (Union)	0.0508	0.0495	0.0520	0.0477	0.0138
	(2.34)	(2.19)	(2.28)	(2.34)	(1.33)
ELF		-0.0599			
		(-0.46)			
Educ			0.0599		
			(0.32)		
Observations	15	15	15	8	16

t statistics in parentheses

Table 9: Robustness for Pres Election

	(1) Basic	(2) Basic	(3) ELF	(4) ELF	(5) Educ
TPP (Union)	0.0138	0.0103	0.0133	0.0113	-0.0260
	(0.57)	(0.43)	(0.52)	(0.41)	(-2.25)
ELF		-0.163			
		(-1.17)			
Educ			-0.0240		
			(-0.11)		
Observations	15	15	15	8	16

t statistics in parentheses

I Considering Non-Reciprocal Ties

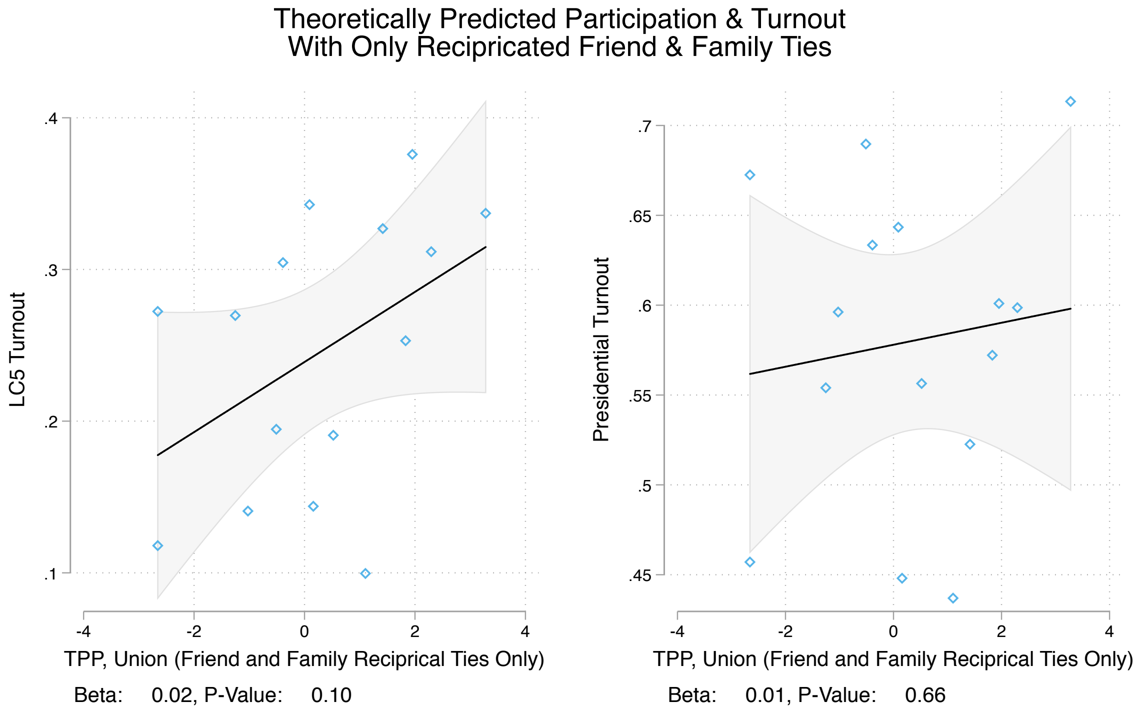
Figure 5 presents results when networks are created using only reciprocated ties to form Friends and Family ties (note the other two inputs into the Union network – the Lender / Solver networks – cannot be restricted in an analogous manner). The figures show results are, if anything, stronger under this restriction.

As shown in Table 10, however, it is not clear that these restrictions are reasonable given the low average degree they generate in the Friend and Family networks. This may be due censoring caused by the limited number of people individuals are allowed to list (5 family members and 5 friends), or failures to recall individuals.

Table 10: Network Summary Statistics: Including Only Reciprocated Friends and Family Ties

	Union	Friends	Family	Lender	Solver
Average Size	210.3	210.3	210.3	210.3	210.3
Average Num Connections	924.1	34.3	156.5	403.3	450.2
Average Degree	8.7	0.3	1.5	3.8	4.2
Min Size	160.0	160.0	160.0	160.0	160.0
Max Size	283.0	283.0	283.0	283.0	283.0

Figure 5



TPP in standard deviations of PCA index across parameter values, turnout in shares.

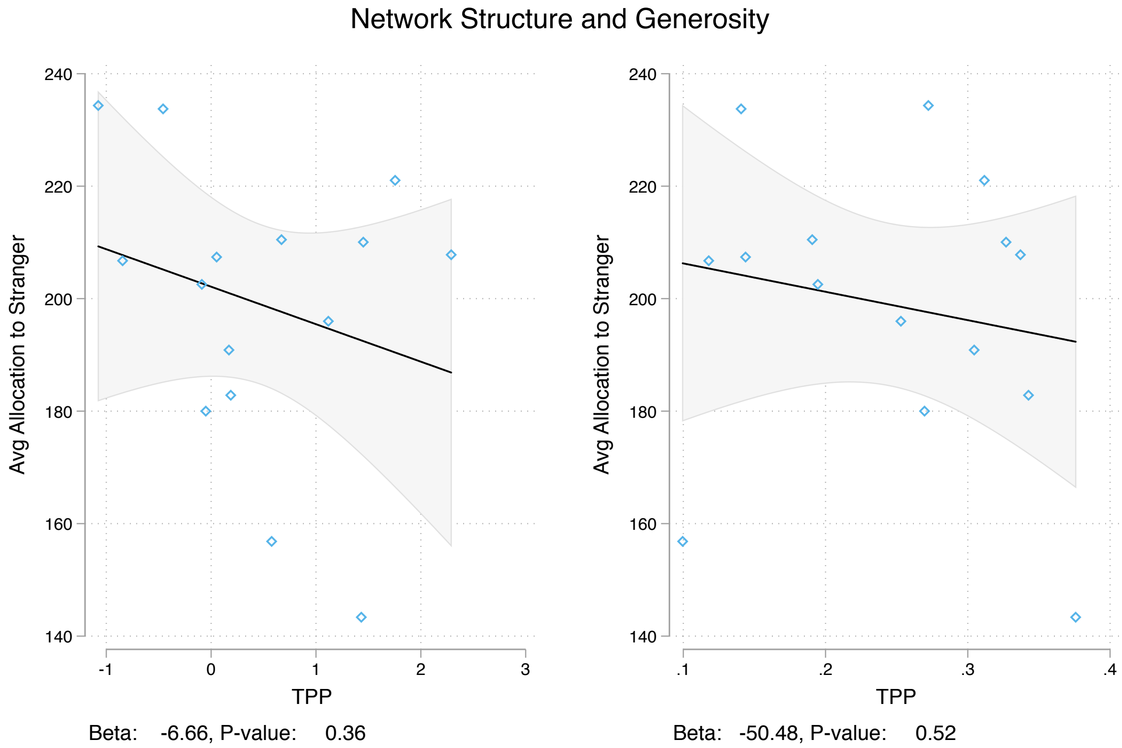
J Divide-The-Dollar Game

The divide-the-dollar game was organized as follows: first, subjects were given ten 100UGX coins. Subjects were then advised that they could split these coins between themselves and a stranger, who they were told will be "someone from Arua whom you do not know personally. We chose the stranger by randomly selecting someone living in Arua district from a long list."

J.1 TPP and Divide-The-Dollar Generosity

Figure 6 the correlates the average allocation to a stranger in each village with TPP. As the figure shows, generosity is, if anything, negatively correlated with TPP.

Figure 6



K Turnout and TPP After Dropping 5 Most Central Nodes

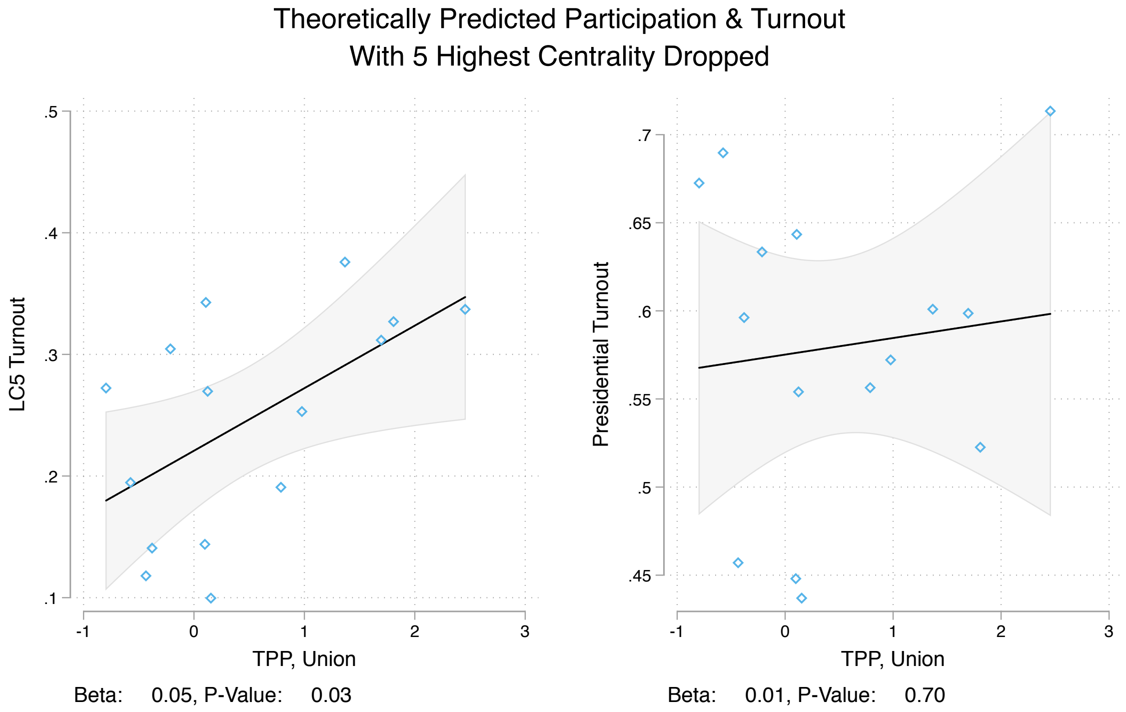
Figure 7 shows the correlation between turnout and TPP when TPP is simulated on village networks after the removal of the five vertices in each village with the highest eigenvector centrality.

L Information Diffusion Simulation

There is no simple statistic – like average degree or average shortest path length – which reliably summarizes the ability of a network to efficiently diffusion information in this type of stochastic manner. (Newman, N.d., p. 19-35) Instead, the ability of a network to support information diffusion is estimated by simulating the diffusion process described above on empirical village networks and the examining the average speed with which information spreads for each village.

More precisely, the simulation proceeds as follows:

Figure 7



TPP in standard deviations of PCA index across parameter values, turnout in shares.

1. At $t = 0$, one vertex v_0 in the network (selected with uniform probability) is endowed with a unique piece of knowledge. It is thus "informed" ($I(v_0) = 1$). All other vertices are assumed to be ignorant of this knowledge ($I(v_j) = 0 \;\; \forall j \in V \backslash 0$).

2. At $t = 1$, information spreads from v_0 to each of the neighbors of v_0, denoted $N(v_0)$ with i.i.d. probability $\frac{p}{|N(v_0)|} \in (0, 1)$.

3. Step 2 is then repeated indefinitely, where at each stage all "informed" vertices spread their knowledge to neighbors with i.i.d. probability p.

The ability of the network to support diffusion can then be specified as the number of people in the network that have become "informed" after s steps of the diffusion model. The larger the number of people "informed" for a given number of steps s, the more efficient a village's network.

Note that the probability of information diffusion from a vertex to her neighbors is normalized by the number of neighbors. This can be thought of as approximating the idea that individuals can only have so many interactions in a given period of time. This normalization more closely approximates the idea that all individuals have the same probability of interacting and sharing information with at least *a* friend in a given period, a dynamic suggested by recent work on information diffusion elsewhere in Uganda (Larson and Lewis, 2017). With that said, results look similar without the normalization.

L.1 Information Diffusion Summary Statistics

Table 11 below shows the correlation in the share of individuals in each village informed at different step thresholds, with different spread probabilities, and with different network specifications. As the table shows, inter-parameter correlations are quite high, and so an index is created for expositional ease by taking the first component of a PCA index for each network specification.

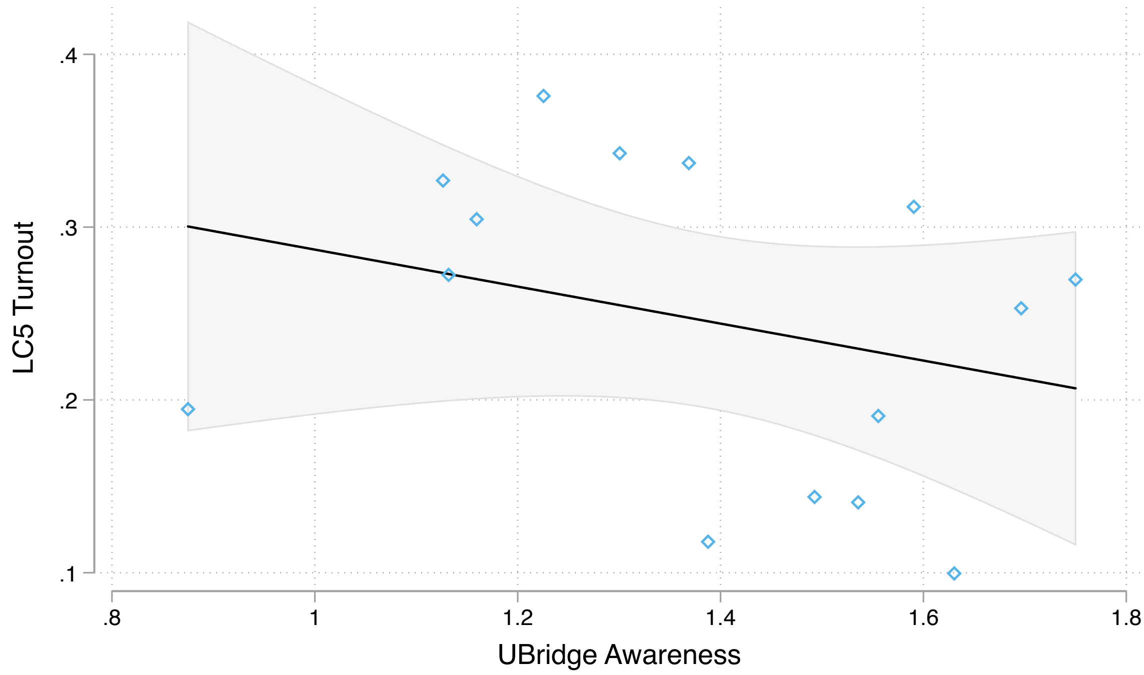
Table 11: Diffusion Correlations across Parameter Values

Variables	p 0.60, 10 steps, Union	p 0.60, 20 steps, Union	p 0.35, 10 steps, Union	p 0.35, 20 steps, Union
p 0.60, 10 steps, Union	1.00			
p 0.60, 20 steps, Union	0.76	1.00		
p 0.35, 10 steps, Union	0.98	0.64	1.00	
p 0.35, 20 steps, Union	0.97	0.86	0.93	1.00

L.2 Information Diffusion and Turnout

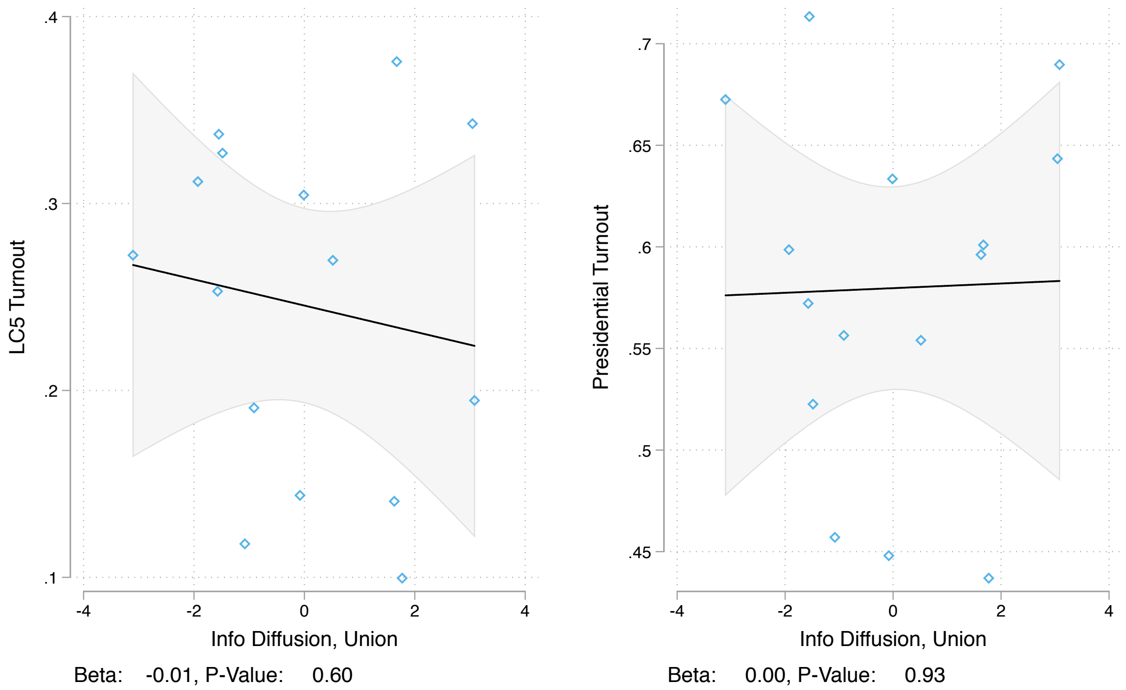
To test the relationship between turnout and information diffusion, we first correlate awareness of the UBridge program with turnout. UBridge was a novel program introduced by USAID to a number of individuals within each village. If the efficiency by which networks diffuse information is driving turnout, then UBridge awareness and turnout should be positively correlated. As shown in Figure 8, they are not. Second, we then correlated our index of simulated information diffusion efficiency with turnout, as shown in Figure 9.

Figure 8: UBridge Awareness & Turnout



Beta: -0.11, P-Value: 0.29
Awareness is share of village residents aware of UBridge, turnout in shares.

Figure 9: Simulated Information Diffusion and Turnout



Info Diffusion in standard deviations of PCA index across parameter values, turnout in shares.