Who Registers? Village Networks, Household Dynamics, and Voter Registration in Rural Uganda

Romain Ferrali * Guy Grossman † Melina R. Platas ‡

Jonathan Rodden §

May 11, 2021

Word Count: 11,467

Abstract

Who registers to vote? Although extensive research has examined the question of who votes, without knowing who is missing from the voter register, our understanding of the determinants of political participation is limited. Studying voter registration in lower income settings is particularly challenging due to data constraints. We link the official voter register with a complete social network census of 16 villages to analyze the correlates of voter registration in rural Uganda, examining the role of individual-level attributes and social ties. We find evidence that social ties are important for explaining registration status within and across households. Village leaders – and through them, household heads – play an important role in explaining the registration status of others in the village, suggesting a diffuse process of social influence. Socioeconomic factors such as income and education do not explain registration in this setting. Together these findings suggest an alternate theory of participation is required.

*Division of Social Science, New York University Abu Dhabi. rferrali@nyu.edu

†Department of Political Science, University of Pennsylvania & EGAP. ggros@upenn.edu

‡Division of Social Science, New York University Abu Dhabi & EGAP. mplatas@nyu.edu

§Department of Political Science, Stanford University. jrodden@stanford.edu

1 Introduction

Who votes? While this is one of the most central questions in political science, the voluminous literature on voter turnout often glosses over an important prior question: *who registers to vote*? Research on voter registration has tended to take place in the context of wealthy democracies where voter registration is not automatic, particularly the United States. While some have suggested that the costs of registration are especially likely to affect those with less education and with lower income, evidence that these individual attributes matter for voter registration is mixed (Hershey, 2009; Leighley and Nagler, 2013). A related strand of work examines the role of social ties— relational rather than individual attributes– in explaining variation in political participation. This work has demonstrated that social ties matter for voter turnout, a relationship observed in both high-income (Rolfe, 2012) and low-income settings (Eubank et al., 2021). However, we know little about whether or how these ties also matter for the prior step of voter registration.

While knowledge gaps exist with respect to voter registration in rich democracies, even less is known in lower income countries. A recent study suggests that individual attributes such as poverty might make constituents more sensitive to registration costs (Harris et al., 2021), but basic descriptive facts about who registers to vote, and importantly, who remains *unregistered*, are elusive. A vexing problem is the lack of reliable data about the population of adults eligible to vote; a challenge that is particularly severe in developing countries where census data are collected infrequently and are of variable quality (Lee and Zhang, 2017).

In this paper, we combine census and official voter register data as well as extensive original survey data from sixteen villages in rural Uganda to examine how and whether individual and relational attributes are associated with registration status. Through an in-person survey with all eligible voters, we collect both individual-level demographic information (e.g., gender, education and age), as well as social network data, which allows us to measure how village residents are related to one another across various domains of interest.

We report two sets of findings. First, there is a limited role for individual attributes

in explaining variation in registration status in this setting. Specifically, factors such as gender, education and income are not significantly correlated with registration status in rural Uganda. Among individual-level factors, only age, being a head of household, and holding a leadership position in the village are clearly associated with registration status (and the latter two are arguably relational attributes). We find that there is a curvilinear relationship between age and registration status. Villagers who are household heads and those who hold leadership positions (for example, elected local leaders, or traditional and religious leaders), are also more likely to be registered to vote.

Second, we find that social ties are important, both within and across households, suggesting a diffuse process of social influence. Within households, irrespective of gender, household heads strongly influence the registration status of others adult household members. Across households, individuals who are closer to the village head and those who occupy a central position within the village network are more likely to be registered. As is common in settings where clientelism is present, in our study area village heads are typically members of political parties and thus mobilize potential voters on their behalf. We find suggestive evidence that registration status is influenced not only through direct ties to village heads, but also via indirect ties to them. Finally, those with ties to non-family lenders are more likely to be registered themselves. These findings show that relational ties are important in explaining registration status, but that ties to multiple types of individuals matter. They further suggest that those individuals who are most likely to be party brokers affect the voter registration process, but this process likely involves both indirect and direct influence.

Our study contributes sorely needed descriptive facts to a nascent literature exploring voter registration status, rather than turnout, in lower income country settings (Harris et al., 2021; Ichino and Schündeln, 2012; Mvukiyehe and Samii, 2017). It also complements a growing body of work examining the role of relational attributes in explaining political behavior. Existing work in this latter domain has examined the role of social networks in voter turnout (Rolfe, 2012; Nickerson, 2008; Eubank et al., 2021), but not for voter

registration, a behavior for which there may be a distinctive process of social influence.

Existing studies of registration in high-income settings on the one hand, and of voter turnout in lower income settings on the other, may not necessarily be informative about who is missing from voter register in lower incomes rural settings that characterize much of the electorate around the world. Indeed, due to the challenges of accessing reliable population data (and social desirability bias in registration status self-reports), we are most likely vastly overestimating voter registration in lower income countries and inadvertently turning a blind eye to those citizens who are absent from electoral politics altogether. Understanding who is not able to participate on election day because they were never registered is important to understanding who is represented through electoral politics and whose voice is heard.

2 Determinants of registration: theory and evidence

Our study builds on two seminal bodies of work that explain political behavior: one focusing on individual attributes and the other on relational attributes and social networks.

Individual attributes and political behavior

Following Downs (1957), an extensive literature examines how individual attributes influence both the motivation and ability to vote (and register to vote), including factors affecting the cost of political participation. A resource model of political participation suggests that those with fewer resources in terms of time, money, and civic skills will be less likely to participate, and that these resources are correlated with socioeconomic status, especially education and income (Brady et al., 1995). Other work examines additional individual attributes that may correlate with resources, particularly membership within marginalized groups, such as gender and religious or ethnic minorities (Grossman et al., 2014). On the basis of this voluminous body of work, there are reasons to believe these individual attributes matter in explaining voter registration status.

However, the role of these factors in explaining political participation and particularly voter turnout vary within and across countries, and therefore may operate differently in lower income (Kasara and Suryanarayan, 2015) and less democratic contexts (Croke et al., 2016). Further, relatively little research has examined whether these factors matter for registration specifically, which typically takes place prior to election day and for which there are no immediate benefits.

Historically, research on the relationship between individual attributes and political behavior has focused disproportionally on the United States, and in this context has repeatedly revealed a strong positive association between political participation and both education and income. However, recent comparative work has demonstrated that the role and relative importance of sociodemographic factors vary across countries, and thus that the relationship between individual attributes and political participation is conditioned by contextual and institutional factors. For example, Gallego (2014) found that while income and education are highly correlated with turnout in the United States, this is not the case in other wealthy democracies like Spain, Denmark, and South Korea. Similarly, Nevitte et al. (2009) found that age is correlated with voting in three-quarters of the 33 elections examined, education in about half, and income in only about one third. The sample was comprised of mostly middle- and high-income countries for which survey data were available.

Research in low- and lower-middle income countries further suggests that factors such as education and income either matter less for political participation or operate differently in poorer, as well as in less democratic settings. Isaksson et al. (2014) found that across 20 African countries, those who are poorer are, if anything, *more* likely to vote, cutting against the resource model of political participation. Similar patterns have been found in India and Latin America (Krishna, 2008).

Attributes such as education may also operate differently in less democratic contexts as compared to consolidated democracies. While education is often found to be positively associated with voting in democratic contexts, more educated individuals may deliberately avoid

turning out to vote in autocracies as a form of dissent Croke et al. (2016). In authoritarian settings, the less educated may also be more easily influenced by vote buying (Blaydes, 2010).

The extent to which findings from the voluminous literature on voter turnout apply to the prior step of voter registration is both poorly understood and contested. Here again, much of the research focuses on the case of the United States, where there exists large variation in laws governing the registration process across states. Where it is not automatic, registration, like turning out to vote, involves costs, which are weighed against the benefits of political participation. It is conceivable, therefore, that those with fewer resources will be less likely to register, and more likely to respond to policy changes that reduce the cost of registration. However, it is challenging to gather reliable individual-level data on unregistered voters, and there is only mixed evidence that reducing the cost of registration disproportionately benefits the poorer or less educated.

Grumbach and Hill (2021) demonstrate that the introduction of election day registration (EDR), which is designed to simplify the registration process, leads to higher registration rates and turnout, especially among young Americans, but Leighley and Nagler (2013) found that EDR expansion did not increase the relative share of low-education individuals among voters. Likewise, (Nagler, 1991) found that restrictive registration laws were associated with lower turnout, but this effect was similar for high- and low-education constituents. While such studies of registration are rare, their findings are consistent with the much larger voting behavior literature in American politics, which generally finds mixed evidence of the effect on turnout of innovations that reduce voting cost. For example, universal vote-by-mail increased overall turnout (Thompson et al., 2020), and this increase was found to be especially pronounced among low-propensity voters like less educated, minority, and young voters (Bonica et al., 2020). However, some earlier studies found that convenience voting is, in general, less helpful for low-propensity voters than for those who were already likely to vote (Berinsky, 2005).

Turning to the context of lower income countries, Harris et al. (2021) found that an intervention reducing the cost of registration in Kenya—conducting mobile registration rather than requiring citizens to registration in-person—increased registration, particularly in poorer communities, though this analysis was conducted at the aggregate rather than individual level. It could be that potential voters are more sensitive to registration costs in poorer areas. However, it could also be that those in lower income settings are more sensitive to clientelistic strategies, such as turnout buying (Nichter, 2008). Indeed, how and whether brokers mobilize voters for voter registration, as compared to better-documented voter turnout, is unclear, a point to which we return below.

Together this body of work underlines a series of reasons to suspect that individual (sociodemographic) attributes could matter for voter registration. This may not necessarily be the case, both because of the distinct way these attributes operate in poorer and less democratic settings, and because voting theories may not apply to the act of registration, for which there may not be immediate benefits. In sum, the predictions about how individual attributes such as education and income matter for registration are cross-cutting.

Social ties and political behavior

Rather than individual attributes, the second strand of work we build on examines the role of social ties and social influence in explaining political behavior (Eubank et al., 2021). Social networks could matter in three ways that are relevant for voter registration: information, influence and mobilization. As for information, citizens must learn to follow a set of rules about how and when to register. Connections to (politically) knowledgeable individuals can reduce information search costs. As for influence, according to social context theory, citizens are more likely to participate politically when they observe sufficient levels of participation among their peers (Rolfe, 2012; Siegel, 2009). Finally, participation might increase when one is connected to mobilizers with a strong interest in making sure that many *others* vote (Nichter, 2008). This literature focuses on within-household (Nickerson,

2008) or within-family dynamics (Cruz et al., 2017), as well as on ties to various types of brokers (Stokes et al., 2013).

Brokers and other party operatives are thought to play an especially important role in mobilizing potential voters. Party brokers can facilitate political participation in numerous ways, typically by reducing the costs and/or by increasing the benefits of participation for individuals (Stokes, 2005) or communities (Rueda, 2015). For example, they can provide transportation to polling or registration places, they can distribute information about candidates and parties, and they can distribute resources in exchange for turning out to vote, with the goal of affecting voting behavior (Brierley and Kramon, 2020; Larreguy et al., 2016). Thus, the nature of ties to village brokers may be an important factor in explaining variation in registration status.

Often, brokers are thought to directly influence voters, and are assumed to have a high degree of information about them (Stokes, 2005). This implies a model of broker behavior that involves brokers directly influencing many of the voters within their sphere of influence. In the context of Uganda, a recent study found that brokers are well-known figures in a given community (Blattman et al., 2019). Brierley and Nathan (2021) found, however, that in the context of Ghana, party brokers are not necessarily well-connected to voters, nor do they have more information than the average voter. This work suggests that brokers may need to work indirectly to maximize their influence on voting outcomes.

Indeed, a strategy that involves directly reaching out to all intended targets is arguably inefficient. If brokers can influence key individuals who then influence their friends, family, and other social ties, brokers will be able to reach far more voters. In other words, party brokers may themselves need intermediaries, and the relational process of political mobilization may be more diffused than previously thought (and conceptualized). By conducting a complete census of the villages in the study area, we are able to examine whether and which social ties matter for voter registration, and in so doing shed additional light on the role brokers may play in this setting.

3 Empirical challenges in the study of registration

Much of the aforementioned research on both individual attributes and relational factors in shaping political participation focuses on turnout rather than the prior step of voter registration. One major reason for the dearth of research on registration is that accessing accurate data on registration rates, much less individual-level correlates of registration, is difficult, and particularly so in low-income country settings.

The calculation of voter registration rates requires two inputs: (a) the number of registered voters, and (b) the number of people who are of voting age in a given year. Reliable estimates for both of these input measures are often not readily available.

First, the denominator, voting age population, is best estimated using census data. However, census data are typically collected every ten years at best and can be of questionable quality. Estimation of adult population figures between census years requires interpolation, which itself entails a number of assumptions. For these reasons, the reliability of voting age population estimates in a given election year is questionable even in the best of conditions (Radcliff, 1996). Subnational measures of registration rates are even more difficult to calculate, as they require additional assumptions about rates of population growth and migration across subnational units. Estimating the voting-age population in a given year at the subnational level is especially problematic for a country like Uganda, with a high population growth rate, variable growth rates across regions, and substantial levels of internal migration and urbanization.

Estimating the number of registered voters, the numerator, is even more problematic than estimating the voting-age population. First, most countries do not make voter registers available to the public. Second, even when they are available, registers are likely to be outdated: they may include individuals who have died or moved away from a given locality, or they may fail to capture those who have moved into a locality. Even in high-income countries like the United States, where one can obtain well-curated individual-level voter files and a wealth of individual-level data collected about registered voters from marketing

firms, individual-level census data are unavailable for reasons of privacy. As a result, we know very little about the unregistered individuals who do not appear in the voter file or accompanying commercial data sets (Jackman and Spahn, 2015).

In the absence of accurate measures of the either the numerator or the denominator, it is difficult or impossible to calculate an accurate voter registration rate even at a national level, much less identify or study those who are unregistered at the micro level. This problem may be particularly acute in lower income countries. For example, Kuenzi and Lambright (2007) show that for one-third of African countries in their sample, the number of official registered voters is larger than the estimated voting age population, suggesting either widespread inflation of voter registers, underestimates of population counts, or both. The result is that the 'true' registration rate is unknown. Worse, without knowing which figure is more or less accurate, even the *direction* of potential bias in this figure is unknown.

Consider, for example, the case of Uganda. In an exploratory exercise, using 2014 census data and 2016 election results published at the parish level, we found that a quarter of the 7,500 parishes in Uganda exhibited a registration rate of 100 percent or greater.[1] This was the case even after increasing the voting-age population to account for voters who came of age between the census and voter registration period. This exploratory exercise underscores the difficulty of using official records to estimate voter registration and the necessity of alternative strategies if we are to empirically address the question of who registers.

How can these challenges be overcome? In this paper we collect original survey data from a complete census we conducted in 16 rural villages in Uganda, and merge it with the official voter register in these villages.[2] Specifically, we match individuals in our village census to

[1]Parishes in Uganda comprise of 3-10 nearby villages.

[2]The 16 villages were drawn from an earlier study the authors conducted in the district as part of a multi-year program called Governance, Accountability, Participation, and Performance (GAPP). GAPP was implemented by RTI International and funded by the United States Agency for International Development (USAID). The number of villages was determined by resource constraints. For more information on village selection criteria, see Supplementary Information (SI), section A.

the voter register, which provides a major advantage, allowing us to advance the empirical examination of determinants of voter registration.[3] Most importantly, our measure of voter registration is behavioral rather than self-reported. An individual in our sample is coded as registered if and only if her name exists in the final certified register itself. Most existing measures of both turnout and registration in similar contexts rely on self-reports, which typically vastly overestimate these behaviors.[4]

By combining our own village census with the official voter register we are able to calculate a more accurate registration rate than is possible using official data or survey data in isolation.[5] To our knowledge, ours is the first study to merge a complete bespoke census with a large voter registration database, allowing for the examination of a rich set of covariates not only for registered voters, but crucially, for unregistered voters as well. With these data in hand, we are able to contribute to a better understanding of the variation in individuals' voter registration status in a low-income country setting, examining two types of explanations that have been posited in the existing literature: individual factors such as SES, and relational factors such as social ties.

[3]We obtained the final register provided to presidential candidates in the run-up to the February 2016 general elections. Importantly, the voter register includes the unique identifiers of all registered voters at both the individual-level (voter ID, surname, first and middle names, date of birth, and sex), and locality-level (constituency, sub-county, parish, village and polling station names). The survey data includes not only individual level information on (almost) all village residents, but also information on ties between any two villagers on multiple dimensions (e.g., family, friendship, lending).

[4]Tellingly, 84% of Ugandan respondents in round 7 of Afrobarometer's nationally representative survey report they have voted in the 2016 general election, while the official turnout rate was 67% among registered voters, or 47% of the estimated adult population.

[5]Only 3% of respondents in round 7 of Afrobaromter's national representative survey in Uganda report they were not registered to vote for the 2016 general election; according to our own estimates, registration rates in Uganda are somewhat less than 70%. In other words, self-reporting of voter registration, at least in Uganda and we suspect elsewhere, is vastly exaggerated.

4 Context

We conduct our study in Uganda, a predominantly rural, low-income country governed by a dominant-party regime. Dominant party regimes comprise an increasing share of regimes globally and are thus substantively important to investigate. Uganda's ruling party, the National Resistance Movement (NRM), has been in power since 1986, and enjoys wide pockets of popular support, especially in rural areas.

While elections under the NRM have not been completely free and fair—the NRM has on multiple occasions resorted to the manipulation of state resources, intimidation, and politicized prosecutions of opposition leaders (Izama and Wilkerson, 2011), and in the most recent election, the alleged abduction of hundreds of suspected opposition supporters (Watch, 2021)—irregularities in vote tallies are typically concentrated in a relatively small number of polling stations, usually NRM strongholds or areas that are difficult for other election observers to access (Klimek et al., 2012; Callen et al., 2016). Overall, turnout rates in general elections range between 60 and 70% of registered voters.[6]

Existing research has documented several turnout strategies in Ugandan elections. For example, candidates use both targeted benefits and promises of group benefits to increase electoral support (Grossman and Lewis, 2014). Political parties in Uganda also make use of paid local brokers, who are well-known figures in the area. According to Blattman et al. (2019), virtually all Members of Parliament in Uganda employed brokers in the 2016 parliamentary elections. However, unlike select other contexts (Brierley and Nathan, 2021), brokers are not selected by a formalized process. As such, much remains unknown about who these brokers are and how they operate in practice. Further, relatively less is known about the mobilization strategies employed for voter registration, which takes place months before a given election, rather than turnout, which takes place over a shorter period of time

[6]For turnout data by year and election type, see `https://www.idea.int/data-tools/country-view/293/40`.

and where there are typically many more domestic and international observers present.

In Uganda, like many countries, the burden of registering to vote falls on individuals, as there is no automatic voter registration. To register to vote, Ugandans need to travel to a registration center and present proof of citizenship and age (18 years and older).[7] Apart from travel costs and waiting time, registration to vote is free of charge.[8] Voter files in Uganda are carried from previous elections, and citizens are encouraged to check their registration status prior to upcoming elections.[9] Voter files are updated prior to elections to remove those who passed away, update the information of those who have migrated and would like to change the location of their assigned polling station, and to add new voters. The register for the 2016 election was compiled and cleaned between 2014 and May 2015, with a final display of the register at each polling station in July and August 2015.[10] We conducted our village censuses in March and April 2016, shortly after the February 2016 general election.

Figure 1 shows the registration rate across the 16 study area villages, compiled by combining our village survey data with the official voter register for Arua district, where all the villages are located. Arua is a relatively rural district, located in the northwest of the country, bordering the Democratic Republic of Congo to the west and just over 400 kilometers by road from the country's capital. Average official district-level turnout (as a share of registered voters) in Uganda during the 2016 elections was 67 percent, and for Arua district was 59 percent. Average district-level vote share for the incumbent president was 64 percent nationally and 58 percent in Arua district. Our estimates of village-level voting patterns

[7]For a complete list of requirements, please consult the voter registration handbook, which the Ugandan electoral commission posts online at `http://ec.or.ug/pub/voter_education.pdf`. The registration process is subject to change and indeed has been changing with the advent of national identification cards.

[8]Unlike some countries where there is wide variation in the voter registration procedural requirements from one locality to another (Kim et al., 1975), registration is centralized and uniform in Uganda, and to our knowledge, the burden of registration does not vary greatly from one community to another.

[9]Ugandans can now verify online their voter registration status at `https://www.ec.or.ug/search/byid`.

[10]The roadmap for the 2016 election is available online at `https://bit.ly/3jmW7Ch`.

indicate that the majority of sampled villages are NRM leaning, but are not ruling party strongholds.[11] The median estimated vote share for the incumbent president in the 2016 elections was 60 percent. This suggests that party brokers may have reasons to try to be selective in their mobilization strategies, rather than targeting all village members.

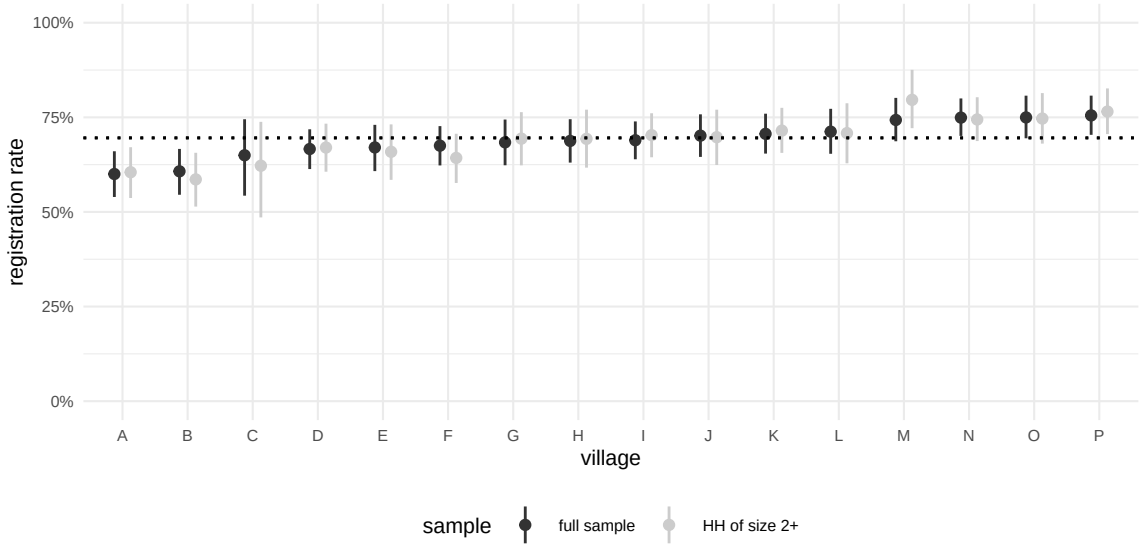

Figure 1: **Registration rates across villages.** The dashed line represents the mean registration rate (69%); it is identical in either sample. Bars are 95% credible intervals, estimated using a Bayesian linear random effect model with random intercepts.

The mean voter registration rate across villages in our sample is 69 percent, indicating that there is still a substantial portion of the adult population unregistered. Village-level census data beyond Arua were not available to us, but we compare parish-level census data in the parishes where the sampled villages are located to parishes in the rest of rural Uganda, excluding cities and municipalities. The parishes in which the study was conducted are similar to other rural parishes in most respects, including population size, educational attainment, access to electricity, reliance on subsistence farming, and prevalence of malnourishment, as shown in Supplementary Information (SI), Figure SI-1. Sampled parishes are somewhat

[11] Estimating village-level vote share for a given candidate is difficult as results are reported at the polling station level, which does not exhibit a one-to-one correspondence with villages in many case. Details about the mapping of polling station election results to villages can be found in Eubank et al. (2021).

more likely to have thatched roofs as compared to iron sheets, which could be an indicator of relatively lower economic status and have a slightly higher share of women.

Importantly, we find that within-village variation in registration rates is much larger than across-village variation: the intraclass correlation is 0.007% (full sample) and 0.011% (for households (HHs) of size 2+). We thus focus our inquiry on within-village variation, using village fixed effects in all regression models to absorb village characteristics that could be correlated with voter registration status.

5 Data

We turn to describe the data and estimation strategy we use to explain within-village variation in voter registration status in rural Uganda. Our key dependent variable is an individual's registration status, which is a binary variable that equals one if the respondent is registered to vote in the 2016 elections, and zero otherwise.

While studies of registration are rare, past research in developing countries suggests that some of the individual- and household-level characteristics associated with turnout in middle- and high-income countries are also important in low-income countries. These include, income (Kasara and Suryanarayan, 2015), age (Kuenzi and Lambright, 2011), gender (Robinson and Gottlieb, 2021), education (Croke et al., 2016), and pro-sociality (Blattman, 2009). By focusing on the registration phase, we explore the possibility that these factors operate in part through their association with voter registration. While we are able to accurately measure registration status, we do not have information on whether an individual voted on election day, as this information is not publicly available in Uganda.

We measure the following individual-level covariates: *age* is a continuous measure of respondent's age in years; *female* is an indicator of the respondent's sex; *secondary education*, a binary variable that equals 1 if the respondent attained at least secondary education; *Catholic* is a binary variable that equals 1 if the respondent identifies as Catholic, which is

the dominant religion in our study area, and 0 if she identifies with another religion (usually Protestant); *Lugbara* is a binary variable that equals 1 if the respondent identifies as Lugbara, which is the dominant ethnic group in our study area,[12] and 0 if she identifies with another ethnic group; *income* is a subjective 5-points categorical measure of wealth ranging from 1 (low) to 5 (high);[13] and *use phone* is a binary variable that equals 1 if the respondent has used a mobile phone in the past 12 months.[14]

We also measure individual covariates that capture involvement in communal affairs; these include: *leader*, a binary variable that equals 1 if the respondent occupies a leadership position within the village;[15] *participation* a continuous summary index aggregating across various recent non-electoral forms of communal engagement;[16] and *pro-sociality*, a behavioral proxy-measure of care for the community, measured as the share of 10 monetary units

[12]We note that at the national level Uganda has no majority ethnic group, such that Lugbara are a minority at the national level. However as our study area is predominantly Lugbara we are not able to examine whether ethnicity matters more broadly for registration, though this is certainly worth examining in future studies.

[13]Income is measured by the following question: "In comparison to other typical households in this village, how would you describe your household's economic situation?", with answers ranging from "Much worse" (1), to "About the same" (3), to "Much better" (5).

[14]Our survey was conducted immediately following the 2016 elections, and our field team suggested that asking about political party affiliation was too sensitive at this time. For this reason we are not able to capture this otherwise important individual-level measure.

[15]We consider the following leadership positions: LC1 (village head), LC3 (subcounty elected leader), LC5 (district elected leader), head of school management committee, head of parent-teacher association, traditional or religious leadership.

[16]We consider attending a village meeting, contributing money to a village project or a village member, contributing labor to a village project, reporting a problem to a village leader, and reporting a problem to the local government, in the past 12 months. The summary index is constructed following Anderson (2008), which gives more weight to more separating components of the index. The index is continuous, standardized and centered in the full sample and has an in-sample range of $[-.9, 1.5]$.

endowment $\in (0, ...1)$ contributed in a standard dictator game that all respondents played with an anonymous same-village recipient.

In addition, we measure two location-based covariates. *Distance (km)* is a continuous measure of the Euclidean distance in kilometers between respondent's geocoded dwelling and the nearest voter registration center. *Geography* captures instead indirect social influence using a spatial lag that counts the number of registered voters within the village besides individual i, and assigns less weight to those who reside farther away from that individual.[17]

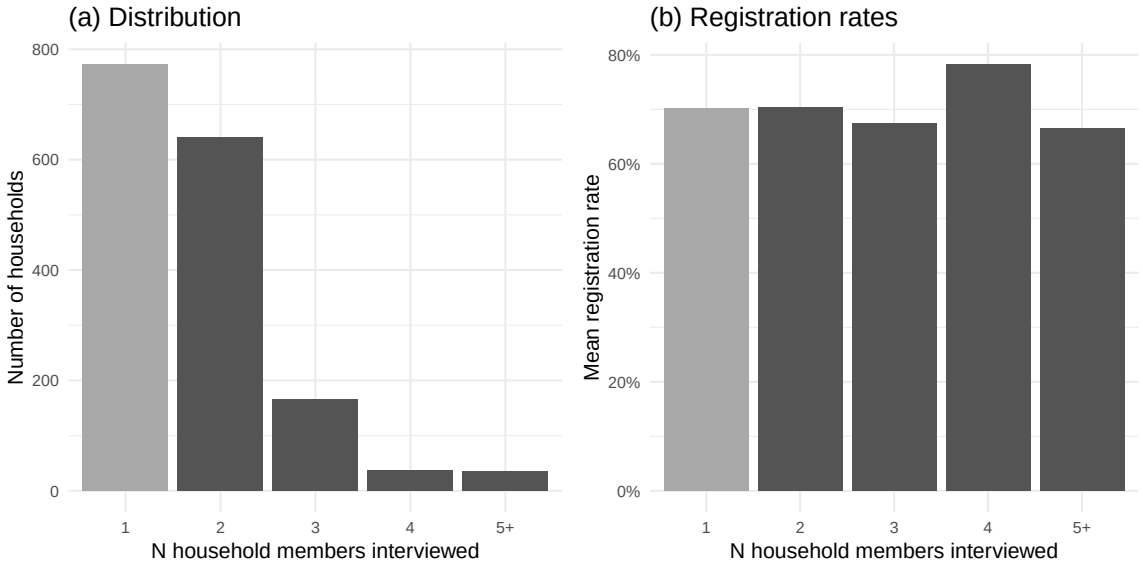

Figure 2: **Distribution of household sizes and registration rates.** The grey bars represent households of size 1, which we later exclude from the analysis. Those 773 households break down about evenly into three groups: first, the households for which only one member could be interviewed ($n = 260$); second, households that actually have one member ($n = 324$); and third, households that were not included in the initial listing exercise and were later added to the sample ($n = 189$).

Finally, we include three household-level factors: *household head* is a self-reported binary measure that is equal to 1 for the head of the household and zero for all other household members; *degree HH*, which is a count variable of the number of adults within a single household, and *% registered HH* which is our proxy measure of peer effects at the household

[17]With $y_i \in \{0,1\}$ i's registration outcome and d_{ij} the distance between i and j, The spatial influence (*geography*) is $\text{geo}_i = \sum_{j \neq i} \frac{y_j}{\log d_{ij}}$.

level. Households are as defined by local leaders with whom we worked to conduct the household listing and enumeration. Figure 2 shows the distribution of household sizes in the study villages and their registration rates, where the size of the household is determined by the number of adults living in the household.

To explore the possibility that registration is driven by network ties beyond the household, we rely on a network module we administered to all village residents as part of the in-person surveys. This module—administered as part of a longer survey that also included modules on respondent demographics and household characteristics, among other—included a set of questions about four different types of social ties respondents may have with other co-villagers. We use a simple name generator technique (Knoke and Yang, 2008) to elicit information on four kinds of relationships: (1) *family* ties, (2) *friendship* ties, (3) *lenders*: to whom they would go to borrow money, and (4) *problem solvers*: to whom they would go to solve a problem regarding public services in the village. This technique involves asking respondents for up to five names of those individuals whom they consider social ties in each of these categories.[18] Through this exercise, it could be that i named j as a peer, while j did not. In other words, this exercise gives us four "directed" networks, i.e. networks that distinguish ties from i to j and ties from j to i. We then collapse those directed networks into undirected networks, i.e. networks in which i and j are either tied or not. To do so, we say that there is a tie between i and j if at least one named the other as a peer.

Because network concepts have simpler definitions with undirected networks, they will be the main focus of our analysis. However, we also report results in which we consider directed networks. Comparing household roster data with network surveys, we believe we reached over 80% of village residents. Following standard practice (Larson and Lewis, 2017), individuals who did not complete a network survey were dropped from the analysis (See SI Section A for details). Table 1 provides descriptive information for the variables used in the empirical analysis.

[18]The exact wording used to elicit these ties is provide in SI Section A.3.

Variable	HH size 1	Sample	Δ_0	HH head	Non-head	Δ_1
Dependent variable						
% registered	0.701	0.696	0.005	0.78	0.639	0.141***
Individual						
% female	0.72	0.526	0.194***	0.163	0.772	-0.609***
age	39.322	37.02	2.302**	43.247	32.809	10.439***
income	2.394	2.586	-0.191***	2.48	2.657	-0.178**
% 2ary education	0.216	0.235	-0.019	0.309	0.184	0.125***
% Catholic	0.629	0.581	0.048	0.561	0.594	-0.033
% Lugbara	0.948	0.946	0.002	0.942	0.949	-0.007
pro-sociality	0.202	0.201	0.001	0.213	0.193	0.02**
leader	0.125	0.151	-0.026	0.24	0.091	0.148***
participation index	-0.076	0.037	-0.113***	0.194	-0.069	0.263***
distance to registration center (km)	0.847	1.017	-0.17**	0.952	1.062	-0.11**
Network						
degree centrality	12.34	17.75	-5.409***	23.843	13.629	10.214***
closeness centrality	0.002	0.002	0	0.003	0.002	0***
clustering coefficient	0.421	0.371	0.05***	0.333	0.396	-0.063***
eigenvector centrality	0.128	0.169	-0.041***	0.221	0.134	0.088***
Sample size						
N	773	2184		881	1303	

Table 1: **Descriptive statistics**. Column 2 ('Sample') reports descriptive statistics for all members of households of size 2 and above. We use this sample in our regression analysis below. The column Δ_0 reports the difference between the columns 'HH size 1' (those excluded from the empirical analysis) and 'Sample'. The column Δ_1 reports the difference between the columns 'HH head' and 'Non-head,' among households of size 2 and above. *Income* is a subjective 5-point categorical measure of wealth ranging from 1 (low) to 5 (high); *Participation index* a continuous summary index aggregating across various recent non-electoral forms of communal engagement with an in-sample range of $[-.9, 1.5]$. Standard errors are clustered at the village level. *p<0.1; **p<0.05; ***p<0.01

Estimation

To explore the relationship between registration status (our outcome of interest) and the above individual attributes and relational ties, we fit the following linear probability model:

$$y_{ij} = \alpha_j + x_i'\beta + f_j(i)'\gamma + \epsilon_{ij}$$

with y_{ij} a binary variable that equals 1 if respondent i in village j is registered, and equals 0 otherwise. The vector x_i is a vector of individual attributes. The function $f_j : N_j \to \mathbb{R}^k$ maps to each respondent $i \in N_j$, the set of respondents in village j, a vector of network statistics derived from the network of village j. As such, while parameters β capture the association between individual attributes and registration, γ parameters capture the association between registration and network attributes. Among those attributes, two are particularly important: *degree*, the number of peers of respondent i, and the *percentage of registered peers*. Because the dynamics of registration may depend on unobserved village-level characteristics, we perform within-village comparisons and include a village-level fixed effect α_j. We cluster standard errors at the village level and use a wild bootstrapping procedure, due to the small number of clusters.

Note that in this specification we use the *percentage* of registered peers, and not the *number* of registered peers, which amounts to estimating a relative threshold model instead of an absolute threshold model (Centola and Macy, 2007). Indeed, because they only count registered peers, absolute threshold models implicitly assume that only such peers influence outcomes. Conversely, because they consider the share of registered peers, relative threshold models implicitly assume that both registered and non-registered peers may influence the outcome. Theoretically, relative threshold models better capture the dynamics of influence that may underlie registration, since non-registered peers may also influence an individual. As a robustness check (Table SI-6), we re-estimate our main results using an absolute threshold model instead and find that both measures produce comparable results (and models fit).

Note that this specification cannot be given a causal interpretation. This is especially the case for network parameters γ, which may suffer from three sources of bias: homophily, contextual confounding, and the reflection problem (Aral et al., 2009). Nonetheless, the large number of individual controls included in x_i, as well as the village-level fixed effect α_j help to alleviate at least some concerns about omitted variable bias. While this exploratory exercise is admittedly correlational, it aims to answer a basic and important question—*who registers?*— with a goal of generating hypotheses that can inform future research.

6 Results

We begin by presenting our findings with respect to a set of individual attributes that have been found to be associated with political participation in a variety of contexts (Table 2, columns 1 to 4).

We find little evidence that factors such as education, income, gender, or membership in a minority ethnic or religious group are associated with registration status, casting doubt that resource-based theories of political participation apply to the behavior of registration in this context as they do in rich democracies.

Note, however, that some individual attributes do matter for registration status, including age, being a household head, and holding a leadership position within the village. Regarding age, we find a curvilinear relationship with registration (Figure 3), though the small sample of older respondents means that we should be cautious in interpreting the relationship between age and registration status among this population. It does seem, however, that those above 25 are significantly more likely to be registered than those between 18 (the legal age for registration) and 25. This finding is consistent with a robust curvilinear relationship between age and turnout observed in high-income consolidated democracies (Bhatti et al., 2012).

	Dependent variable: registered						
	All	All (2+)	Head	Non-Head	All (2+)	Head	Non-Head
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
female	0.017	0.008	0.016	0.00000	0.014	0.027	0.005
	(0.027)	(0.028)	(0.035)	(0.033)	(0.029)	(0.037)	(0.035)
age	0.015***	0.018***	0.002	0.028***	0.018***	−0.0001	0.028***
	(0.003)	(0.005)	(0.007)	(0.006)	(0.005)	(0.007)	(0.006)
age^2	−0.0002***	−0.0002***	−0.00004	−0.0003***	−0.0002***	−0.00002	−0.0003***
	(0.00004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
income	−0.006	0.002	0.026**	−0.010	0.001	0.027**	−0.011
	(0.007)	(0.007)	(0.011)	(0.009)	(0.006)	(0.011)	(0.009)
2ary education	0.021	0.025	0.011	0.054	0.024	0.017	0.049
	(0.023)	(0.024)	(0.034)	(0.041)	(0.023)	(0.032)	(0.041)
Catholic	0.032	0.037	0.041	0.022	0.037	0.038	0.021
	(0.040)	(0.044)	(0.060)	(0.044)	(0.042)	(0.061)	(0.041)
Lugbara	0.076	−0.018	0.029	−0.022	−0.016	0.040	−0.025
	(0.047)	(0.050)	(0.066)	(0.080)	(0.050)	(0.068)	(0.086)
leader	0.052**	0.035	0.011	0.092**	0.036	0.010	0.097***
	(0.026)	(0.024)	(0.034)	(0.037)	(0.024)	(0.034)	(0.035)
pro-sociality	−0.012	−0.016	0.009	−0.042	−0.013	0.015	−0.037
	(0.056)	(0.048)	(0.057)	(0.083)	(0.045)	(0.056)	(0.080)
participation index	0.020	0.025	0.003	0.037	0.028	0.003	0.040
	(0.014)	(0.018)	(0.028)	(0.025)	(0.017)	(0.029)	(0.025)
distance (km)	0.011	0.018	−0.003	0.035*	0.014	−0.009	0.029
	(0.012)	(0.016)	(0.014)	(0.020)	(0.016)	(0.014)	(0.020)
geography	−0.005	−0.002	−0.007	0.003	−0.006	−0.013	−0.001
	(0.007)	(0.007)	(0.009)	(0.008)	(0.008)	(0.010)	(0.008)
household head	0.074***	0.082***			0.095***		
	(0.023)	(0.026)			(0.028)		
degree HH					0.008	0.022	0.008
					(0.005)	(0.014)	(0.006)
% registered HH peers					0.061*	0.089***	0.089**
					(0.035)	(0.032)	(0.039)
Observations	2,753	1,940	785	1,155	1,940	785	1,155
R^2	0.049	0.065	0.033	0.094	0.069	0.043	0.099
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 2: **Determinants of individual voter registration across 16 Ugandan villages**. Model 1 considers the full sample. Models 2 to 7 focus on households of size 2 and more, for which within-household peer effects are well-defined. Within these households, models 3 and 6 focus on household heads, while models 4 and 7 focus on non-head members.

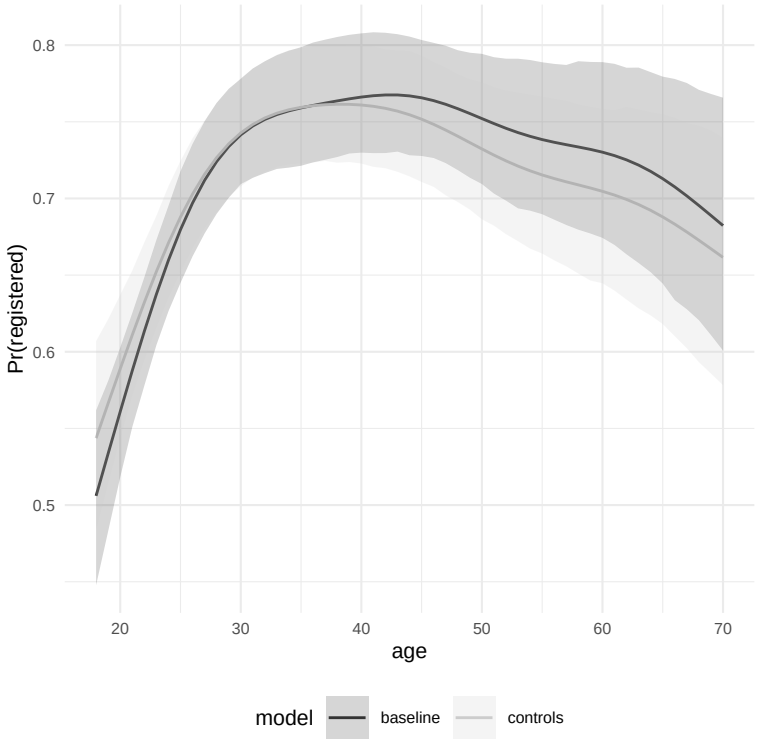

Figure 3: **Relationship between age and voter registration status.** Estimates from generalized additive models (GAM) with thin-plate splines. Both models (with and without control) include village fixed effects. The specification with controls includes the covariates reported in Table 2, model 1.

We estimate that being the head of a household increases the likelihood of being registered to vote by 8 percentage points. Those who hold leadership positions (as defined above) are also more likely to be registered to vote, though this relationship seems driven primarily by those individuals who hold a leadership position but are not household heads, a relatively small group. Still, among this group the effect is quite large, 9.2 percentage points, as shown in Table 2, column 4.

As mentioned, we do not find evidence that traditional proxies of marginalization at the individual-level—such as being female, relatively poor, a member of local religious minority (i.e. not being Catholic) or local minority ethnic group (i.e. not being Lugbara), or living at the village periphery[19]—matter for registration status among in our rural sample (Table 2,

[19]Residential-based peripherality is proxied by both *geography* and *distance*, as defined above.

columns 1 to 4). Similarly, (not) having attained secondary-level education does not correlate with registration status. Importantly our null finding for education is likely not a function of how education attainment is coded.[20] An examination of the bivariate correlations between these attributes and registration status suggests that multicollinearity does not explain the absence of significant relationships observed in the multivariate model (SI, Table SI-5). Our index of political participation is not associated with registration in the regression analysis, but it does exhibit a high degree of correlation in the bivariate analysis. The lack of a clear relationship between marginalization and registration are important when considering that past work has demonstrated that these marginalization factors do matter for voting and other forms of participation in Uganda (Grossman et al., 2014).

We urge caution in the interpretation of this latter set of findings, however, for a few reasons. First, it could be that the effects of an attribute like income or education are cross-cutting. For example, low levels of education, particularly when proxying illiteracy, could make it more difficult to register to vote in practical terms, but on the other hand, it could be that those with more education are less likely to engage politically in a quasi-democratic context because they view such efforts as futile (Croke et al., 2016). Thus, it could be that education does indeed have causal effects on registration, but that these effects cancel each other out. The same logic applies to income (registration costs reduce, but vote buying increase the propensity of registration for poor constituents). Second, it could be that there is insufficient variation in either education or income for us to pick up an association between these attributes and registration, but that one would emerge if we included places

[20]When using an alternative categorization of education levels we find some weak evidence that those with no education are less likely be registered than those with some primary or some secondary education (SI Table SI-7), but this relationship is not very robust. We reproduce Table 2, but use a categorical variable for education. We find evidence that being more educated positively correlates with registration only when considering the entire sample. This correlation disappears when considering smaller samples (i.e. columns 2 to 4). Admittedly, such pattern may owe to either within-household interactions dwarfing the impact of education, or lack of power.

with greater variation, particularly urban areas. Third, individual attributes may have an effect on registration for some sub-samples of potential voters. For example, when focusing on the sample of household heads (column 3), we find some evidence of an income effect: a one unit increase on the 5-point categorical income scale is associated with 2.5 percentage point increase in the probability of registration (Table 2, column 3).

Together, our findings on individual attributes suggest that existing models of political participation—especially resource-based models—do not find strong support in our examination of voter registration in this rural setting. This does not mean that these attributes are entirely irrelevant, but they may matter less for the outcome of registration that for turnout, less in rural Uganda than in more urban or wealthy settings, or perhaps operate differently in lower income or less democratic settings.

Social ties and voter registration: a diffuse process of influence

Next, we turn to our findings with respect to a set of relational factors. Our results provide evidence that social ties matter and suggest a pattern of diffuse social influence in shaping registration behavior. Specifically, we find evidence of peer effects within and across households, as well as evidence that social ties to some notable individuals, who we believe are most likely to be political brokers, shape registration status.

Within-household peer effects

We focus on households with at least two adult members, since the notion of within-household peer effects is not defined for households of size 1. Above, we found that household heads are more likely than non-heads to be registered. Here, we find that whether the head of household is registered is positively associated with the registration status of other household members, suggesting a role of within-household peer effects, and indicating that there are clusters of registered and unregistered households.

First, we note that voter registration increases dramatically with the share of "other"

household adult members being registered (Table 2, column 5). For both household head (column 6) and non-household head members (column 7), moving from no other household members registered to all registered, increases registration by about 8.9 percentage points, when accounting as is customary for household size (*degree HH*). In sum, within-households peer effects on political engagement are consequential.

Second, we investigate how much influence is exerted by different household members (Table 3). To do so, we focus on non-head members in households with at least 2 adult members (that is to say, the full sample of multi-person households), and consider first the role of any member (column 1), then the role of the household head only (column 2). Considering column 2, we find that if the household head is registered to vote, this increases the likelihood that other non-head adults are registered from 0.57 to 0.67 (9.6 pp, or 17%), a magnitude that is comparable to the joint effect of any member (column 1).

Focusing on large households (3 adult members or more; columns 3 to 5) allows us to evaluate more directly the finding that household heads convey disproportionate influence, by comparing peer effects of the household head against those of 'other' household members that are not the head.

Note that the sample size is rather small for households with three or more adult members. Although results are not statistically significant, we replicate our findings from the full sample of multi-person households (columns 3 and 4). In column 5, we estimate both effects jointly and find that there is a large and positive coefficient, though not statistically significant, for the variable indicating wither the head of household is registered, but negative for the percent of non-heads who are registered, as shown in Table 3, column 5. This suggests that within-household registration dynamics are driven by household heads, and not by "other" non-heads members.

	Dependent variable: registered				
	Size 2+	Size 2+	Size 3+	Size 3+	Size 3+
	(1)	(2)	(3)	(4)	(5)
female	0.005	0.006	−0.043	−0.043	−0.046
	(0.035)	(0.034)	(0.042)	(0.040)	(0.042)
age	0.028***	0.028***	0.029***	0.030***	0.030***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
age^2	−0.0003***	−0.0003***	−0.0003***	−0.0003***	−0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
income	−0.011	−0.011	−0.002	−0.002	−0.002
	(0.009)	(0.009)	(0.013)	(0.013)	(0.014)
2ary education	0.049	0.046	0.088	0.083	0.086
	(0.041)	(0.041)	(0.062)	(0.058)	(0.062)
Catholic	0.021	0.020	0.004	0.001	0.002
	(0.041)	(0.041)	(0.060)	(0.058)	(0.059)
Lugbara	−0.025	−0.022	0.069	0.072	0.072
	(0.086)	(0.088)	(0.066)	(0.070)	(0.070)
leader	0.097***	0.098***	0.070	0.071	0.071
	(0.035)	(0.035)	(0.071)	(0.071)	(0.072)
pro-sociality	−0.037	−0.039	0.004	−0.001	−0.001
	(0.080)	(0.079)	(0.100)	(0.098)	(0.098)
participation index	0.040	0.037	0.015	0.012	0.010
	(0.025)	(0.025)	(0.029)	(0.027)	(0.027)
distance (km)	0.029	0.029	0.070***	0.067***	0.070***
	(0.020)	(0.020)	(0.017)	(0.015)	(0.016)
geography	−0.001	−0.002	0.010	0.008	0.009
	(0.008)	(0.008)	(0.011)	(0.010)	(0.011)
degree HH	0.008	0.011**	0.006***	0.008***	0.008***
	(0.006)	(0.005)	(0.002)	(0.002)	(0.002)
% registered	0.089**		−0.031		
	(0.039)		(0.108)		
head registered		0.096***		0.058	0.061
		(0.033)		(0.047)	(0.051)
% registered non-head					−0.042
					(0.091)
Observations	1,155	1,155	586	586	586
R^2	0.099	0.102	0.129	0.131	0.132
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3: **Effect of household head.** This table only considers non-head members. Column 1 reproduces column 7, Table 2 to facilitate comparisons.

Network centrality and registration status

In addition to examining the role of within-household peer effects, the unique nature of our network data allows moving beyond within-household peer effects, and exploring the relationship between individuals' position in their village network and registration status.

First, we find that villagers with a greater number of connections in the community are more likely to be registered: in Table 4, the coefficient for "degree"—that is, the number of peers that a respondent has—is always positive and significant. The first column in Table 4 is our within-household model, and "degree" simply refers to the size of the household. Columns 2 through 6 consider outside-household peer effects, first considering any kind of tie to a non-household member (column 2), and then, respectively, the respondent's reported ties to family, friends, lenders/borrowers, and people that respondent would turn to, or that would turn to that respondent, to solve a problem.[21] The coefficients for "degree" in columns 2 through 6 indicate that no matter which types of ties we consider, those who are more connected to others in the community are more likely to be registered. We demonstrate in the appendix that this finding is robust to using alternative ways of capturing an individual's network centrality, such as closeness, clustering and eigenvector centrality (Table SI-8, and Appendix A.1 for a glossary of those measures).

It is not surprising that better-known individuals, who occupy relatively central positions in the village social network, are more likely to be registered to vote, and perhaps more politically active in general. Yet full network data of this kind are rare, and we know of no other study documenting the relationship between network position and voter registration. Moreover, this finding also serves as an important reality check on both our outcome measure, as well as our measures of network ties.

Second, we find that some outside-household ties are more consequential than others. The models presented in Table 4 also include the "percent of registered peers" for the respondent in each type of network. The coefficients are indistinguishable from zero for the family, friends, and problem-solver networks. While individuals who are central in these networks are more likely to be registered, we do not see evidence of influence such that individuals with larger shares of registered peers are more likely to be registered.

[21] See SI, Section A.3 for verbatim excerpts from our in-person survey used to construct adjacency matrices capturing the four different types of within-village network ties.

	Dependent variable: registered					
	Household	Non-HH	Family	Friends	Lender	Solver
	(1)	(2)	(3)	(4)	(5)	(6)
female	0.014	0.019	0.030	0.012	0.013	0.009
	(0.029)	(0.028)	(0.027)	(0.027)	(0.028)	(0.029)
age	0.018***	0.016***	0.015***	0.010**	0.015***	0.018***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***	−0.0002***	−0.0001**	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.0001)	(0.0001)
income	0.001	−0.001	−0.001	−0.003	−0.002	0.001
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
2ary education	0.024	0.013	0.019	0.017	0.021	0.021
	(0.023)	(0.023)	(0.022)	(0.024)	(0.024)	(0.025)
Catholic	0.037	0.042	0.042	0.044	0.042	0.039
	(0.042)	(0.042)	(0.043)	(0.042)	(0.042)	(0.044)
Lugbara	−0.016	−0.014	−0.013	−0.018	0.002	−0.017
	(0.050)	(0.047)	(0.044)	(0.046)	(0.046)	(0.049)
leader	0.036	−0.005	−0.006	−0.015	0.003	0.024
	(0.025)	(0.028)	(0.026)	(0.026)	(0.025)	(0.026)
pro-sociality	−0.013	−0.004	−0.002	0.008	−0.006	−0.013
	(0.045)	(0.044)	(0.042)	(0.038)	(0.044)	(0.047)
participation index	0.028	0.011	0.010	0.001	0.010	0.021
	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
distance (km)	0.014	0.014	0.011	0.007	0.013	0.018
	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)	(0.016)
geography	−0.006	−0.008	−0.010	−0.015**	−0.008	−0.003
	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)
household head	0.095***	0.065**	0.052**	0.044	0.065**	0.078***
	(0.028)	(0.027)	(0.026)	(0.029)	(0.028)	(0.028)
degree	0.008	0.004***	0.016***	0.033***	0.014***	0.001**
	(0.005)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)
% registered peers	0.061*	−0.028	0.041	0.038	0.079***	0.032
	(0.035)	(0.068)	(0.032)	(0.036)	(0.026)	(0.048)
Observations	1,940	1,940	1,940	1,940	1,940	1,940
R^2	0.069	0.078	0.086	0.113	0.082	0.067

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: **Comparison of within-household dynamics with other possible peer effects.** Sample is restricted to household with size 2 and above. Definition of ties can be found in the Supplementary Information (SI), section A.3.

However, we do find strong evidence of such influence for one specific type of network—that of lender-borrower relationships. In fact, a comparison of columns 1 and 5 in Table 4 indicates that this type of peer effect is similar in magnitude to the rather large within-

household peer effects described above.[22]

If politicians are relying on local intermediaries to mobilize registration among likely allies, we would expect to see influence flowing not necessarily through friends or family members, but through a class of influential and politically connected community members. Below, we consider more formally the possibility that those identified as local lenders are playing such a mobilization role.

Investigating the role of political brokers

Existing research suggests that political brokers play a central role in mobilizing potential voters. If this is the case, we would expect to find that those who have social ties with brokers are more likely to be registered than those who are not. While we do not have a measure of whether a given individual is a broker, past work in Uganda suggests that elected village and subcounty heads are very likely to be brokers (Blattman et al., 2019). Additionally, our finding that, out of non-household ties, only the lending network carries significant peer effects (shown in Table 4, column 5) hints to the possibility that lenders and borrowers may have a patron-client relationship, in which lenders act as more proximate political brokers. As such, we examine two kinds of relationships: lending relationships, as well as relationships to the village head, also called a Local Council 1 chairperson, or shortly – LC1.

We first confirm that lenders likely act as patrons to their borrowers. As discussed above, only among the network of lenders, and not among other types of networks, do we find evidence that social ties are associated with registration status. Lenders are also significantly more likely to be registered to vote (79%) than the average person, lending further support to the possibility that (at least some) lenders act as something akin to political brokers.[23]

[22]This finding is robust to various model specifications, including adding controls for household size and within-household peer effects (Table SI-9).

[23]For an extended discussion on features of the lender network see SI, Section C.

	Dependent variable: registered		
	(1)	(2)	(3)
female	0.013	0.006	0.006
	(0.029)	(0.029)	(0.028)
age	0.015***	0.018***	0.018***
	(0.006)	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)
income	−0.002	0.00004	0.0001
	(0.006)	(0.007)	(0.007)
2ary education	0.020	0.026	0.027
	(0.024)	(0.024)	(0.024)
Catholic	0.042	0.037	0.038
	(0.042)	(0.042)	(0.043)
Lugbara	−0.0001	−0.011	−0.014
	(0.044)	(0.050)	(0.049)
leader	0.004	0.032	0.033
	(0.025)	(0.024)	(0.024)
pro-sociality	−0.007	−0.008	−0.013
	(0.044)	(0.046)	(0.048)
participation index	0.009	0.016	0.020
	(0.017)	(0.018)	(0.018)
distance (km)	0.013	0.016	0.018
	(0.015)	(0.016)	(0.016)
geography	−0.007	−0.004	−0.003
	(0.007)	(0.007)	(0.007)
household head	0.066**	0.079***	0.080***
	(0.027)	(0.027)	(0.027)
degree > 0	0.041		
	(0.059)		
degree	0.014***		
	(0.002)		
% registered lenders	0.058*		
	(0.033)		
out degree > 0		−0.033	
		(0.041)	
out degree		0.014***	
		(0.005)	
% registered out lenders		0.082***	
		(0.024)	
most popular lender not registered			0.023
			(0.042)
most popular lender registered			0.072**
			(0.030)
Observations	1,940	1,940	1,940
R^2	0.082	0.071	0.069
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5: **Effect of lender.** Model 3: whether the most popular lender of agent i is registered. Reference category is i has no lender.

However, the analysis shown in Table 4, column 5, made two simplifying assumptions. First, it combined individuals who have no lending ties with individuals who have lending ties but whose alters – the other individual with whom someone reports a social tie – are not registered to vote. Second, the analysis in Table 4 considered *undirected* lending ties, thereby omitting information about who was the lender and who was the borrower for a given social tie. To relax these simplifying assumptions and further investigate the idea that village lenders likely act as political brokers, we conduct additional analyses. First, we explicitly model those individuals that have at least one lending tie, shown in Table 5, column 1. Second, in column 2, we also take into account the direction of the tie, considering specifically outward ties, that is, ties *to* a lender. Addressing both issues confirms our previous finding that lenders exert influence on their borrowers (column 2).

Furthermore, individuals may borrow from several lenders, whose "popularity" may vary, in the sense that such lenders may have a different number of borrowers. Should lenders act as brokers, then more popular lenders should wield more influence. To test this implication, we zoom in on each individual's most popular lender; that is, their lender with most borrowers (column 3). Comparing the impact of all lenders to that of one's most popular lender (i.e., columns 2 vs. 3), we find that most of the effect of lenders is simply a function of whether or not one's most popular lender is registered to vote. Finally, we show that the only type of lenders that matter for voter registration are those who are *not also family members*, further suggesting that it is only the lenders who are most likely to be brokers (rather than simply family members) that exert political influence and mobilize registration (see Table SI-10).

In Table 6 we examine instead the association between ties to the LC1 and voter registration. Examining relationships to the LC1 is challenging since LC1s are well-connected, leaving few people without ties to the LC1. As such, we restrict our attention to two analyses; one that considers direct lending relationships to the LC1 and one that focuses on indirect relationships, since influence most likely trickles down from the LC1 to an individual through a series of patron-client relationships.

	Dependent variable: registered	
	(1)	(2)
female	0.009	0.012
	(0.029)	(0.030)
age	0.018***	0.018***
	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***
	(0.0001)	(0.0001)
income	0.002	0.001
	(0.007)	(0.007)
2ary education	0.025	0.026
	(0.024)	(0.024)
Catholic	0.037	0.039
	(0.044)	(0.045)
Lugbara	−0.018	−0.018
	(0.051)	(0.051)
leader	0.035	0.042*
	(0.024)	(0.023)
pro-sociality	−0.016	−0.013
	(0.048)	(0.049)
participation index	0.025	0.020
	(0.018)	(0.018)
distance (km)	0.018	0.017
	(0.016)	(0.017)
geography	−0.003	−0.005
	(0.007)	(0.008)
household head	0.081***	0.079***
	(0.027)	(0.028)
distance = 1	0.012	0.087***
	(0.023)	(0.030)
distance = 2		0.072**
		(0.034)
distance = 3		0.075**
		(0.034)
Observations	1,940	1,940
R^2	0.065	0.070
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 6: **Effect of distance to LC1 on registration status.** In Model 1, the reference category for the variable "distance" is i has no tie to an LC1 (i.e. i is more than 1 step away from an LC1). In Model 2, the reference category for that variable is i is more than 3 steps away from an LC1. Both models include village fixed effects.

First, we examine (directed) lending ties to LC1s. Specifically, we compare individuals who report that they will reach out to the LC1 if in need of a loan to those who do not; that is, we compare individuals who are one step away from the LC1 to individuals who are further away. This comparison suggests that those individuals are not significantly more

likely to register to vote than villagers who are further away, as shown in Table 6, column 1. While at first sight, this may suggest that LC1s hold little influence on voter registration, it is also possible that their political influence extends further than their immediate connections.

As such, we then investigate whether the LC1's influence extends further than his immediate connections (i.e., borrowers), by considering separately individuals that are 2 and 3 steps away from the LC1 (column 2). We find that the LC1's influence extends up to 3 steps away, with little decay from steps 1 to 3, implying a sharp decay in influence for those individuals that are farther than 3 steps away from the LC1; that is, the reference category (Table 6, column 2).[24]

Together, our findings on lenders, LC1s, and within-household ties are consistent with a diffuse process of influence, whereby LC1s act as patrons to a first circle of clients, who then act as patrons to a second circle of clients, and so on. Such process extends up to three degrees away from the LC1. Note that those circles of patron-client relationships that radiating from the LC1 are likely to involve household heads. While we do not find evidence that household heads are more likely to be influenced by LC1s than non-heads,[25] we do find that household heads are more likely to be connected to LC1s in the first place, both in the lending and strong ties networks (Table SI-11).

For this reason, it appears that social influence travels from the LC1 to household heads through a patron-client network, ultimately reaching other family members through the disproportionate influence that household heads exert upon other household members. While existing work suggests brokers work through direct contact with potential voters, we find evidence of a much more indirect process of influence—one that is potentially more efficient.

[24]Our sample has too few individuals that are more than 3 steps away from the LC1 to pin down precisely where influence decays.

[25]Available upon request.

7 Discussion

In this study, we examine the relationship between voter registration status and both individual attributes and relational ties in a sample of Ugandan villages. Our analysis is premised on the idea that learning about who registers to vote can shed light on the logic of electoral mobilization in lower income settings, as well as identifying whose voices are not being heard and represented in these contexts. We make three key contributions.

First, we highlight the challenges of measuring voter registration, particularly in lower income settings, and provide a novel strategy to accurately measure voter registration in a sample of villages that does not depends on self-reports nor on adult population estimates. By combining the official voter register with survey data conducted with all adults in the villages in our sample, we are able to both estimate a more accurate voter registration rate for these villages as well as examine individual predictors of voter registration. We find that actual registration rates are significantly lower than registration estimates based on both self-reports (using publicly available surveys such as the Afrobarometer) and on rates based on official registration statistics. We thus urge caution in relying on official voter registration rates given limitations in the accuracy of information about the voting age population and in the number of people who have been officially registered to vote, and particularly registration rates reported at subnational levels.

Second, we do not find evidence that individual attributes commonly thought to affect political participation, including education and income, are associated with registration status in our study area. Neither do we find evidence that groups thought of as being politically marginalized, including women or members of ethnic or religious minorities, are less likely to be registered. These findings underscore the limitations of generalizing core findings from high income countries, and particularly the United States, to lower income countries and to less democratic settings. An interesting question for future analysis is whether proxies of marginalization such as gender, income, or educational disparities emerge as consequential at the stage of voting in rural Uganda, even though such disparities do not seem to be highly

consequential at the registration stage.[26]

Nevertheless, we urge caution in interpreting our results as suggestion that factors such as education and income do not matter for voter registration in lower income settings. These factors could have positive and negative effects on voter registration, canceling each other out, and may matter more in urban settings, which constitute a growing share of the population in rapidly urbanizing lower income countries. Future research could go further in examining the conditions under which, if at all, sociodemographic factors such as education and income matter for voter registration in lower income settings.

Third, voter registration status in rural Uganda is strongly related not only to one's social ties to other members of one's household (Table 2), but also to other members of the community (Table 4). Importantly, we find that intra-household dynamics are crucial for understanding voter registration, and that this relationship is driven by the household head, irrespective of the head's gender. In addition, registration is increasing in how connected an individual is to others, and how central a individual is in her village's social network (Appendix, Table SI-8), which is consistent with social rather than individualized mobilization (Eubank et al., 2021).

Further, our analysis reveals evidence of a diffuse process of social influence in the registration process, with individuals who lend money and elected village leaders (LC1s) playing key but often indirect roles in mobilizing residents to register to vote. Previous work attributed much influence to a small number of core brokers. We find that while the distance to the village leader matters in explaining registration status, there is little decay until one is quite removed, suggesting a large role for indirect influence. Household heads are more likely than the average person to have a close tie to the village head, and household heads themselves exert influence on the registration of other household members. This indirect

[26]Exploratory analysis using the Afrobarometer round 7 from Uganda suggest that women and more educated respondents in rural Uganda are somewhat less likely to report that they voted, yet this should be taken with a grain of salt as voting in the Afrobarometer is self-reported.

and diffuse process of influence is intuitive in the sense that it would be inefficient for a broker to target each and every potential voter. Rather, a diffuse and indirect process allows brokers themselves to rely on intermediaries to influence others to register. Future work could employ a similar strategy to examine whether the results reported here travel to other contexts, as well as the implications of the distinct mobilization strategies underlying voter registration and voter turnout.

References

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Aral, S., Muchnik, L., and Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549.

Berinsky, A. (2005). The perverse consequences of electoral reform in the united states. *American Politics Research*, 33(4):471–491.

Bhatti, Y., Hansen, K. M., and Wass, H. (2012). The Relationship between Age and Turnout: A Roller-coaster Ride. *Electoral Studies*, 31(3):588–593.

Blattman, C. (2009). From Violence to Voting: War and Political Participation in Uganda. *American Political Science Review*, 103(2):231–247.

Blattman, C., Larreguy, H., Marx, B., and Reid, O. R. (2019). Eat widely, vote wisely? lessons from a campaign against vote buying in uganda. *NBER Working Paper No. 26293*.

Blaydes, L. (2010). *Elections and distributive politics in Mubarak's Egypt*. Cambridge University Press.

Bonica, A., Grumbac, J. M., Hill, C., and Jefferson, H. (2020). All-mail voting in colorado increases turnout and reduces turnout inequality. *Working paper, Stanford University*.

Brady, H. E., Verba, S., and Schlozman, K. L. (1995). Beyond SES: A Resource Model of Political Participation. *American Political Science Review*, 89(2):271–294.

Brierley, S. and Kramon, E. (2020). Party campaign strategies: Rallies, canvassing and handouts in ghana. *African Affairs*, 119(477):587–603.

Brierley, S. and Nathan, N. (2021). The connections of party brokers. *Journal of Politics*, Forthcoming.

Callen, M., Gibson, C. C., Jung, D. F., and Long, J. D. (2016). Improving Electoral Integrity with Information and Communications Technology. *Journal of Experimental Political Science*, 3(1):4–17.

Centola, D. and Macy, M. (2007). Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734.

Croke, K., Grossman, G., Larreguy, H. A., and Marshall, J. (2016). Deliberate Disengagement: How Education Can Decrease Political Participation in Electoral Authoritarian Regimes. *American Political Science Review*, 110(3):579–600.

Cruz, C., Labonne, J., and Querubin, P. (2017). Politician Family Networks and Electoral Outcomes: Evidence from the Philippines. *American Economic Review*, 107(10):3006–37.

Downs, A. (1957). *An Economic Theory of Democracy*. Harper & Row, New York.

Eubank, N., Grossman, G., Platas, M., and Rodden, J. (2021). Viral voting: Social networks and political participation. *Quarterly Journal of Political Science*, Forthcoming.

Gallego, A. (2014). *Unequal Participation Worldwide*. Cambridge University Press New York, USA.

Grossman, G., Humphreys, M., and Sacramone-Lutz, G. (2014). "i wld like u wmp to extend electricity 2 our village": On information technology and interest articulation. *American Political Science Review*, 108(3):688–705.

Grossman, G. and Lewis, J. I. (2014). Administrative unit proliferation. *American Political Science Review*, 108(1):196–217.

Grumbach, J. and Hill, C. (2021). Rock the Registration: Same Day Registration Increases Turnout of Young Voters. *Journal of Politics*, Forthcoming.

Harris, J. A., Kamindo, C., and van der Windt, P. (2021). Electoral administration in fledgling democracies: Experimental evidence from kenya. *Journal of Politics*, forthcoming.

Hershey, M. R. (2009). What We Know about Voter-ID Laws, Registration, and Turnout. *PS: Political Science and Politics*, 42(1):87–91.

Ichino, N. and Schündeln, M. (2012). Deterring or displacing electoral irregularities? spillover effects of observers in a randomized field experiment in ghana. *The Journal of Politics*, 74(1):292–307.

Isaksson, A.-S., Kotsadam, A., and Nerman, M. (2014). The gender gap in african political participation: Testing theories of individual and contextual determinants. *Journal of Development Studies*, 50(2):302–318.

Izama, A. and Wilkerson, M. (2011). Uganda: Museveni's Triumph and Weakness. *Journal of Democracy*, 22(3):64–78.

Jackman, S. and Spahn, B. (2015). Unlisted in America. Unpublished manuscript.

Kasara, K. and Suryanarayan, P. (2015). When Do the Rich Vote Less Than the Poor and Why? Explaining Turnout Inequality across the World. *American Journal of Political Science*, 59(3):613–627.

Kim, J.-O., Petrocik, J. R., and Enokson, S. N. (1975). Voter Turnout Among the American States: Systemic and Individual Components. *American Political Science Review*, 69(1):107–123.

Klimek, P., Yegorov, Y., Hanel, R., and Thurner, S. (2012). Statistical Detection of Systematic Election Irregularities. *Proceedings of the National Academy of Sciences*, 109(41):16469–16473.

Knoke, D. and Yang, S. (2008). *Social Network Analysis*. Sage Publications, Los Angeles.

Krishna, A. (2008). *Poverty, Participation, and Democracy: A Global Perspective*. Cambridge University Press.

Kuenzi, M. and Lambright, G. M. (2007). Voter turnout in africa's multiparty regimes. *Comparative political studies*, 40(6):665–690.

Kuenzi, M. and Lambright, G. M. (2011). Who votes in africa? an examination of electoral participation in 10 african countries. *Party Politics*, 17(6):767–799.

Larreguy, H., Marshall, J., and Querubin, P. (2016). Parties, Brokers, and Voter Mobilization: How Turnout Buying Depends Upon the Party's Capacity to Monitor Brokers. *American Political Science Review*, 110(1):160–179.

Larson, J. M. and Lewis, J. I. (2017). Ethnic networks. *American Journal of Political Science*, 61(2):350–364.

Lee, M. M. and Zhang, N. (2017). Legibility and the Informational Foundations of State Capacity. *Journal of Politics*, 79(1):118–132.

Leighley, J. E. and Nagler, J. (2013). *Who votes now?: Demographics, issues, inequality, and turnout in the United States*. Princeton University Press.

Mvukiyehe, E. and Samii, C. (2017). Promoting democracy in fragile states: Field experimental evidence from liberia. *World Development*, 95(C):254–267.

Nagler, J. (1991). The effect of registration laws and education on us voter turnout. *The American Political Science Review*, pages 1393–1405.

Nevitte, N., Blais, A., Gidengil, E., and Nadeau, R. (2009). Socioeconomic status and nonvoting: A cross-national comparative analysis. *The comparative study of electoral systems*, pages 85–108.

Nichter, S. (2008). Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot. *American political science review*, 102(1):19–31.

Nickerson, D. W. (2008). Is Voting Contagious? Evidence from Two Field Experiments. *American Political Science Review*, 102(1):49–57–57.

Radcliff, B. (1996). Turnout and the economy: A comment on jackman and miller's "voter turnout in the industrial democracies in the 1980s". *Comparative Political Studies*, 29(6):718–723.

Robinson, A. L. and Gottlieb, J. (2021). How to Close the Gender Gap in Political Participation: Lessons from Matrilineal Societies in Africa. *British Journal of Political Science*, 51(1):68–92.

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

Rueda, M. R. (2015). Buying Votes with Imperfect Local Knowledge and a Secret Ballot. *Journal of Theoretical Politics*, 27(3):428–456.

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

Stokes, S. C. (2005). Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina. *American Political Science Review*, 99(3):315–325.

Stokes, S. C., Dunning, T., Nazareno, M., and Brusco, V. (2013). *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge University Press.

Thompson, D., Wu, J., Yoder, J., and Hall, A. (2020). Universal vote-by-mail has no impact on partisan turnout or vote share. *Proceedings of the National Academy of Sciences*, 17(25):14052–14056.

Watch, H. R. (2021). Uganda: End enforced disappearances of opponents.

ONLINE APPENDIX

— Supporting Information for "Who Registers to Vote?" —

A Village Networks

A.1 A glossary of network concepts

A network (or *graph*), is a collection of nodes and of ties between these nodes. The graph $g = (G, N)$, whereby N is the set of nodes, G is the set of ties, and a tie is a pair $(i, j), i, j \in N$. Networks can also be represented by an $(N \times N)$ *adjacency matrix* m, where $m_{ij} = 1$ if there is a tie from i to j, and $m_{ij} = 0$ otherwise. The *size* of g is the number of nodes.

A graph can be *directed* or *undirected*. In the former case, there is a distinction between a tie from i to j and a tie from j to i. That is, we do not require that m is symmetric. In the latter case, there is no distinction, and we require that m is symmetric. In what follows, we define the network concepts used in the paper in the case of an undirected network.

- *Neighbor*: j is a neighbor (or a peer) of i if they are connected; that is, if $(i, j) \in G$. The neighborhood of i is the set of i's neighbors.
- *Degree*: the degree d_i of i is the number of neighbors i has. That is, $d_i = \sum_{j \neq i} m_{ij}$.
- *Isolate*: i is an isolate if it has a degree of 0.
- *Density*: captures the amount of ties in g, relative to its size. A network of size n has $T_g = n(n-1)/2$ ties. Let $t_g = \sum_{i<j} m_{ij}$ be the amount of ties in g. The density of g is $D_g = t_g / T_g$.
- *Clustering coefficient*: the extent to which the friends of i are friends with each other. Formally, it is the amount of triangles in i's neighborhood normalized by the amount of triangles in i's neighborhood. It writes $c_i = \sum_j \sum_k m_{ij} m_{ik} m_{jk} / \sum_j \sum_k m_{ij} m_{ik}$, with $i \neq j, i \neq k, j < k$.

- *Path*: a path between i and j is a route from i to j on the graph g. Formally, it is a sequence of ties $(i_1,i_2),(i_2,i_3),...,(i_{K-1},i_K)$ such that $(i_k,i_{k+1}) \in G$ for each $k \in \{1,...,K-1\}$, with $i_1 = i, i_K = j$, and each node in the sequence $i_1,...,i_K$ is distinct.

- *Connected* graph: a graph is connected if there is a path between any $i,j \in N$

- *Path length*: the number of steps it takes to get from i to j on some path. Formally, the length of path $p = (i_1,i_2),(i_2,i_3),...,(i_{K-1},i_K)$ is $K-1$.

- *Distance*: the distance l_{ij} between i and j is the length of the shortest path between i and j.

- *Closeness centrality*: how close is node i from the rest of the graph? The closeness centrality of i is the mean distance between i and all other nodes of the graphs. It writes $L_i = \sum_{j \neq i} l_{ij}/(N-1)$. The concept is not well-defined when the graph g is not connected.

- *Betweenness centrality*: how much do people have to go through node i? Betweenness centrality is, for any $j,k \neq i$, the amount of shortest paths that go through i. The concept is not well-defined when the graph g is not connected.

A.2 Village selection

An important question is whether the 16 villages we sampled are representative of Ugandan villages. To answer the question, we leverage census data to examine where sampled villages lie in the distribution of several important village-level characteristics. The results, reported in Figure SI-1 show that for all examined dimension, more than half sampled villages lie within the 20 and 80th percentile of the distribution. The dimensions for which sampled villages are least representative are education and age, with sampled villages being slightly better educated and older than the population.

A.3 Network Construction

First, we provide verbatim excerpts from our in-person survey used to construct adjacency matrices capturing within-village network ties.

"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, 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."

Second, we report in Figure SI-2 the degree distribution across the four types of networks, as well as in the union network. Finally, since these networks are constructed using a name generator, top coding (i.e. naming the maximum number of respondents allowed by the name generator) may be an issue, because it may artificially truncate the degree distribution. In Table SI-1, we report, for each of our four networks, the percentage of respondents that reported the maximum number of 5 alters. The table shows that the prevalence of top

coding is low: it affects about one quarter respondents in the friendship network, and less than one fifth in the family network, and about one tenth in the remaining two networks.

Network	Pct. top coding
family	0.269
friend	0.142
lender	0.056
solver	0.056

Table SI-1: **Top coding in network ties elicitation.** This table report the percentage of respondents that reported 5 alters by type of ties. Difference in means are reported in the Δ column, with standard errors clustered at the village level; *p<0.1; **p<0.05; ***p<0.01. Top coding (i.e. reporting 5 alters) has a relatively low prevalence. It affects about one quarter respondents in the friendship network, and less than one fifth in all other networks.

A.4 Handling missing network data

Understandably, we were unable to interview every single individual in the village. This means there are villagers for whom we only observe a fraction of their network: they were mentioned as ties by other respondents, but were not interviewed in-person. About 18% of named individuals fall in this category. Following standard practice, we exclude those nodes from the analysis. Table SI-2 reports the number of individuals we surveyed in each village, the number of individuals mentioned by at least one person ("alters"), and the number of adults living in each village, according to 2014 census data. This information allows calculating the number of missing nodes.

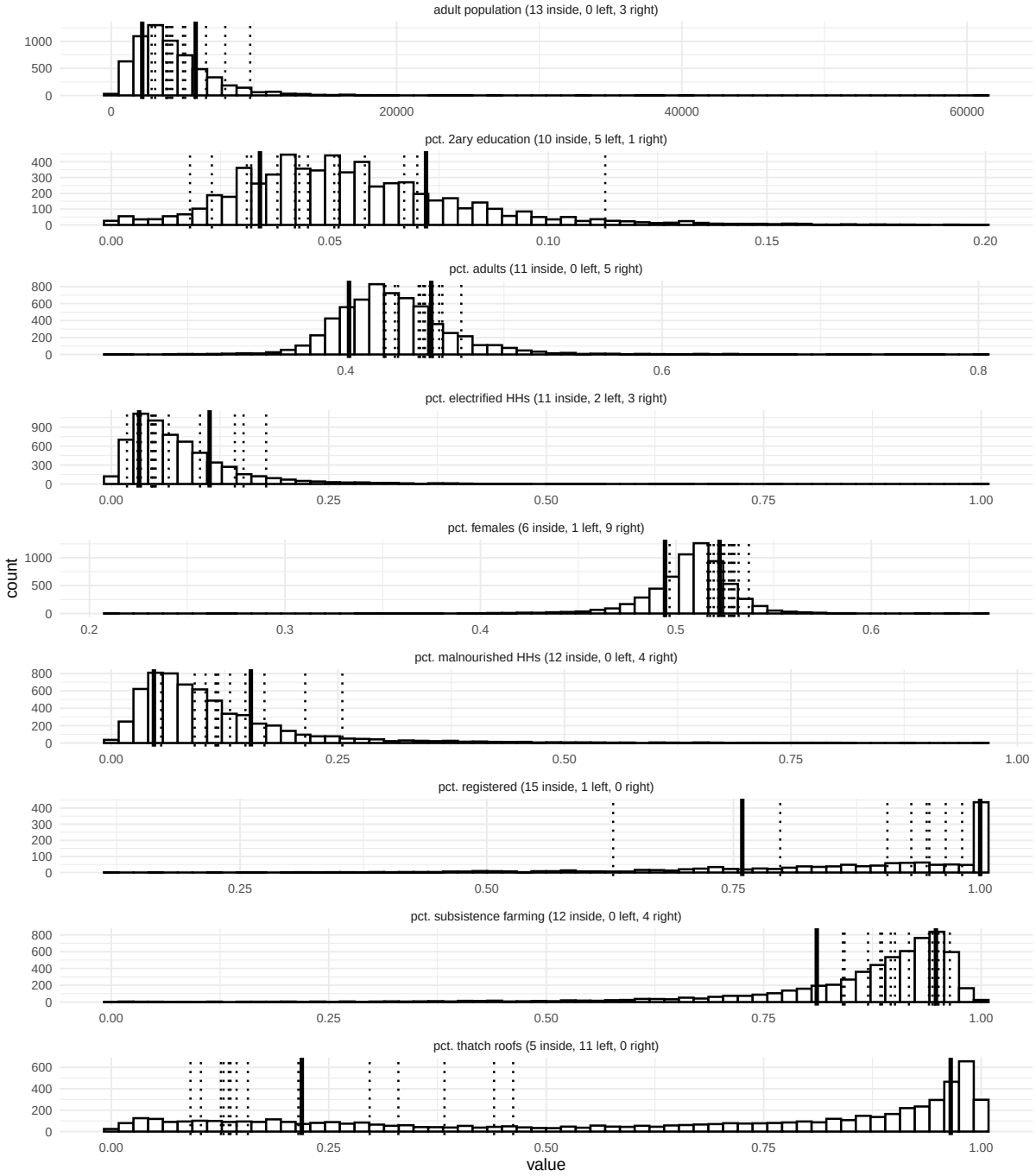

Figure SI-1: **Representativeness of sampled villages.** The thick lines represent the 20 and 80th percentiles of the distribution of each variable in all of Uganda (except for registration, where we report only the Arua province). The dotted lines represent the value for each sampled village. Panel titles report, in parenthesis, the number of sampled villages that fall (1) within the 20 and 80th percentiles ("inside"), (2) below the 20th percentile ("left"), and (3) above the 80th percentile ("right"). Sampled villages are slightly more female and less likely to have thatch roofs than the population.

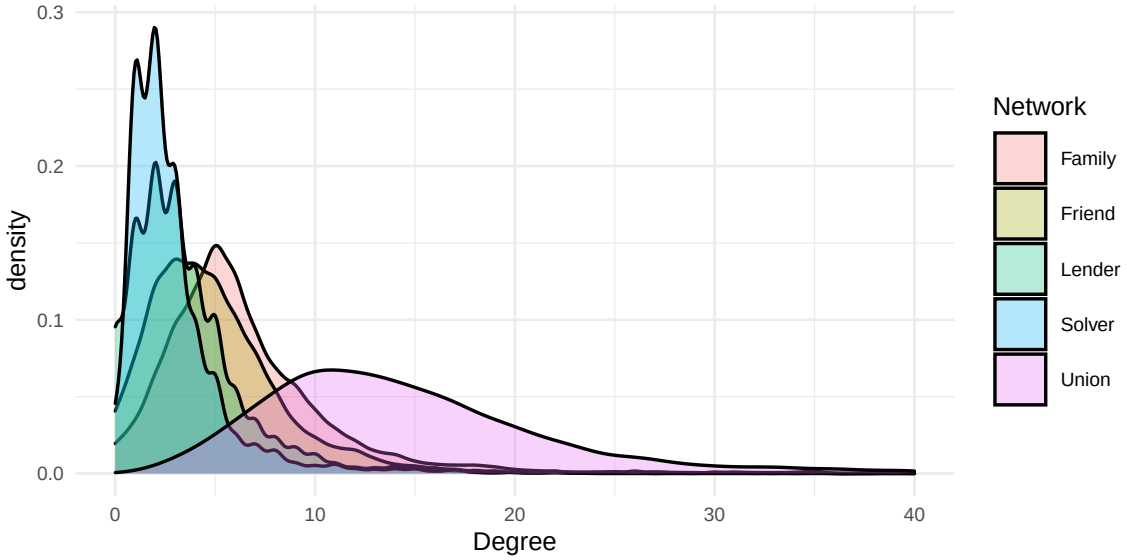

Figure SI-2: Degree distribution by network type.

Village	N interviewed	N alters	Adult population	Pct. non-interviewed alters	Pct. non-interviewed population	In-degree interviewed	In-degree non-interviewed	Δ
A	30	41	31	0.27	0.03	18.73	2.45	16.28^{***}
B	205	296	258	0.31	0.21	107.05	2.74	104.31^{***}
C	192	276	233	0.30	0.18	100.41	4.21	96.19^{***}
D	229	307	242	0.25	0.05	119.27	3.19	116.08^{***}
E	160	216	161	0.26	0.01	84.35	3.86	80.49^{***}
F	237	325	295	0.27	0.20	121.77	2.05	119.72^{***}
G	163	212	203	0.23	0.20	85.61	4.35	81.26^{***}
H	185	264	285	0.30	0.35	98.02	3.8	94.22^{***}
I	254	321	315	0.21	0.19	132.22	3.64	128.57^{***}
J	197	279	274	0.29	0.28	102.19	3.34	98.85^{***}
K	225	291	198	0.23	-0.14	117.21	3.47	113.74^{***}
L	168	306	230	0.45	0.27	88.49	3.01	85.48^{***}
M	189	262	183	0.28	-0.03	98.99	3.58	95.41^{***}
N	283	372	320	0.24	0.12	147.25	5.73	141.52^{***}
O	204	281	266	0.27	0.23	106.72	3.92	102.79^{***}
P	263	381	385	0.31	0.32	136.41	4.74	131.68^{***}
All	3184	4430	3879	0.28	0.18	111.97	3.68	108.3^{***}

Table SI-2: **Network sampling.** N alters reports the number of individuals mentioned as alters in the network survey. Adult population from 2014 census data. The Δ column reports the difference in mean in-degree between interviewed and non-interviewed individuals. Standard errors are heteroskedastic-robust for within-village differences and clustered by village for across-village differences; *p<0.1; **p<0.05; ***p<0.01.

Variable	Sample mean	Census mean	Δ
% 2ary education	0.23	0.15	0.08**
% female	0.58	0.54	0.04***
age	37.39	36.78	0.61

Table SI-3: **Comparing sample to census.** Sample mean and census mean for gender, age and secondary education broken down by village and for all villages. The sample over-represents females and educated individuals. Standard errors clustered at the village level; *p<0.1; **p<0.05; ***p<0.01.

Using census data, we are able to examine whether we are particularly likely to miss certain types of people, which could have implications for the interpretation of our results. We find that our sample somewhat over-represents females and educated individuals (Table SI-3), and also that the in-degree of those we did not interview is lower than those we interviewed (Table SI-2). However, Table SI-4 shows that females and educated individuals do not respond differently to peer effects, suggesting that sampling does not introduce bias in our estimation of peer effects.

	Dependent variable: registered			
	(1)	(2)	(3)	(4)
female	0.032	0.014	0.014	0.038
	(0.037)	(0.029)	(0.030)	(0.037)
age	0.020***	0.020***	0.018***	0.018***
	(0.005)	(0.005)	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
income	0.002	0.001	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.006)
2ary education	−0.010	0.021	−0.013	0.025
	(0.065)	(0.025)	(0.060)	(0.023)
Catholic	0.035	0.035	0.037	0.037
	(0.042)	(0.043)	(0.042)	(0.043)
Lugbara	−0.015	−0.016	−0.016	−0.015
	(0.050)	(0.050)	(0.050)	(0.050)
leader	0.036	0.035	0.036	0.037
	(0.024)	(0.025)	(0.024)	(0.025)
pro-sociality	−0.012	−0.012	−0.012	−0.013
	(0.046)	(0.045)	(0.045)	(0.046)
participation index	0.029	0.028	0.028	0.028
	(0.018)	(0.017)	(0.018)	(0.017)
distance (km)	0.014	0.014	0.014	0.014
	(0.016)	(0.016)	(0.016)	(0.016)
geography	−0.005	−0.006	−0.006	−0.006
	(0.008)	(0.008)	(0.008)	(0.008)
household head	0.098***	0.097***	0.096***	0.095***
	(0.027)	(0.027)	(0.028)	(0.028)
degree HH	0.008	0.008	0.008	0.008
	(0.005)	(0.006)	(0.005)	(0.005)
% registered	0.157	0.153	0.047	0.077**
	(0.122)	(0.125)	(0.039)	(0.036)
age × % registered	−0.003	−0.002		
	(0.003)	(0.003)		
secondary education × % registered	0.044		0.053	
	(0.070)		(0.066)	
female × % registered	−0.027			−0.034
	(0.031)			(0.028)
Observations	1,940	1,940	1,940	1,940
R^2	0.070	0.070	0.069	0.069
Note:				*p<0.1; **p<0.05; ***p<0.01

Table SI-4: **Sensitivity to peer-effects for sub-populations.** This table reproduces model (5) from Table 2 in the main text but adds interaction terms for age, education, and gender. Neither age, education, nor gender affects sensitivity to peer effects.

B Additional Tables and Figures

In this section, we provide additional tables and figures for robustness and completeness of the empirical analysis. In Figure SI-3, we show the bivariate relation between registration status individual-level covariates and in Table SI-5, we show the correlation matrix of our independent variables.

In Table SI-6, Table SI-7 and Table SI-8, we demonstrate the robustness of our main results from Table 2. Specifically, we show that our key findings are insensitive to using absolute threshold instead of relative threshold models, to using an alternative measure of education attainment, and to alternative measures of network centrality, respectively. In Table SI-9, we show that the results from Table 4 (comparison between various network ties) are robust to adding controls for household size and within-household peer effects. Table SI-10 complements Table 4 (model 1), by adding specifications that consider both lending ties that overlap with family and lending ties that exclude family ties.

Finally, in Table SI-11, we compare those individuals that are close and far from the LC1.

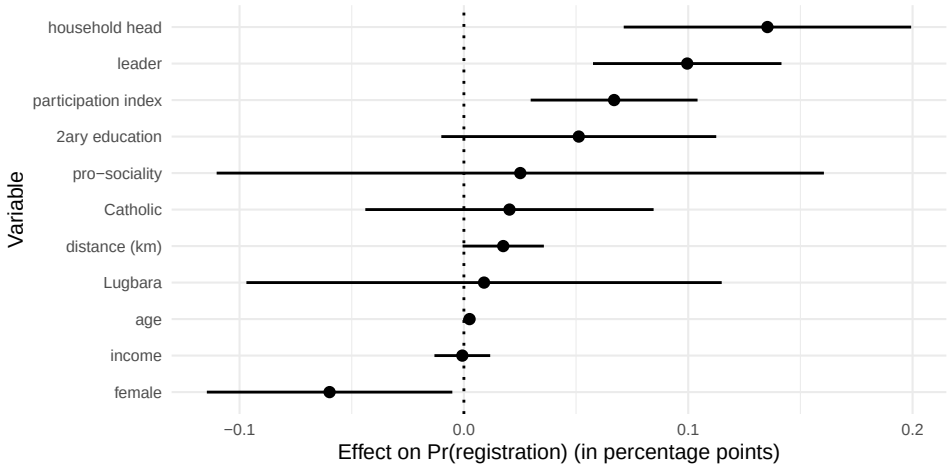

Figure SI-3: **Bivariate relationships with registration.** This figure reports the coefficients of separate regressions models that each regress registration status on the the column variable. All models include village-level fixed effects. Bars represent 95% confidence intervals derived from standard errors that are clustered at the village level.

	regis.	female	age	income	2ary edu.	no school.	Catho.	Lug.	leader	pro-soc.	pol. particip.	dist.	geo.
registered													
female	-0.03												
age	0.05**	-0.03											
income	-0.03	0.00	-0.19***										
2ary education	0.02	-0.25***	-0.10***	0.20***									
no schooling	-0.02	0.23***	0.34***	-0.17***	-0.22***								
Catholic	0.03	0.01	-0.03	-0.05**	-0.15***	0.12***							
Lugbara	0.03	0.01	0.05**	-0.01	-0.01	0.06**	0.01						
leader	0.06***	-0.14***	0.25***	0.00	0.02	0.00	0.03	0.03					
pro-sociality	0.01	-0.03	0.10***	0.04	0.07***	0.02	0.01	0.02	0.02				
participation index	0.06**	-0.24***	0.08***	0.11***	0.20***	-0.14***	-0.08***	0.00	0.17***	0.05*			
distance (km)	0.03	-0.04*	-0.05**	0.01	0.05**	-0.03	-0.05**	0.01	-0.03	0.03	0.00		
geography	0.07***	-0.01	-0.01	-0.02	-0.04*	-0.02	-0.04*	0.07***	0.00	0.03	0.07***	-0.23***	
household head	0.11***	-0.41***	0.31***	-0.11***	0.09***	0.00	-0.02	-0.03	0.15***	0.02	0.14***	-0.08***	0.01

Table SI-5: **Bivariate Correlations between DV and main control variables.** Lack of correlation between registration and gender, income, education, being a member of the majority religion (Catholic), being a member of the majority ethnic group (Lugbara) in the multivariate models (Table 2) does not owe to multicollinearity. *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: registered		
	All	Head	Non-Head
	(1)	(2)	(3)
female	0.013	0.028	0.006
	(0.030)	(0.037)	(0.035)
age	0.018***	−0.0001	0.028***
	(0.005)	(0.007)	(0.006)
age^2	−0.0002***	−0.00002	−0.0003***
	(0.0001)	(0.0001)	(0.0001)
income	0.001	0.027**	−0.010
	(0.006)	(0.012)	(0.009)
2ary education	0.023	0.014	0.050
	(0.024)	(0.032)	(0.041)
Catholic	0.036	0.038	0.021
	(0.044)	(0.061)	(0.044)
Lugbara	−0.016	0.038	−0.021
	(0.051)	(0.067)	(0.083)
leader	0.035	0.008	0.096***
	(0.024)	(0.034)	(0.036)
pro-sociality	−0.014	0.015	−0.040
	(0.047)	(0.056)	(0.082)
participation index	0.028	0.004	0.041*
	(0.018)	(0.029)	(0.024)
distance (km)	0.015	−0.010	0.031
	(0.017)	(0.014)	(0.021)
geography	−0.005	−0.013	0.0002
	(0.008)	(0.010)	(0.009)
household head	0.091***		
	(0.029)		
degree HH	−0.001	−0.015	0.001
	(0.013)	(0.023)	(0.015)
N registered HH peers	0.015	0.060**	0.014
	(0.024)	(0.027)	(0.027)
Observations	1,940	785	1,155
R^2	0.067	0.041	0.095
Note:			*p<0.1; **p<0.05; ***p<0.01

Table SI-6: **Absolute threshold models.** This table reproduces models 4, 5, 6 in Table 2, but uses absolute threshold models instead of relative threshold models. The fit of those relative threshold models is comparable to that of absolute threshold models.

	Dependent variable: registered			
	All	All (2+)	Head	Non-Head
	(1)	(2)	(3)	(4)
female	0.024	0.007	0.043	−0.002
	(0.026)	(0.027)	(0.039)	(0.034)
age	0.015***	0.019***	0.002	0.028***
	(0.003)	(0.005)	(0.006)	(0.006)
age^2	−0.0002***	−0.0002***	−0.00004	−0.0003***
	(0.00004)	(0.0001)	(0.0001)	(0.0001)
income	−0.007	0.002	0.026**	−0.010
	(0.006)	(0.006)	(0.012)	(0.009)
Catholic	0.035	0.037	0.045	0.021
	(0.041)	(0.044)	(0.059)	(0.044)
Lugbara	0.080*	−0.018	0.029	−0.025
	(0.046)	(0.050)	(0.067)	(0.081)
leader	0.049*	0.034	0.006	0.093**
	(0.025)	(0.025)	(0.033)	(0.037)
pro-sociality	−0.011	−0.014	0.002	−0.041
	(0.055)	(0.049)	(0.056)	(0.084)
participation index	0.018	0.026	0.001	0.039
	(0.015)	(0.017)	(0.028)	(0.024)
distance (km)	0.011	0.019	−0.006	0.036*
	(0.012)	(0.016)	(0.014)	(0.020)
geography	−0.006	−0.003	−0.010	0.003
	(0.007)	(0.008)	(0.010)	(0.008)
household head	0.075***	0.083***		
	(0.023)	(0.027)		
some 1ary edu	0.049*	0.002	0.089	−0.023
	(0.025)	(0.035)	(0.081)	(0.034)
some 2ary edu	0.077*	0.044	0.132	0.036
	(0.040)	(0.049)	(0.097)	(0.060)
some 3ary edu	0.045	−0.007	0.044	0.016
	(0.045)	(0.045)	(0.084)	(0.056)
Observations	2,753	1,940	785	1,155
R^2	0.050	0.066	0.039	0.094
Note:				*p<0.1; **p<0.05; ***p<0.01

Table SI-7: **Education.** This table reproduces models 1 to 4 of Table 2, but uses a categorical variable for education. The reference category is "No schooling."

	Dependent variable: registered			
	(1)	(2)	(3)	(4)
female	0.020	0.022	0.021	0.034
	(0.027)	(0.027)	(0.027)	(0.027)
age	0.016***	0.015***	0.015***	0.014***
	(0.005)	(0.005)	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
income	−0.001	−0.001	−0.001	−0.004
	(0.006)	(0.006)	(0.006)	(0.007)
2ary education	0.013	0.005	0.013	0.001
	(0.024)	(0.023)	(0.023)	(0.022)
Catholic	0.042	0.042	0.042	0.037
	(0.043)	(0.041)	(0.044)	(0.040)
Lugbara	−0.014	−0.011	−0.013	−0.003
	(0.047)	(0.046)	(0.046)	(0.044)
leader	−0.006	−0.009	−0.006	−0.013
	(0.027)	(0.028)	(0.028)	(0.027)
pro-sociality	−0.003	−0.004	−0.008	0.001
	(0.043)	(0.044)	(0.044)	(0.046)
participation index	0.011	0.007	0.010	0.006
	(0.016)	(0.016)	(0.016)	(0.016)
distance (km)	0.013	0.014	0.013	0.013
	(0.015)	(0.015)	(0.015)	(0.016)
geography	−0.008	−0.009	−0.009	−0.012*
	(0.007)	(0.007)	(0.007)	(0.007)
household head	0.067**	0.064**	0.065**	0.059**
	(0.027)	(0.028)	(0.027)	(0.028)
degree	0.004***	0.003***	0.003***	−0.003
	(0.001)	(0.001)	(0.001)	(0.002)
closeness		2.511***		
		(0.735)		
clustering			−0.236***	
			(0.075)	
eigenvector				1.040***
				(0.262)
Observations	1,940	1,940	1,940	1,940
R^2	0.079	0.083	0.082	0.089
Note:				*p<0.1; **p<0.05; ***p<0.01

Table SI-8: **Network covariates**. Note: all network covariates but degree are normalized to fall between 0 and 1.

	Dependent variable: registered				
	Non-HH	Family	Friends	Lender	Solver
	(1)	(2)	(3)	(4)	(5)
female	0.025	0.036	0.019	0.019	0.015
	(0.029)	(0.027)	(0.027)	(0.028)	(0.030)
age	0.016***	0.015***	0.010**	0.015***	0.018***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)
age^2	−0.0002***	−0.0002***	−0.0001**	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.00004)	(0.0001)	(0.0001)
income	−0.001	−0.002	−0.003	−0.003	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2ary education	0.012	0.018	0.016	0.020	0.021
	(0.023)	(0.021)	(0.023)	(0.023)	(0.024)
Catholic	0.041	0.041	0.044	0.041	0.038
	(0.041)	(0.042)	(0.040)	(0.041)	(0.041)
Lugbara	−0.013	−0.012	−0.016	0.003	−0.015
	(0.047)	(0.044)	(0.047)	(0.045)	(0.050)
leader	−0.003	−0.004	−0.014	0.006	0.027
	(0.028)	(0.025)	(0.026)	(0.025)	(0.026)
pro-sociality	−0.002	0.001	0.012	−0.003	−0.010
	(0.041)	(0.039)	(0.035)	(0.041)	(0.045)
participation index	0.014	0.012	0.004	0.013	0.024
	(0.016)	(0.017)	(0.016)	(0.016)	(0.017)
distance (km)	0.010	0.007	0.002	0.008	0.013
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)
geography	−0.011	−0.013	−0.018**	−0.011	−0.006
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
household head	0.078***	0.065**	0.058**	0.079***	0.092***
	(0.028)	(0.027)	(0.029)	(0.028)	(0.029)
degree	0.004***	0.015***	0.033***	0.014***	0.001*
	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)
% registered peers	−0.024	0.043	0.044	0.080***	0.034
	(0.069)	(0.032)	(0.035)	(0.025)	(0.048)
degree HH	0.008*	0.009**	0.010**	0.009*	0.008
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
% registered HH peers	0.048	0.043	0.050	0.051	0.058
	(0.037)	(0.036)	(0.037)	(0.035)	(0.036)
Observations	1,940	1,940	1,940	1,940	1,940
R^2	0.081	0.088	0.117	0.085	0.070
Note:					*p<0.1; **p<0.05; ***p<0.01

Table SI-9: **Comparison with other peer effects.** This table reproduces Table 4, but controls for household size and within-household peer effects.

	Dependent variable: registered			
	Both	Lender & Family	Lender & not Family	Lender & Solver & not Family
	(1)	(2)	(3)	(4)
female	0.013	0.011	0.010	0.009
	(0.028)	(0.029)	(0.027)	(0.029)
age	0.015***	0.017***	0.015***	0.018***
	(0.005)	(0.005)	(0.005)	(0.005)
age^2	−0.0002***	−0.0002***	−0.0002***	−0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
income	−0.002	−0.0001	−0.002	0.002
	(0.006)	(0.006)	(0.006)	(0.007)
2ary education	0.021	0.022	0.023	0.024
	(0.024)	(0.024)	(0.024)	(0.024)
Catholic	0.042	0.040	0.040	0.038
	(0.042)	(0.043)	(0.043)	(0.044)
Lugbara	0.002	−0.013	0.006	−0.014
	(0.046)	(0.048)	(0.046)	(0.048)
leader	0.003	0.021	0.009	0.028
	(0.025)	(0.025)	(0.025)	(0.026)
pro-sociality	−0.006	−0.009	−0.004	−0.014
	(0.044)	(0.045)	(0.046)	(0.048)
participation index	0.010	0.018	0.012	0.023
	(0.017)	(0.018)	(0.016)	(0.018)
distance (km)	0.013	0.015	0.013	0.018
	(0.016)	(0.016)	(0.015)	(0.015)
geography	−0.008	−0.005	−0.007	−0.003
	(0.008)	(0.008)	(0.007)	(0.007)
household head	0.065**	0.071***	0.069**	0.082***
	(0.028)	(0.027)	(0.027)	(0.027)
degree	0.014***	0.015***	0.017***	0.022***
	(0.002)	(0.004)	(0.002)	(0.005)
% registered peers	0.079***	0.037**	0.086***	−0.051*
	(0.026)	(0.019)	(0.021)	(0.030)
Observations	1,940	1,940	1,940	1,940
R^2	0.082	0.072	0.083	0.067
Note:				*p<0.1; **p<0.05; ***p<0.01

Table SI-10: **Lenders.** This table reproduces model 4 of Table 4, and adds specifications that consider only lending ties that overlap with family ties (model 2), and only lending ties that do not overlap with family ties (model 3).

	Lending network			Strong ties network		
Variable	Close	Far	Δ	Close	Far	Δ
Dependent variable						
% registered	0.721	0.618	0.098***	0.721	0.605	0.106***
Individual						
% HH head	0.424	0.341	0.087***	0.446	0.249	0.21***
% female	0.499	0.605	-0.102***	0.486	0.668	-0.185***
age	36.710	37.970	-0.769	37.615	34.865	3.182***
income	2.596	2.553	0.088	2.563	2.667	-0.088
% 2ary education	0.233	0.239	0.009	0.244	0.203	0.043
% Catholic	0.580	0.581	-0.019	0.570	0.619	-0.03
% Lugbara	0.945	0.950	0.002	0.946	0.947	0.002
pro-sociality	0.145	0.169	-0.034	0.147	0.167	-0.02
leader	0.200	0.204	-0.005	0.204	0.190	0.012
participation index	0.078	-0.088	0.161***	0.071	-0.088	0.149***
distance to registration center (km)	1.024	0.996	-0.083	1.040	0.936	0.014
Network						
degree centrality	18.119	16.616	0.605	17.731	17.816	-0.654
closeness centrality	0.002	0.002	0**	0.002	0.002	0**
clustering coefficient	0.365	0.388	-0.03**	0.371	0.371	-0.006
eigenvector centrality	0.178	0.143	0.029***	0.173	0.156	0.015**
Sample size						
N	1647.000	537.000		1711.000	473.000	

Table SI-11: **Comparing individuals that are far & close to the LC1.** Columns 'Close' and 'Far' report descriptive statistics for all members of households of size 2 and above, broken down by whether those members are at most 3 steps away from the LC1 ('Close') or at least 4 steps away from the LC1 ('Far'), either in the directed lending network or the strong ties network. Columns Δ report the difference between columns 'Close' and 'Far'. Standard errors are clustered at the village level. Individuals that are close to the LC1 are significantly more likely to be household heads. *p<0.1; **p<0.05; ***p<0.01

C Lending networks

It is worth noting that, while consistent with the idea of brokers' mobilization, our findings are somewhat more nuanced owing to the diffused nature of lending networks, common in Uganda and elsewhere in East Africa. In Table SI-12 we contextualize the lender networks

in our study area. Importantly, it is not the case that villagers necessarily all name the same small set of 'fixers'. We identify, for each respondent i, the most popular lender—i.e. the lender mentioned by most respondents—among all lenders mentioned by individual i. If all respondents borrowed from the same small set of 'fixers', there would be a lot of redundancy, and the exercise would yield a small number of lenders. Instead we find that this exercise yields 1,600 different individuals. Note also that lenders, while diffused, are on average more likely to be registered to vote (79%) than the average villager.

population	N	pct. registered	mean connectedness
most popular lenders	16	1.00	0.14
LC lenders	54	0.85	0.05
most popular lenders to ego	1600	0.79	0.02
all lenders	1875	0.78	0.02
all individuals	2184	0.70	0.01

Table SI-12: **Lenders' Descriptive Information.** Note that while our regression analysis above was limited to those residing in households of size of 2+ (given our interest in household dynamics), we consider all villagers in the study area, irrespective of their household size, when computing popularity as lenders, and their associated registration rates. LC stands for Local Councilor chairperson at village, parish or the subcounty level.