BROKERING VOTES WITH INFORMATION SPREAD VIA SOCIAL NETWORKS*

RAÚL DUARTE
Harvard University

FREDERICO FINAN
UC Berkeley

HORACIO LARREGUY
ITAM

LAURA SCHECHTER
UW Madison

February 2025

Politicians rely on political brokers to buy votes throughout much of the developing world. We investigate how social networks facilitate these vote-buying exchanges. Our conceptual framework suggests brokers should be particularly well-placed within the network to learn about non-copartisans' reciprocity in order to target transfers effectively. As a result, parties should recruit brokers who are central among non-copartisans. We combine village network data from brokers and citizens with broker reports of vote buying, allowing us to use broker and citizen fixed effects. We show that networks diffuse information about citizens to brokers who leverage it to target transfers. In particular, among those citizens who are not registered to their party, brokers target reciprocal citizens about whom they can learn more through their network, and these citizens are more likely to support the brokers' party. Moreover, recruited brokers are significantly more central than other citizens among non-copartisans, but not among copartisans. These results highlight the importance of information diffusion through social networks for vote buying, broker recruitment, and ultimately for political outcomes.

JEL Codes: D72, O12

Keywords: vote buying, brokers, social networks

*Duarte: rduarte@g.harvard.edu ; Finan: ffinan@econ.berkeley.edu ; Larreguy: horacio.larreguy@itam.mx ; Schechter: lschechter@wisc.edu . We are grateful to Arun Chandrasekhar, Cesi Cruz, Joseph Cummins, Cynthia Kinnan, Julien Labonne, John Marshall, Pablo Querubín, David Yanagizawa-Drott, and seminar participants at ETH, Harvard, ITAM, Jinan, Maryland, MIT, NEUDC, New Economic School, Notre Dame, Nottingham, Pittsburgh, Princeton, PUC Chile, Texas A&M, UC Irvine, UCLA, UC Riverside, UCSD, UW Madison, Warwick, Yale, Zurich, and the 8th European Meeting on Networks for useful comments. Larreguy acknowledges IAST funding from the French National Research Agency (ANR) under grant ANR-17-EURE0010 (Investissements d'Avenir program). Schechter thanks the Wisconsin Alumni Research Foundation, the Russell Sage Foundation, and the Delta Foundation for funding.

1 Introduction

The secret ballot is a cornerstone of fair and free elections. It enables citizens to vote without intimidation or fear of reprisal and is considered an essential check against vote buying (Baland and Robinson, 2008; Robinson and Verdier, 2013). Despite the almost universal adoption of the secret ballot, vote buying remains pervasive throughout the developing world, and its persistence is not yet well understood.[1]

Political brokers play an important role in sustaining vote buying (Finan and Schechter, 2012; Larreguy, 2013; Lehoucq, 2007; Schaffer, 2007; Stokes, 2005). Political parties commonly use brokers as intermediaries to exchange targeted benefits for votes. These brokers, who are often community leaders with extensive local knowledge, are thought to exploit their social connections to facilitate vote-buying exchanges. They are believed to leverage social networks to buy votes using various complementary tactics, including a) monitoring of voter electoral behavior (Cruz, 2019; Stokes, 2005), b) targeting of citizens who are more likely to reciprocate at the polls (Finan and Schechter, 2012; Lawson and Greene, 2014), c) targeting of copartisans who are unlikely to vote (Nichter, 2008), and d) targeting of opinion leaders (Auerbach and Thachil, 2020; Cox, 2015; Schaffer and Baker, 2015). But whether and how brokers leverage their social networks to sustain vote buying, and how this in turn shapes broker recruitment as a function of broker location in these networks, has yet to be assessed empirically. This gap is due in large part to the lack of data on both the social networks and the vote-buying decisions of the political brokers themselves.

We overcome these data limitations using a novel survey of political brokers and citizens in rural Paraguay. With these data, we estimate the extent to which the network positions of brokers and citizens predict the brokers' knowledge about citizens and the brokers' targeting decisions. We compute a network statistic at the broker-citizen level called *hearing*. *Hearing* measures how much information a specific broker would expect to hear about a specific citizen given their locations in the social network and the network's structure, according to the information diffusion model of Banerjee et al. (2013). *Hearing* takes into account all the paths, both direct and indirect, that information might take between a citizen and a broker and the decay it would experience along the way, and so is more relevant to information sharing than other broker-citizen-level network measures such as social distance, which only captures information diffusion via the shortest direct

[1]While reliable data are hard to come by, the share of the population offered a vote-buying transfer and the amount offered is high across multiple countries. One-third of respondents in the Philippines were offered between \$1 and \$30 in the 2010 elections (Cruz, 2019), 40% of respondents were offered a quarter of the minimum monthly wage during Uganda's 2016 general election (Blattman et al., 2024), and 34% of respondents in Mexico were offered something in the 2018 general elections (Montes, 2018).

path between a citizen and a broker.

While other studies have measured the social networks of citizens or candidates, this is, to our knowledge, the first study to systematically measure the social networks of political brokers and citizens. This is a significant improvement over the existing literature because vote buying is ultimately based on dyadic relationships. In contrast to previous studies, our network measure captures the relative network positions of brokers and citizens, rather than only measuring the position of one type of actor. This feature not only allows us to estimate the information-diffusion role of networks in facilitating vote-buying exchanges but also offers insights into the importance of network position for who becomes a political broker, a subject about which little is known and to which this research contributes.

There are several key features of our survey data. First, we elicit information about citizens from multiple brokers. This allows us to control for both broker and citizen fixed effects, eliminating broker- and citizen-specific confounders. With broker fixed effects we exploit only within-broker variation in vote buying – and thus control for broker-specific confounders such as their position in the network or willingness to admit to vote buying. Including citizen fixed effects further restricts our analysis to within-citizen variation, accounting for citizen-specific confounders such as the citizen's socioeconomic status and network position. Second, in contrast to the extant literature, we measure vote buying as reported by the brokers rather than by the citizens.[2] This approach is essential for addressing important confounding factors such as some forms of misreporting, omitted variable bias, and network endogeneity, which is a key contribution of this research. Third, because we also ask several of the same questions about citizen characteristics to citizens, we can test how much information about citizens is diffused through the network to brokers.

Another key feature of our data is an experimental measure of intrinsic reciprocity collected from an incentivized anonymous game years before the election we are studying. An intrinsically reciprocal person finds pleasure in increasing the material payoffs of someone who has helped her, whether or not it affects the present value of her own material payoffs. Likewise, she will enjoy decreasing the material payoffs of someone who has harmed her. This is contrasted with instrumental reciprocity, which is motivated by forward-looking self-interest (Sobel, 2005).

Our analysis delivers four main findings consistent with a simple vote-buying and broker-recruitment model. First, our broker-citizen measure of *hearing* strongly predicts how well brokers know their fellow villagers, particularly characteristics important for targeting vote-buying

[2] One notable exception is Ravanilla et al. (2021), who collected broker-reported measures of vote buying during the 2016 Philippine elections.

exchanges. Second, *hearing* significantly predicts broker targeting and whether a citizen claims to support the broker's party thereafter. Our preferred estimates imply that a one standard deviation increase in *hearing* accounts for approximately 0.32 standard deviations of the vote-buying index. Third, using our experimental measure of reciprocity and publicly available party registration data, we find that brokers are more likely to target reciprocal citizens not registered to their party but about whose reciprocity level they can more easily learn due to the architecture of the social network. In contrast, brokers target citizens registered to their party regardless of their reciprocity level or their position in the network. Fourth, consistent with these targeting patterns, we find that parties recruit brokers who have higher levels of *hearing* among non-copartisans relative to other villagers. Average broker diffusion centrality (the sum of the broker's *hearing* with all citizens) is 20 percentiles greater than that of other households. However, their diffusion centrality is significantly higher (40 percentiles) among non-copartisans, but no different among copartisans.

While our data and experimental design have many advantages, there are also some limitations that we address as best we can. First, our network data is incomplete as it comes from a sample rather than a census, which can lead to bias. We argue that the incomplete network data will lead to attenuation both by reviewing the econometric literature and by re-running our regressions with data from villages with different sampling rates.

Second, we do not have random variation in network structure, and network formation is endogenous. The inclusion of broker- and citizen-level fixed effects helps account for some network endogeneity (for example if unobserved citizen heterogeneity impacts both network formation and the outcome variable). Network endogeneity at the broker-citizen dyad level remains an issue. We note that our broker-citizen measure of *hearing* is based on links between all citizens in the network, not just links involving the broker. This implies that it is more difficult for a broker to endogenously increase his level of *hearing* with a citizen because he can not force others to link with one another. We show that our results are robust to i) controlling for social distance, which exploits more direct (and perhaps more endogenous) links than does our measure of *hearing*, ii) controlling for direct links between the broker and citizen, iii) excluding information spread directly to the broker from the citizen himself when calculating *hearing*, and iv) recomputing the networks and *hearing* excluding ties resulting from non-political financial ties, both monetary and in-kind, which one could argue are more endogenous.

Third, the data on vote buying is self-reported by the brokers. This is a substantive improvement over the existing literature relying on citizen reports, with different citizens potentially being more or less willing to report vote buying. By measuring vote buying as reported by brokers at the broker-citizen level, we can use fixed effects to control for reporting bias that is specific to the broker and the citizen. Still, one might worry that brokers report differently depending on the person about whom they are reporting. First, we show that our estimates are robust to controlling for link-specific covariates. Second, we show that results are similar with a secondary outcome of whether the citizen claims to support the broker's party, which does not rely on broker reporting.

This study contributes to various literatures. First, it contributes to research on the role of social networks in explaining vote buying and electoral outcomes. Some papers focus on citizens' social networks and show that well-connected citizens are more likely to be targeted (Cruz, 2019; Fafchamps and Labonne, 2020; Ravanilla et al., 2021). Other papers focus on candidates' or brokers' social networks and show that more central brokers are better able to influence vote choice (Szwarcberg, 2012), or that more central candidates get more votes (Cruz et al., 2017). In contrast with previous studies that focus on the network position of one individual, we construct the complete network and create a broker-citizen-level measure of the connection between the two. This allows us to study targeting using fixed effects at both the broker and citizen levels to control for important broker- and citizen-level confounders that have plagued previous studies. Moreover, we show that brokers' centrality with respect to non-copartisans matters for broker recruitment.[3]

Second, a growing body of literature studies the effects of social networks on political outcomes other than vote buying. Social networks affect perceptions and voting behavior through the dissemination of information about unemployment (Alt et al., 2022), electoral violence (Fafchamps and Vicente, 2013), and elections in general (Fafchamps et al., 2020). Arias et al. (2019), Bond et al. (2012), Collier and Vicente (2014), and Enríquez et al. (2024) further show that social networks coordinate the electoral behavior of individuals around information campaigns. Our findings similarly highlight the importance of social networks in diffusing information from citizens to brokers.

Third, there is a rich literature on the determinants of citizen targeting. Some suggest politicians target citizens with weak ideological attachment (Dixit and Londregan, 1996; Lindbeck and Weibull, 1987), while others argue they target core supporters (Cox and McCubbins, 1986; Nichter, 2008). More recent work highlights the targeting of reciprocal citizens, supporters unlikely to turn out, and opinion formers (Auerbach and Thachil, 2020; Cox, 2015; Finan and Schechter, 2012; Lawson and Greene, 2014; Nichter, 2008; Schaffer and Baker, 2015). The importance of brokers for deciding who to target has been discussed by Larreguy et al. (2016) and Stokes et al. (2013).

We build off Finan and Schechter (2012) who show that brokers target reciprocal citizens. Our study extends Finan and Schechter (2012) in at least three significant ways. First, we show that brokers learn which citizens are reciprocal from information diffused through the social network.

[3]This might explain why brokers are not more central than the average citizen in a context characterized by turnout buying (Brierley and Nathan, 2021).

Second, we show that the effect in Finan and Schechter (2012) comes specifically from brokers targeting reciprocal non-copartisans about whom they can more easily learn information due to the network architecture. Third, we study brokers' placement in their social networks and show that parties recruit brokers who are most central among citizens not registered with their party.

Finally, we contribute more generally to a growing literature on social networks' role in sustaining informal transactions. Networks play a key role in various settings (Chuang and Schechter, 2015; Jackson, 2014; Munshi, 2014) particularly for their role in information diffusion (Alatas et al., 2016; Alt et al., 2022; Banerjee et al., 2013, 2019), social learning (Chandrasekhar et al., 2020; Conley and Udry, 2010), and transaction enforcement (Bloch et al., 2008; Chandrasekhar et al., 2018; Jackson et al., 2012; Schechter and Yuskavage, 2012). Our study provides further evidence of the ability of social networks to facilitate informal and, in this case, illicit transactions.

We structure the remainder of our paper as follows. In section 2, we provide background information on political brokers and vote buying in Paraguay. In section 3, we present the conceptual framework and the predictions that we take to the data. In section 4, we describe the data and the construction of the citizen-broker network measures. In section 5, we describe our empirical strategy, and in section 6 we present our main results, robustness checks, and tests for alternative mechanisms. Section 7 concludes.

2 Background

Paraguay was under the dictatorship of Alfredo Stroessner of the Colorado party from 1954 to 1989. Until 2008, when independent bishop Fernando Lugo won the presidency, the Colorado party had controlled the national government for sixty-one years. Paraguay is effectively a two-party system. The Colorado and Liberal parties are by far the strongest, although smaller parties have recently gained modest popularity. As a result of the 2006 municipal elections – the elections we study – half of the villages in our sample elected a Colorado mayor and the other half a Liberal mayor. The brokers in our sample are also evenly distributed between the two strongest parties.[4]

Paraguay has 17 departments, broken into 238 municipalities. Each municipality consists of an urban town and approximately twenty rural villages. Each municipality has a mayor, the 2006 election of whom is the election we focus on. Citizens choose whether to register to vote and whether to affiliate with a party, all of which is publicly available information. Because polling stations are in central locations, often in urban towns, and contain citizens from multiple villages,

[4]Among the smaller parties, the National Union of Ethical Citizens (UNACE), which was founded from a faction of the Colorado party, also has one broker who operates in a village in our sample. We present descriptive statistics about the brokers in our sample in section 4.

candidates and brokers cannot know how a village voted in aggregate.

Political parties in Paraguay are not strongly ideological, and there is little policy differentiation between them (Lachi and Rojas-Schaffer, 2018; Parks, 2018). Political campaigns tend to be highly personalized, and vote buying is thought to be an effective electoral strategy (New York Times, 2023; Paraguay, 2018). This was evident in the focus groups conducted with politicians and brokers by Lachi with *Transparencia Paraguay* in 2005 and by Lachi and Rojas-Schaffer in 2018. For example, a broker of the Liberal party in the municipality of General Morínigo commented that "elections in Paraguay are decided by the voters who are mobilized with money. A very small percentage of the voters are loyal. The incentivized voters define [the election]." Vote buying is also becoming increasingly important to win elections, as a broker of the Colorado party from the municipality of General Aquino explained: "there are three groups of voters: the captive, the thinkers, and those that can be bought. Relative to previous elections the captive voters have declined, and the voters that can be bought have increased."

Political brokers, who in Paraguay are known as *operadores políticos*, act as intermediaries between candidates and citizens, exchanging money and favors for promises to vote accordingly. Their ability to facilitate these exchanges is due, in large part, to how embedded they are within their community. As a politician of the Liberal Party in the municipal council of San Lorenzo noted, "political brokers are fundamental since they know their zone well." The mayor of Coronel Oviedo describes brokers as people "who know the neighborhood, who accompany the candidate and show him the people and the neighborhood. Since candidates cannot know everything, these contact people are important for candidates." When asked how brokers learn about citizens, a Colorado broker in the municipality of General Aquino noted that "it is all about ñe'embegue (gossip)." Lachi (2009) concludes from his series of focus groups that political brokers "are essential for electoral campaigns and that their value is directly proportional to their integration within their communities." More details about the selection and role of brokers as described in the focus groups conducted by Transparency Paraguay are included in the supplemental materials of Finan and Schechter (2012). Political parties typically have multiple brokers operating within the same village. According to ethnographic work of Dosek (2019), brokers of the same party typically coordinate on which citizens to target.

Brokers leverage their local knowledge to target citizens. A Liberal party official in Asunción mentioned that brokers "know who [their] party supporters are." Similarly, a Liberal official in the governor's office of Coronel Oviedo argued that brokers know "which Colorado and Liberal voters would sell their vote." Importantly, brokers suggest that the citizens who they target are likely to reciprocate with their vote. For example, a Liberal broker in the municipality of General Morínigo mentioned that, "while some voters take the money and vote for another candidate, the number of voters like that is small." A Colorado broker in the municipality of General Aquino further indicated that the citizens they target "always thank favors."

To win elections, brokers need to target non-copartisan citizens in addition to copartisan core supporters. For example, Liberal brokers from the municipality of General Morínigo recognize that their party "helps" non-copartisans more than their own copartisans. They confirm that elections are won with votes from non-copartisans, which is why they target citizens beyond their own party.

In terms of compensation, while some brokers are incentivized with the prospect of public employment, the focus groups make clear that there are not enough public employment jobs for all the brokers and so brokers need to be motivated with more immediate financial compensation. The focus groups state that brokers receive regular wages for two to four months around election time. Monthly salaries are said to range between \$50 and \$100. For comparison, a day's wages in agriculture in our villages in this period was a bit less than \$2 per day. Brokers are also given money and gifts to give to the citizens. Due to the impossibility of having signed receipts, brokers do not have to account for how they spend the money they receive from candidates.

3 Conceptual Framework

Building on the qualitative evidence presented in Section 2, we argue that brokers focus on mobilizing copartisans regardless of what they learn about them through their networks, but they specifically target non-copartisans whom they identify through their networks as likely to be reciprocal. Consequently, candidates recruit brokers who are central among non-copartisans. For a more detailed understanding, we formalize this argument in a stylized vote-buying model in Appendix A.

The idea behind our argument is as follows. The job of a broker is to deliver votes. Brokers have an incentive to target citizens who they are confident will vote for their party in a secret ballot. Brokers may engage in both turnout buying and vote buying. Turnout buying involves compensating a citizen for the citizen's cost of voting. This sort of transfer targets citizens who support the broker's party but would not turn out to vote without a transfer. Vote buying involves targeting citizens who would normally have voted for the other party but, with a transfer, would vote for the broker's party. The anecdotal evidence presented above and in previous literature emphasizes that these citizens are of a reciprocal type.

As one can see even from this simple setup, a broker's knowledge about his fellow villagers is critical for effectively delivering votes to his party. Some information, such as turnout and party registration, are public information in Paraguay as in many other settings. Other information, such as reciprocity, must be learned through social networks. We assume, and subsequently test, that a broker's knowledge about a specific citizen is determined by the number of times a broker hears information about the citizen, which is a function of their relative positions in the village network.

Given this environment, first, a broker mobilizes his copartisans with turnout-buying transfers. If citizens support the broker's party but face a high cost of voting, then the broker may want to compensate them for their voting costs to induce them to vote. In this case, the broker does not need to identify reciprocal citizens because turning out to vote is observable and easily contractible.

Among non-copartisans, a broker targets reciprocal citizens with vote-buying transfers. Reciprocity is crucial in this process. When a broker targets a non-copartisan, he is essentially compensating the citizen for voting against her own interests. If a citizen is not reciprocal, she will not feel obligated to vote for the broker's party. However, reciprocity is not a publicly known trait, and the broker must learn about it through his social network. As a result, parties have the incentive to select brokers who are more centrally located, particularly among non-copartisans. These brokers are uniquely positioned to learn about the reciprocity of non-copartisans and target them accordingly. In contrast, whether brokers are central or peripheral relative to copartisans is irrelevant since the enforcement of turnout buying does not rely on citizen reciprocity.

Underlying the discussion are two important assumptions. First, we assume for simplicity that when deciding whom to target, brokers do not consider the targeting decisions of brokers from other parties. This assumption would be reasonable if the targeting budget of the incumbent party relative to the opposition were sufficiently large that the incumbent could act as a monopolist (Baland and Robinson, 2008), or if the political marketplace is sufficiently segmented where parties compete over the recruitment of brokers who can deliver blocks of votes within their neighborhoods.[5] If, however, one were to generalize our model to include competition across brokers of different parties, then brokers may have an incentive to target co-partisans who they believe the other party brokers believe are reciprocal. We show, however, empirically that brokers do not in fact incorporate the *hearing* of brokers from different parties in their targeting decisions and that they rarely target the same citizen.

Second, we assume that a broker's network position is essential for learning about citizen reciprocity, but that party registration and turnout are public information since that is the case in our context, as in many others. If party registration and turnout were not easily observable (either officially or unofficially), then one would have to extend the learning model to include partisanship and turnout. If that were the case, turnout buying would become less easily contractible and brokers would then have an incentive to target reciprocal copartisans.

[5]In this case, our model can be viewed as a sub-game to a more extensive model of party competition over brokers.

Predictions

Our framework implies several predictions that we can test with the data. First, the expected number of times a broker hears about a citizen, as determined by their relative network positions, predicts how well the broker knows the citizen. Second, neither a citizen's reciprocity level nor the expected number of times the broker hears about the citizen affects the targeting of copartisans. Third, conditional on brokers targeting non-copartisans, brokers will target reciprocal non-copartisans with whom they are connected in the network in such a way that they can hear more about them. Fourth, conditional on brokers targeting non-copartisans, the party will recruit brokers who are more centrally located among non-copartisans so that, on average, they hear more about them.

In order to test these predictions, we need data on i) the number of times a broker is expected to hear about a citizen as determined by their relative network positions, ii) the citizen's party registration, iii) the citizen's level of reciprocity, and iv) the broker's targeting decisions. The latter three variables are directly available in our data. For the first variable, one must construct a model of information diffusion in a network and apply it to network data. We measure the extent to which a broker is expected to hear about a citizen using *hearing* as defined by Banerjee et al.'s (2013) model of information diffusion. We describe all of these variables in more detail in Section 4.

4 Data

We combine vote-buying data from brokers and citizens originally collected for Finan and Schechter (2012), with social network data originally collected for Ligon and Schechter (2012). Our sample consists of brokers and citizens from ten villages, each village chosen randomly from a different municipality, across two departments. Households were selected to be surveyed in 1991 with subsequent rounds collected in 1994, 1999, 2002, and 2007. In 2002, incentivized experiments were also conducted. In 2007, more households were added so that at least 30 households were surveyed in each village. The 2007 survey also included a section with questions about the 2006 Paraguayan municipal elections, which was conducted with the same household member who participated in the 2002 experiments whenever possible. We refer to the household member who answered this section of the 2007 survey as the 'citizen' throughout the paper.

In 2010, we returned to the villages and asked the households in our original sample to identify the brokers working in their villages.[6] At this point, the villagers knew us well. We had conducted

[6]The original survey was conducted in three departments, while the 2010 follow-up was only conducted in the two more easily accessible departments. Because the original survey was intended to study land reform and property

incentivized experiments with them multiple times, and in both 2002 and 2007 they had interacted with one co-author who is a Guaraní-speaking American.[7] Many of the enumerators had conducted multiple rounds of the survey, and over the decades the villagers had learned that there were no negative consequences from talking to us and that we were to be trusted. We then felt comfortable inquiring about the brokers operating in the villages in which we worked.

Households identified 43 brokers working in the ten villages, and we were successful in interviewing 38 of them. Four of the interviewed brokers did not live in the villages in which they worked, and so we do not know how they fit in the village social network. It turned out that 20 of the interviewed brokers were members of households that were part of our panel data sample, and thus, we had directly surveyed their social connections in 2007. Other directly surveyed households mentioned an additional 12 brokers as social ties, and so we have indirect information about these brokers' social connections. Thus, we are only missing social network data for two of the surveyed brokers living in the villages in which we worked.

The 38 brokers in the study are mostly male, with only two females. Twenty of the brokers are from the dominant Colorado party, while 17 are from the opposition Liberal party. One is from the relatively small UNACE party. Twenty-two of the brokers in our study worked for the party of the incumbent mayor. They are all farmers, as is almost everybody who lives in the villages. They are, on average, 51 years old, with eight years of education. The average among the citizens in our sample is 50 years old and five years of education, so the brokers are slightly more educated. The brokers have lived in the villages for a long time, with 24 having lived there all their life, and the other 14 living there an average of 20 years (with a minimum of 6 years). They have worked as a broker between 5 and 48 years, with an average of 18 years.[8]

Network measures

Our data include social networks collected in 2007 from 10 villages. In each village, between 30 and 48 households were surveyed, delivering direct sampling rates ranging between 12 and 91% (with a cross-village mean of 47%). If we also consider non-surveyed households who were mentioned as a social connection by at least one survey respondent, then we have network information on between 54 and 100% of households in each village (with a cross-village mean of 88%). All dyads are coded as either connected or unconnected; no relationships are coded as missing.

rights, farmers owning more than 25 hectares of land were over-sampled in 1991.

[7] Guaraní is the indigenous language of Paraguay.

[8] For comparison, for the sub-sample of citizens whose residence history we have, 54% have lived in the village all their life, and the others have lived there an average of 25 years with a minimum of 3. Thus, it is not obvious that middlemen have lived in the villages longer than the average citizen in our sample.

Social connections include family ties, godparenting ties, support networks, and non-political financial – monetary or in-kind – transactions in the last year, including informal transfers, gifts, and loans.[9] We include all types of social connections since restricting them to a particular type would lead to fewer interpersonal connections and disconnected sets of nodes. Appendix Figure C1 provides an example of the social network of the households in one of the villages in our data, representing 257 connections between 81 households (of which 39 were directly surveyed). It is interesting to look at the network placement of the two brokers in that village. Both stand out visually as being quite central. One resides in the household labeled 9, was directly surveyed, and is linked with 14 households. The other resides in the household labeled 73, and although it was indirectly surveyed, it is linked with 10 households (meaning it was mentioned by 10 of the 39 directly surveyed households).

We measure the extent to which broker b hears information about citizen i using *hearing* as defined in Banerjee et al. (2013). We define T as the number of periods that information flows in the network and p as the probability with which it flows in a given period between two directly connected households after one household received the information in the previous period. Define the adjacency matrix $\mathbf{g}$ as the matrix where each row and column represents one household in the village and an element equals 1 if the row and column households are connected and 0 if they are not. The expected number of times that broker b hears a piece of information originating from citizen i if information is diffused according to this process is the ib^{th} entry of the matrix $\mathbf{H} = \sum_{t=1}^{T} (p\mathbf{g})^t$, or H_{ib}, which we call *hearing*. The broker's diffusion centrality, H_b, is the sum of their *hearing* with all citizens, or the sum of the elements of column b.

For the intuition on how this measure is computed, consider the process of diffusion of a piece of information originating from citizen i to broker b as illustrated in Figure 1. Sub-figure (a) shows the social network and the information in period 0. Sub-figure (b) shows that in the first period, the two nodes directly connected to i (colored in gray) find out the information with probability $p \in (0,1]$. Because broker b cannot hear the information in the first period, $H_{ib}(1) = 0$. Sub-figure (c) shows that in the second period, those who received the information in the first period transmit it to the nodes with which they are directly connected (colored in gray) with probability p. In subsequent periods, those who received the information in the previous period transmit it to the nodes with which they are directly connected with probability np, where n is the number of times

[9] More specifically, we construct undirected village-level networks where two households are connected if a member of each household belongs to the same family (i.e., parents, children, siblings); a member of one household is the godparent of the child of someone in the other household; one household would go to the other for monetary assistance in times of need; if in the past year someone in the household provided monetary assistance when someone in the other household fell sick; or if in the past year someone in one household made a non-political monetary or in-kind transfer or lent money to someone in the other household.

they could have received the information in the previous period.[10]

As shown in sub-figure (d), broker b's first chance to hear information originating from citizen i is in period 3. We see that $H_{ib}(3) = 0.09$, which means that, in the third period, we expect the broker to have heard the information 0.09 times. By period 5, which is the last period shown, we expect the broker to have heard the information 0.23 times. This process lasts for T periods, where T is a finite positive integer, as information likely loses relevance with the passage of time.

To construct *hearing*, we must take a stand on two nuisance parameters, T and p. We set T equal to 7, the largest social distance between any citizen and broker in our sample.[11] If T were smaller than the largest social distance, information from some citizen would never reach some broker in their network. Following Banerjee et al. (2013), we set p equal to the inverse of the largest eigenvalue of the adjacency matrix for each village's social network.[12] In our data, this yields information transmission probabilities between 0.08 and 0.14.

As our example illustrates, *hearing* captures the relative network positions of brokers and citizens. Social distance is another broker-citizen-level network measure. It captures only the shortest path between a broker and a citizen, ignoring the less direct paths that information could take as well as the recycling of information that *hearing* takes into account. For this reason, *hearing* provides a better measure of a broker's knowledge about citizens when deciding whom to target.

Besides *hearing*, we construct additional network statistics to address concerns associated with homophily and other potential confounders. At the citizen level, these measures include the clustering coefficient, degree centrality, betweenness centrality, diffusion centrality, and eigenvector centrality. At the broker-citizen level, these measures include the existence and number of support pairs (direct connections in common), an indicator for actual and hypothetical non-political informal financial transaction ties, and social distance. Jackson et al. (2012) show that informal exchanges between a pair of individuals that are locally enforceable and renegotiation-proof require that the pair is "supported" by a common tie. Chandrasekhar et al. (2018) show that the social distance between two individuals explains their ability to sustain informal exchanges in the absence of contract enforcement. We describe these measures in more detail in Appendix B.

Table 1 presents summary statistics for the network measures. The values of these network

[10]Note that while the equation is the same, the description differs slightly from that given in Banerjee et al. (2013), as described in Bramoullé and Genicot (2020).

[11]Banerjee et al. (2013) suggest T equal the largest social distance between any two individuals in the network, which would be 10 in our case. Their logic is that T should be sufficiently large so that every individual in the network can hear something about another individual with positive probability. We adopted their logic to our object of interest, which is broker learning about citizens. We show that results are robust to different choices of both T and p.

[12]Choosing larger values of p leads to total diffusion as T increases, while choosing smaller values of p causes diffusion to die out as T increases. The inverse of the largest eigenvalue is the critical intermediate value between these two processes.

characteristics may be difficult to interpret in isolation. Compared with households in the Indian villages studied in Banerjee et al. (2013), the households in our data have lower average degree (number of direct connections) but higher values of all other network measures. Thus, these villages are more inter-connected than the villages studied in India.[13]

Brokers and the information they have

To test whether *hearing* predicts how well a broker knows a citizen, we complement our broker survey with the data collected from citizens. These data allow us to combine the answers to a series of questions about citizen characteristics that were asked to both brokers and citizens. Appendix B explains the coding of both the broker and citizen answers, as well as our criteria to categorize matching answers. We construct indices of brokers' knowledge about citizens along three dimensions and sum these three indices into an overall knowledge index.

The *covariates index* combines four indicators of a broker's general familiarity with the citizen. These are whether the broker states that he knows the citizen; the broker can correctly name the spouse of the citizen; the broker can accurately state the years of education of the citizen; and the broker can correctly state the amount of land the citizen owns. As Table 1 summarizes, brokers can identify the citizens and their spouses in 89% and 77% of cases, respectively. Similarly, they can correctly assess citizens' education and land 81% and 42% of the time, respectively.[14] The *political index* is an indicator for whether the broker correctly assesses the strength of the citizen's party preference. Brokers are accurate about how strongly citizens prefer a specific party 59% of the time.[15] Lastly, the *social preferences index* combines two variables that indicate a broker's knowledge about the social preferences of each citizen. These are indicators for whether the broker knows the extent to which the citizen would retaliate wrong-doing (59% match) and for whether the broker knows whether the citizen trusts at least half of those in their village (66% match).

Overall, brokers are very knowledgeable about citizens. Such knowledge is substantively greater than random guessing, as noted in Table II and Section 4.1 of Finan and Schechter (2012).[16]

[13]Some of the calculated network measures may have lower values in both the Paraguayan and Indian villages than in reality since their calculation comes from a sample rather than a census of households.

[14]We consider a mismatch to signal that the broker did not know the information. Of course, a mismatch could also imply that the broker knew the right answer, but the citizen misrepresented the truth.

[15]How strongly a citizen prefers a party differs from their official party registration, which is public information.

[16]This accurate assessment of fellow villagers by central people has also been documented in other settings. For example, Takasaki et al. (2000) find that village informants in the Peruvian Amazon can accurately state the physical and human capital of fellow villagers. Similarly, using network data from 631 Indonesian villages, Alatas et al. (2016) show that more connected people are better at ranking villagers in terms of their economic well-being. Alix-Garcia et al. (2021), in turn, find that household and informant-based asset indices are highly correlated, and their association does not vary systematically across informant characteristics.

Also, because we tested brokers' knowledge on randomly selected villagers, these results are likely a lower bound of the extent to which brokers know their own clients.

Political variables

We measure vote buying in the 2006 Paraguayan municipal elections as reported by 32 brokers in 2010. Each broker provided information for the same approximately 30 randomly chosen citizens in the village where they lived and worked. This yields 295 citizens, none of whom lived in the same household, and a total of 932 broker-citizen pair observations.[17]

Our main outcome variable is a broker-citizen-level standardized index that sums two measures of vote buying as reported by brokers about each citizen. These measures are indicators for whether a broker approached a citizen and whether he offered the citizen something during the electoral campaign. For the first variable, we asked the broker whether, during the political campaign before the 2006 municipal elections, he went to talk with the person about the campaign. These visits tend to involve the broker visiting the person in their home to discuss the candidate, to encourage the citizen to vote for the candidate, to give propaganda items (such as stickers, calendars, and hats) to the citizen, and to give non-propaganda vote-buying transfers to the citizen. For the second variable, we also asked the broker whether, on behalf of his candidate, he offered any non-propaganda items to that citizen. We mentioned that this could include giving foodstuffs, medicines, or money; paying their water or electricity bills; and having their fields plowed by a tractor. The correlation between these indicator variables is 0.35 ($p < 0.0001$).

While vote buying is illegal, the brokers discussed this information with us freely. Vote buying is a widely used and accepted electoral strategy as evidenced by the high participation rates in our survey and the focus group transcripts conducted by Transparency Paraguay. Very few (if any) politicians have been punished for vote buying. In addition, one of the coauthors who speaks Guaraní, has been working in these villages since the early 2000s, collecting multiple rounds of data often with the same enumerators. Over the decades, the villagers have understood that are no negative consequences from talking to us and that we are to be trusted, and so the brokers were willing to talk with us about the vote-buying transfers they engaged in.

An additional outcome is an indicator for whether the citizen reports supporting the broker's party in the 2007 survey – conducted a few months after the 2006 elections. With this outcome, we test the political effects of vote-buying exchanges as facilitated by the information-diffusion role

[17] We exclude 16 broker-citizen observations for which the broker and citizen were either the same person or lived in the same household as one another.

of networks.[18]

As Table 1 indicates, the average broker approached 48% of the citizens and offered something to 27% of them. Combining both vote-buying outcomes, brokers approached or offered something to 54% of citizens; that is 75% of citizens registered to their party and 38% of non-copartisans. Citizens support the same party as the broker in 46% of observations. Both parties are active in targeting, with the Colorados being slightly more active than the Liberals. At the same time, it is relatively rare (8% of the population) for both parties to target the same citizen.[19]

More generally, the ten municipalities we worked in are quite varied. The winning party, the competitiveness of the election, and the turnout rate all vary significantly. For example, of the ten municipalities in our study, the smallest winning vote margin in 2006 was 0.3% while the largest was 75.7%. Four municipalities had Colorado mayors in both 2001 and 2006, three had Liberal mayors in both 2001 and 2006, two had a Liberal incumbent but then a Colorado won in 2006, while one had a Colorado incumbent but then a Liberal won in 2006. We do not know vote shares of the specific villages in which we worked, since voting outcomes are at the larger polling station level. In the 26 polling stations where citizens from the 10 villages were registered to vote, Colorado party vote shares ranged from 36% to 90%, and Liberal party vote shares from 8% to 60%. In our sample, Colorado party support ranged from 40% to 83%, and Liberal party support from 7% to 50%.

Turnout of registered voters in the ten municipalities ranged from 42% to 61%. Turnout of the registered voters in the 26 polling stations where voters from the 10 villages in our sample cast their vote similarly ranged from 41% to 68%. In our sample, we have data indicating that at least 82% of the citizens are registered to vote. Of those who are registered to vote, 93% are affiliated with a party. The turnout rate for the citizens in our sample who are registered to vote was 68%.

Mediating variables

Beyond assessing how well brokers know citizens, we test whether citizen characteristics, and what brokers hear about citizens, predict targeting. In particular, we focus on citizens' partisanship and reciprocity, two characteristics that have been shown to explain citizen targeting. With respect to partisanship, we use official, publicly-available data on citizens' political affiliations to create

[18]Unfortunately, our data does not allow us to analyze whether the citizen voted for the broker's party because we do not know how he or she voted.

[19]Among individuals registered to the Colorado party, 45% only receive an offer from the Colorado party, 9% only from the Liberal party, 9% from both, and 37% from neither. Among individuals registered to the Liberal party, 34% only receive an offer from the Liberal party, 18% only from the Colorado party, 6% from both, and 42% from neither. Finally, among individuals registered to neither party, 20% only receive an offer from the Colorado party, 9% only from the Liberal party, 9% from both, and 62% from neither.

an indicator that the citizen is not registered to the same party for which the broker works. In our sample, the proportion of registered Colorados and Liberals is 59% and 30%, respectively.[20] In the subsequent analysis, we treat this measure as exogenous. We do this because party registration is persistent within families, with 90% of Paraguayans registering with the same party as at least one of their parents (Lachi and Rojas-Schaffer, 2018). In addition, citizens tend to affiliate with a party when they first register to vote and rarely change their affiliation thereafter.

As our second mediating variable, we use the experimental measure of citizen reciprocity developed by Finan and Schechter (2012). In 2002, a sub-sample of the citizens in our dataset participated in a trust game. The first mover was given 8,000 Gs (1,000 Gs were worth about 20 cents at that time) and had to decide whether to send nothing, 2,000, 4,000, 6,000, or 8,000 Gs to a second mover, who received the amount tripled. The second mover could keep all the money or return as much as she wanted. Before finding out how much money she would receive, the second mover had to outline a contingency plan (i.e., how much of 6,000 Gs, 12,000 Gs, 18,000 Gs, and 24,000 Gs she would return), which was implemented accordingly. All players played once as a first mover and once as a second mover.

A reciprocal person will increase the material payoffs of someone who has helped her, and decrease the material payoffs of someone who has harmed her. To separate altruism from reciprocity, Finan and Schechter (2012) calculate the reciprocity of a second mover by subtracting the share that she would return if she received 6,000 Gs from the average share that she would return if she received 12,000, 18,000, or 24,000 Gs, and they censor this measure below zero. So, if a second mover always returns half, no matter how much the first mover sends her, her reciprocity level will be zero. The more generous the second mover is when receiving a high transfer from the first mover compared to how stingy she is when receiving a low transfer, the higher her level of reciprocity. This measure is available for 85 citizens and 271 broker-citizen observations.

5 Empirical Strategy

We answer three main questions. First, do social networks diffuse information about citizens that brokers leverage to sustain vote buying? Second, what useful information do the brokers learn through the network that informs their targeting decisions? Third, do parties recruit brokers who, on average, have higher *hearing* among non-copartisans? For the first two questions, we use dyad-level data where each observation is a broker-citizen pair for the sample of all citizens

[20]In terms of party registration, our sample is fairly representative of the two departments where the villages are located. The proportion of registered Colorados in those two departments is 63% and 65%, while the proportion of registered Liberals is 29% and 34%.

whose households we interviewed. For the third question, we use individual-level data where each observation is a directly or indirectly surveyed household in the village.

To answer the first question regarding whether social networks diffuse information about citizens that brokers leverage to sustain vote buying, we estimate regressions of the following form:

$$y_{ib} = \alpha + \beta H_{ib} + \theta_i + \eta_b + X'_{ib}\delta + \varepsilon_{ib}, \tag{1}$$

where y_{ib} is an outcome defined for citizen i and broker b, H_{ib} is the *hearing* measure that captures the information-diffusion role of social networks, X_{ib} is a vector of observable dyad-level characteristics, θ_i is a citizen fixed effect, and η_b is a broker fixed effect. We use two-way clustering of our standard errors, clustering at both the broker and citizen levels.[21] Outcomes include how well the broker knows the citizen and whether the broker targeted the citizen.

Our ability to include both broker and citizen fixed effects when estimating Equation (1) is an important innovation over the existing vote-buying literature. By including broker fixed effects, we can control for unobserved broker-level determinants of a brokers' ability to engage in vote buying, such as their relative importance in the social network (Szwarcberg, 2012). Similarly, with citizen fixed effects, we can account for any unobserved citizen-level determinants of vote buying, such as the likelihood that the citizen turns out to vote (Nichter, 2008), the citizen's social preferences (Finan and Schechter, 2012; Lawson and Greene, 2014), and the citizen's relative importance in their social network (Schaffer and Baker, 2015). While previous studies have allowed for either citizen or candidate/broker fixed effects, this is the first study we know of to incorporate both.

Brokers and citizens with more extensive networks will naturally have higher *hearing*, so after we include fixed effects, *hearing* captures their relative network positions. For this variation to be considered exogenous, we need that $E[H_{ib}\varepsilon_{ib}|\theta_i, \eta_b, X_{ib}] = 0$. Depending on how the network is formed, this assumption may not hold. In Section 6, we discuss the conditions under which this assumption is reasonable and present some robustness checks in support of it.

To answer the second question regarding what useful information brokers learn through the network that they use to target citizens, we measure how *hearing*, party registration, and reciprocity differentially impact broker targeting. To do so, we estimate the following dyad-level regression:

$$y_{ib} = \alpha + \beta_1 H_{ib} + \beta_2 R_i + \beta_3 P_{ib} + \beta_4 H_{ib} R_i + \beta_5 H_{ib} P_{ib} + \beta_6 R_i P_{ib} + \beta_7 H_{ib} R_i P_{ib} + X'_{ib}\delta + \eta_b + \varepsilon_{ib}, \tag{2}$$

where R_i measures the citizen's level of reciprocity and P_{ib} measures whether the broker and citizen are not registered to the same political party as one another. In specifications in which we control

[21] Our final estimation sample contains 932 dyads representing 32 brokers and 295 citizens.

for citizen fixed effects, we drop the separate control for citizen reciprocity. We predict that $\beta_7 > 0$, or that for citizens who are not registered to their party, brokers are most likely to target those who are reciprocal and about whom they can learn that personal information. The framework also predicts that $\beta_4 = 0$. In other words, targeting decisions should not depend on the *hearing* and reciprocity levels of citizens registered to the same party as the broker.

Finally, to answer the third question regarding whether parties recruit brokers with higher *hearing* among non-copartisans, we use individual-level data to estimate the following regression:

$$y_{iv} = \alpha + \beta B_{iv} + \gamma D_{iv} + \eta_v + \varepsilon_{iv}, \tag{3}$$

where y_{iv} are individual measures of centrality, including degree centrality, betweenness centrality, eigenvector centrality, and diffusion centrality, all defined for citizen i in village v. The sample is all households in the village, whether or not we interviewed them. The most important outcome is diffusion centrality, the sum of the individual's *hearing* with all members of the village. A nice feature of diffusion centrality,[22] as seen in the description in Section 4, is that one can separately calculate how much an individual hears from every individual who is registered to his own political party and how much he hears from every individual who is not. Thus, we consider as separate outcomes diffusion centrality with those who are registered to the same party and diffusion centrality with those who are not. The main explanatory variable of interest, B_{iv}, measures whether the household contains a broker. We also control for village fixed effects and we control for D_{iv}, whether the household was directly surveyed (as opposed to our only knowing about their location in the social network through the reports of their fellow villagers). From this, we can test our prediction that parties recruit brokers who are more central than the average citizen, and are especially central with respect to citizens who are not registered to the same party as they are.

6 Results

We begin this section by showing that *hearing*, our network-based measure of information diffusion between a broker and a citizen, predicts how much a broker knows about a citizen. We then provide evidence that *hearing* is robustly associated with whether the broker targets the citizen for vote buying (our primary outcome of interest) and whether the citizen claims to support the party of the broker just after the election. Consistent with our theoretical framework, we find that brokers target citizens who are not registered to their party and who they hear are more reciprocal. The targeting of citizens registered to their own party is unrelated to what brokers hear about

[22]This also applies for degree but not eigenvector centrality or betweenness centrality.

them through the network. We conclude with a descriptive analysis of the broker location within their social networks. Again, consistent with our framework, we show that brokers are on average more centrally located than citizens, particularly among citizens who are not registered to their own party.

Relationship between *hearing* and broker knowledge

Do brokers know more information about citizens with whom they have higher *hearing*? Panel A of Table 2 reports the relationship between *hearing* and different measures of a broker's knowledge about each citizen. *Hearing* should predict broker information in all domains, not only political preferences and social preferences such as reciprocity. We explore as outcomes the indices measuring brokers' knowledge of citizens' demographics, political preferences, and social preferences, as well as an index aggregating all three indices.[23]

The coefficient on *hearing* is significant, positive, and robust across all outcomes and specifications reported. These results suggest that information diffusion through the social network facilitates the broker's acquisition of information about citizens. A one standard deviation increase in *hearing* is associated with a 0.21 standard deviation increase in overall knowledge (column (2)), a 0.17 standard deviation increase in knowledge of demographic characteristics (column (3)), a 0.20 standard deviation increase in politically-relevant information (column (4)), and a 0.14 standard deviation increase in knowledge about citizen social preferences (column (5)).

Relationship between *hearing* and vote buying and party support

Given the evidence that *hearing* predicts broker knowledge, we next explore the association between *hearing* and vote buying. The vote-buying index aggregates the indicators for whether the broker reported offering the citizen something during the electoral campaign and whether the broker reported approaching the citizen to talk about the electoral campaign. It is standardized to have mean 0 and standard deviation 1. Column (1) of Panel B of Table 2 includes broker fixed effects, and column (2) adds citizen fixed effects. Continuing with this more robust specification, in columns (3) and (4) we present estimates of the relationship between *hearing* and each component of the index.[24]

Throughout columns (1) to (4), *hearing* is positively associated with the vote-buying index and its constituent elements. Our preferred specification in column (2) implies that a one standard

[23]In Appendix Table C1, we present results on the constituent components of the indices.

[24]Appendix Table D1 shows that results in Table 2 are robust to List et al.'s (2019) multiple hypothesis adjustments.

deviation increase in *hearing* accounts for a 0.32 standard deviation increase in the vote-buying index. These results suggest that the diffusion of information through social networks plays an important role in determining which citizens the brokers target with vote buying.

If social networks enable brokers to target citizens effectively, then the citizens who brokers hear more about should also be more likely to support that broker's party near the time of the election. We test this in column (5). *Hearing* is positively associated with the citizen reporting that she supports the broker's party. A one standard deviation increase in *hearing* is associated with a 12 percentage point (28%) increase in the likelihood that the citizen supports the broker's party. While we do not know which party the citizen votes for, this result suggests that vote buying is effective at persuading citizens to support the broker's party.

Whether a broker targets a citizen might depend not only on how much he hears about the citizen, but also on how much any other brokers from his party hear about the citizen. In addition, it may depend on whether he expects the other brokers from his own party to target the citizen and whether he expects competing brokers to target the citizen. We explore the possibility of strategic interactions between brokers of the same party and of different parties by assessing whether the *hearing* of and targeting by other brokers within the same village predicts the targeting of a citizen.

Table 3 runs regressions similar to those in column (2) of Panel B of Table 2, but also controls for the standardized mean *hearing* of brokers from the same party and the standardized mean *hearing* of brokers from other parties in column (1). In column (2), more descriptively, we instead control for the standardized mean vote-buying targeting index of brokers from the same party and of brokers from other parties. Column (2) is not cleanly identified, as other brokers' vote buying is an endogenous outcome included as a control variable, but the results are suggestive.

Column (1) of Table 3 shows that *hearing* of brokers from the same party, but not from other parties, is positively associated with vote buying. This implies that brokers from the same party share information with one another. In contrast, brokers from opposing parties either do not share information with one another or do not take into account the knowledge they have about the information that brokers from opposing parties have. Additionally, the results in column (2) suggest that when other brokers (either from the same party or other parties) target a citizen, a broker is less likely to target the same citizen. While brokers from the same party share information, which helps ensure vote-buying transfers are targeted to the right citizens, they also make sure not to duplicate their efforts and both make transfers to the same person. Brokers from opposing parties are even less likely to target the same citizens, which is consistent with the summary statistics we presented in Section 2. Overall, this pattern suggests that if cross-party brokers do compete, they do so in a way that segments the vote-buying market.

Robustness Checks

In this section, we address identification and measurement concerns related to our principal finding that brokers are more likely to target individuals about whom they receive more information through the network. We start by discussing the forms of network endogeneity that might affect our estimate and provide a series of robustness tests. Next, we discuss and test for concerns related to the measurement of vote buying and *hearing*.

Identification Concerns

To help frame the discussion, let's consider a simple econometric model of network formation (Graham, 2017). The decision to form a link, D_{ij}, between individuals i and j depends on the utility individual i receives from linking with j, and vice versa. Under the assumption of transferable utility, individuals i and j form a link if the total surplus from doing so is positive:

$$D_{ij} = \mathbb{1}(g(X_{ij}, a_i, a_j) - U_{ij} \geq 0) \quad (4)$$

where X_{ij} represents a set of observable dyadic attributes (e.g., homophily), a_i and a_j denote unobserved attributes of individuals i and j, and U_{ij} represents an idiosyncratic random shock. The function $g(\cdot)$ rules out network effects because it only depends on the characteristics of i and j.

A common parameterization for $g(\cdot)$ assumes the function is additive in the unobserved attributes:[25]

$$g(X_{ij}, a_i, a_j) = X_{ij}'\delta + a_i + a_j.$$

Under this specification, citizen and broker fixed effects control for any endogeneity associated with the formation of the network. Alternatively, we might assume the following specification:

$$g(X_{ij}, a_i, a_j) = X_{ij}'\delta + a_i + a_j + a_i a_j, \quad (5)$$

which allows for an unobserved match quality effect. In this case, network endogeneity may be an important concern since the error term ε_{ib} in Equation (1) likely contains this unobserved link component, $a_i a_b$.

To rule out this potential form of endogeneity, we perform a series of robustness checks presented in Table 4. As we specified in Equation (5), individuals with similar traits may choose to form links, a concept known as homophily. If brokers are more likely to target citizens who share

[25] See for example Auerbach (2022), Graham (2017), Goldsmith-Pinkham and Imbens (2013), and Johnsson and Moon (2021).

similar traits and are also more likely to hear information about them because they are part of their network, then homophily could be a source of bias. In column (1), we test whether results are robust to allowing for homophily by including several dyad-level controls that capture similarity in age, gender, party registration, geographic proximity, and network centrality. The addition of these broker-citizen controls leaves the original point estimate essentially unchanged, even though the R^2 of the regression increases by 20 percent.

Another test of network endogeneity is to examine whether the estimate is robust to controlling for the direct link D_{ib} between the broker and the citizen. By conditioning on D_{ib}, we partial out its effect on *hearing*, and use identifying variation from indirect links, which are arguably more exogenous. As we can see from column (2), even when we control for D_{ib} directly, the estimate of the main relationship maintains economic and statistical significance.

Instead of controlling for D_{ib}, which is arguably a bad control, one could instead drop the broker's direct link with the citizen when constructing that dyad's *hearing*. One concern with this approach, however, is that it induces more sparsity in our networks which, as we will discuss below, can attenuate the coefficient and increase the standard errors. In column (3), we instrument the original *hearing* measure, with *hearing* that is calculated excluding the direct link between that broker and citizen. The resulting coefficient estimate remains consistent with the original estimate.

In columns (4) and (5), we control for the social distance between the broker and the citizen (i.e., the fewest number of links that separate i from b) both linearly and non-parametrically. Controlling for social distance is important for two reasons. First, *hearing* captures information flows throughout the entire network and not simply along the shortest path. By including both *hearing* and social distance, we can test which one matters more for vote-buying targeting. Second, after controlling for social distance, we are exploiting variation along indirect links between a broker and citizen, rather than just the shortest path. This variation is likely to be less prone to biases due to endogenous network formation by the broker because the broker has less control over the connections that other citizens make with one another. Adding those controls changes the coefficient very little.

Social networks are useful for both spreading information and enforcing agreements. The ability of brokers and citizens to use their social networks to enforce informal transactions, such as vote buying, might confound our results. In columns (6) and (7), we control for the number of friends that the broker and citizen have in common, both linearly and non-parametrically. The number of shared friends has been used as a proxy for the network's capacity to enforce contracts and commitments. The results from these columns suggest that our findings remain robust even after accounting for the role networks play in contract enforcement. Because enforcement might be facilitated by the spread of information about bad behavior (Breza et al., 2019), it is not possible to fully separate information diffusion from enforcement, and so these results are suggestive.

We next test the robustness of our results to non-political financial transaction ties. The social networks are a combination of two types of connections: long-term ties, such as family and godparent relationships; and transaction ties, such as actual non-political monetary and in-kind transactions in the year before the survey, including loans and gifts, and hypothetical assistance such that one household would go to the other in times of need. In column (8), we recalculate *hearing* in the network based only on the first type of connection - family and godparent relationships. These links are the least endogenous to political vote-buying targeting. Limiting the network to fewer link types introduces measurement error into the network, likely biasing our estimate downward. Nevertheless, even under this alternative network structure, the point estimates remain economically and statistically significant, though slightly attenuated. In column (9), we control for an indicator for the second type of connection – transaction ties and the results remain robust.

Relatedly, it is worth noting that village social networks are not centered on party lines. We test for sorting along party lines using the Freeman segregation index (FSI) (Freeman, 1978). This index measures the extent of segregation among those registered to one of the two main parties.[26] Given two distinct groups in a population, the FSI is defined as $1-\frac{p}{\pi}$, where p denotes the observed proportion of between-group connections and π the expected proportion if connections were generated randomly. The FSI ranges between 0 (a randomly generated network) and 1 (a network with fully segregated groups). The average FSI in our networks is 0.112, indicating that partisan segregation is low. As a benchmark, the FSI of partisan segregation within Twitter networks in the 111th United States Congress is 0.590 (Sparks, 2010).

Measurement concerns

A central concern in the vote-buying literature is that self-reported measures of vote buying are subject to social desirability or reporting bias, which may itself be correlated with the explanatory variable of interest, in our case *hearing*. This is unlikely to be a source of bias in our setting for at least two reasons. First, in contrast to the previous literature, our measures of vote buying are reported by the brokers themselves. We were able to collect this type of information because of years of work in the villages that established trust between the researchers and brokers. Second, the inclusion of broker and citizen fixed effects allows us to control for any social desirability bias

[26] Due to endogeneity concerns, we use official party registration. As a result, we compute the FSI only for citizens who are registered to a party.

that is specific to each broker or to each citizen. For example, broker fixed effects control for any general tendency a broker may have to under-report vote buying. Similarly, citizen fixed effects control for the possibility that brokers may be more likely to remember having targeted citizens with a higher network centrality. Finally, we can dismiss the concern that brokers report differently depending on the person they are reporting about, as Panel B of Table 2 shows that the patterns for broker-reported vote buying and citizen-reported party support are similar.

We next address two concerns related to the measurement of *hearing*: the specific choices of T and p used in the calculation, and the partial sampling of the network. We test whether our results are robust to the choice of T (the number of periods that information originating from a citizen can circulate through the network) and p (the probability that nodes pass on information). Recall that we set T equal to 7, which is the largest social distance between a citizen and a broker in any village network in our sample and p to between 0.08 and 0.14, which is the inverse of the largest eigenvalue of the adjacency matrix for each village's social network. We conduct two complementary exercises to show that the results are not driven by the choice of T. In Panel A in Appendix Table C2, we show that results are robust to setting T equal to the maximum social distance between brokers and citizens in each village's network, which varies between 3 and 7. In Appendix Figure C2a, we show how the relationship between *hearing* and targeting in the main specification varies as T goes from 1 to 50 at intervals of five. The relationship is robust to the choice of T and starts out increasing before it flattens out at around $T = 10$. Similarly, to show that our results are not driven by the choice of p, in Appendix Figure C2b we show how the relationship between *hearing* and targeting varies as p goes from 0.05 to 0.50 at intervals of 0.05. The relationship peaks at around $p = 0.10$, and remains significant from 0.05 to 0.50. Appendix Figure C3 summarizes the results of Figures C2a and C2b through a heat map, with lighter colors denoting larger coefficients.

A third possible concern is that we have a partial sampling of households in the village networks, which may bias our estimates (Chandrasekhar and Lewis, 2016). First, note that the 47% average direct sampling rate in our villages is higher than most of the network literature in developing countries.[27] In addition, we also include network information of indirectly surveyed households, and thus have an average sampling rate of 83%.

To address this potential concern more directly, we assess how our estimates are affected after sequentially dropping the least partially sampled village among those remaining in the sample. Specifically, we begin in column (1) of Table 5 with the full sample and end in column (10) with

[27] According to Chandrasekhar and Lewis (2016), papers using social network data in developing countries have a median sampling rate of 42%, and only slightly over a quarter of them have sampling rates above 47%.

only the village with the highest sampling rate. The relationship between *hearing* and vote buying is highly robust. All of the results except those in the last column, which only has 84 observations, maintain statistical significance, and even in that column the point estimate is in line with the original estimate.

The point estimates in Table 5 increase as we drop the villages with lower sampling rates, suggesting that partial sampling leads to attenuation bias. In results not shown here, we find a positive association between village-level estimates of the coefficient of interest and the direct sampling rate in the village, again suggesting that the main impact of missing network data is attenuation bias. This finding is consistent with a growing literature looking at how sparse or missing network data affects the statistical analysis of networks (Cai, 2023; Chandrasekhar and Lewis, 2016; Fellows and Handcock, 2023; Griffith and Kim, 2023; Hsieh et al., 2024). While these papers differ in their use of topological statistics, two consistent findings emerge from this literature. One finding is that the bias decreases as the sampling rate increases (Chandrasekhar and Lewis, 2016; Fellows and Handcock, 2023). The other finding is that sampling often leads to attenuation bias (Cai, 2023; Griffith and Kim, 2023).[28]

While prior literature suggests that sampling may lead to attenuation bias, it is unclear whether our hearing statistic will exhibit the same properties. To investigate this, we conducted a simulation exercise using data from the village with the highest direct sampling rate of 91%. This village is the closest we get to having data on the full network, and so the estimate of β in that village is subject to very little partial sampling bias. We then randomly drop different shares of observations in the village to see how estimates of β are affected by partial sampling. Specifically, for different partial sampling rates, we drew $1,000$ sub-samples from the set of nodes in that village retaining $x\%$ of the non-broker nodes and retaining all brokers. For each sub-sample, we construct the simulated network that removes any link between two dropped nodes. This takes into account that in our data we have indirect knowledge of links from reports of one node even when the other is not sampled. Next, for every non-dropped node in each simulated network, we computed our *hearing* metric.

Once the data generation process was complete, we estimated the association between vote-buying and *hearing* for the non-dropped nodes in each simulated network, controlling for broker and citizen fixed effects. The regression only included citizen observations that were not dropped in the earlier sampling step. We then averaged the estimated coefficient on *hearing* across the

[28]Because we look at how dyad-level characteristics are correlated with dyad-level outcomes, Cai (2023) is arguably the most closely related paper. Cai (2023) estimates the impact of node-level diffusion centrality (the sum of the dyad-level *hearing*, which is the focus of our paper) on node-level outcomes. He finds that sparseness and measurement error in network data will tend to lead to attenuation bias in the measured impact of diffusion centrality. He also finds that the estimated effects of diffusion centrality are much more robust than estimated effects of eigenvector centrality.

1,000 simulations.

Appendix Figure C4 presents the results of this procedure for different values of $x \in \{0.33, 0.5, 0.66, 0.75\}$. For the full network in that village, we estimate a coefficient on *hearing* of 0.47. This is the coefficient shown in column (10) of Table 5. As we reduce the sampling rate, the coefficients attenuate slightly. Even with a sampling rate as low as 33%, the association between vote-buying and hearing remains close to 0.40, though the estimates become slightly noisier. Overall, these results, in combination with existing literature, suggest that if our estimates are biased, they likely represent a lower bound.

A related concern is that the ties of some of the brokers in our sample (12 out of 32) were indirectly surveyed. Panel B of Appendix Table C2 shows robustness of the estimates to restricting the sample to broker-citizen observations involving brokers whose ties were directly surveyed.

Information diffusion and vote buying

Thus far, we have provided robust evidence that brokers target citizens about whom they are able to hear more information through their network. Next, we look at what type of information matters for brokers' targeting decisions. First, brokers have access to voter rolls and know which citizens are registered to their party. Second, brokers can learn many characteristics about a citizen through their network. In particular, the conceptual framework in Section 3 highlights that brokers rely on their network to learn citizens' reciprocity levels.

We provide evidence that information flows explain how brokers choose which citizens to target. The conceptual framework predicts that brokers are more likely to target non-copartisans who both are reciprocal, and thus will return the favor, and with whom they have higher *hearing* to learn about the citizen's reciprocity. Moreover, they target copartisans independently of *hearing* and reciprocity. We test these predictions by including the double and triple interactions of *hearing*, citizen reciprocity, and whether the citizen is registered to the broker's party as in equation (2).

Columns (1) and (2) of Table 6 replicate the same columns in Panel B of Table 2 but for the smaller sample of broker-citizen pairs for which we have the experimental measure of citizen reciprocity. Columns (3) and (4) add this measure of citizen reciprocity, along with its interaction with *hearing*. In columns (5) and (6), we instead add an indicator for whether the citizen is not registered to the broker's party as well as its interaction with *hearing*. Citizens usually register for a party at the same time that they first register to vote and do not often change official party registration, which lessens concerns that party registration is endogenous. Lastly, columns (7) and (8) add the triple interaction.

The findings in columns (3) and (4) suggest that, on average, brokers target citizens that they learn through the network are more reciprocal. In other words, it is not enough for a citizen to simply be reciprocal to be targeted. A reciprocal citizen must be located in the network relative to a broker in such a way that the broker hears about the citizen's reciprocity.

Results in columns (5) and (6) further support the fact that information about citizens' party affiliation is publicly known. Information diffusion through the network is not necessary for brokers to learn about citizens' party registration.[29] Brokers are more likely to target citizens registered to their party no matter what they hear about them through the network. The coefficient on non-registration to the broker's party is relatively large in magnitude, suggesting that mobilizing copartisans to "get out the vote" is an important part of brokers' targeting strategy (Casas, 2018; Larreguy et al., 2016; Nichter, 2008).

We then look at the joint effects of citizen reciprocity and party registration. We interpret the positive coefficient on the triple interaction in columns (7) and (8) as indicating that the targeting of reciprocal citizens is concentrated among citizens not registered to the broker's party, but about whom the brokers can learn the citizen's reciprocity level.[30] Also consistent with the framework, the interaction between *hearing* and reciprocity does not predict the targeting of citizens registered to the broker's party.

In results not shown here, we run the same regressions replacing reciprocity with each of the citizen characteristics that are components of the broker knowledge index: land ownership, years of education, the frequency with which a citizen would retaliate wrongdoing, the level of trust the citizen has towards other villagers, strong preference for the Colorado party, and strong preference for the Liberal party. None of these exhibit the same patterns as reciprocity, suggesting that brokers target based on reciprocity as heard through the social network, and not based on other characteristics they might hear about through the social network. Overall, these findings suggest that brokers leverage the information diffused through their social network to guide their targeting in a way that is consistent with our framework.

Table 7 follows the same structure of Table 6, but uses as an outcome an indicator for whether the citizen reports that she supports the broker's party after the election. The results follow the same pattern as those in Table 6. We take this as further indication that the vote buying guided by the information that social networks diffuse about citizens to brokers is effective at persuading targeted citizens to vote for the broker's party. Lastly, to further dismiss endogeneity concerns, Appendix Table C3 follows the same structure as Table 6, but uses as an outcome an indicator for whether a non-political informal transaction tie exists between the broker's and citizen's households. In

[29]The coefficient on party registration in these columns is robust to the exclusion of *hearing* and its interaction.

[30]Results are qualitatively similar if the outcomes are whether the broker reported offering something to the citizen and whether the broker reported approaching the citizen.

contrast to the result in Table 6, while *hearing* is naturally associated with non-political transaction ties, its association is not mediated by either citizen reciprocity or party registration.

We also provide suggestive evidence against the alternative interpretation that our results are driven by brokers targeting citizens who are good at transmitting information to other citizens and persuading them to vote for specific candidates. Specifically, the concern is that better-connected citizens might also be able to persuade more citizens in the village. Controlling for citizen fixed effects largely deals with this concern. As an additional test, Appendix Table C4 adds the interaction of *hearing* with citizen degree centrality, diffusion centrality, eigenvector centrality, and betweenness centrality. The coefficient on *hearing* is robust to these additions, while the coefficients on the interaction terms are negative and small in magnitude, further dismissing the idea that brokers are differentially targeting more well-connected citizens in an attempt to purchase persuasion.

Broker selection within village networks

As the conceptual framework in Section 3 highlights, parties have the incentive to select brokers who are more centrally located within their networks, particularly among non-copartisans. Appendix Figure C5 plots kernel density estimates of standardized degree centrality, betweenness centrality, eigenvector centrality, and diffusion centrality for broker households (32 observations) and citizen households (1,000 observations, including those directly surveyed and those whose connections we learned about indirectly). Consistent with the conceptual framework's prediction, in each plot the density for brokers lies clearly to the right of that of the citizens, indicating that brokers have higher centrality than citizens.

In Table 8, we refine this analysis by regressing the network measures of centrality on an indicator for whether a household contains a broker as well as village fixed effects.[31] In columns (1), (2), (3), and (6), we include all households in the village that were either directly or indirectly sampled and control for an indicator for whether the household was directly sampled in our network survey. In columns (4), (5), (7), and (8), we focus on the households for which we have party registration data and calculate diffusion and degree centrality among each household's copartisans and non-copartisans.[32,33] In Panel A, we consider overall standardized centrality measures as outcomes,[34] while in Panel B we consider the within-village percentile of each centrality measure to

[31] Appendix Table D2 shows that the results in Table 8 are robust to multiple hypothesis adjustments.

[32] We only have registration data for 245 of the 295 directly surveyed individuals who we could match with the government data.

[33] We cannot distinguish between copartisans and non-copartisans for eigenvector or betweenness centrality.

[34] Appendix Table C5 presents the summary statistics for the non-standardized centrality measures for all households, whether or not they were interviewed.

provide a better sense of the magnitude of the differences in centrality.

Results in columns (1), (2), (3), and (6) of Panel A show that brokers' network centrality is significantly higher than that of regular citizens, ranging from 0.10 standard deviation higher betweenness centrality to 0.60 standard deviation higher eigenvector centrality. Panel B further indicates that broker centrality is between 8 and 21 percentiles greater than that of other citizens.[35]

In columns (4) and (5), we see that brokers' diffusion centrality is only significantly higher among citizens not registered to their party. Brokers' *hearing* with copartisans is not higher than that of the average citizen. This is consistent with the predictions of the model and the previous results. Brokers have an incentive to target non-copartisans who they believe to be reciprocal. If brokers can only learn who is reciprocal through their social networks, then parties will want to select brokers who have high *hearing* with non-copartisans.[36] Columns (7) and (8), show that degree centrality exhibits a similar pattern as does diffusion centrality. These results suggest that political parties recruit brokers with high centrality, particularly among non-copartisans.

7 Concluding Remarks and Implications

Vote buying is pervasive throughout the developing world, and political brokers are critical intermediaries in this exchange. We use data on village networks and vote buying to show that the amount and content of the information a political broker hears about a citizen through the social network predicts whether the broker tries to buy the citizen's vote. We find that the more information a broker hears about a citizen, the more the broker learns politically relevant information about her, even including her social preferences. Brokers use this knowledge to target reciprocal individuals but only among those not registered to their party. Among copartisans, brokers target citizens regardless of their level of reciprocity. As a result, parties select brokers who are more centrally located among non-copartisans.

While these findings are specific to a setting where citizen turnout and partisanship are observable, we believe our results likely extend to many other countries throughout the developing world where party registration and turnout are verifiable through public records (e.g., Argentina, Brazil, Mexico, and the Philippines). Furthermore, the importance of social networks as conduits of information for targeting decisions is not confined to vote buying. Several studies have shown

[35]These results are consistent with Ravanilla et al. (2021), who find that brokers in the Philippines exhibit higher betweenness centrality than a random sample of villagers. In results not shown here, we find that centrality measures are similar for brokers of the two main parties.

[36]There is no significant correlation between brokers' years of experience and brokers' diffusion centrality. This is suggestive evidence that when brokers are chosen they already have higher centrality rather than growing their diffusion centrality over time in order to better perform their job.

that targeting centrally-located individuals can be crucial for technology adoption (Beaman et al., 2021), vaccination (Banerjee et al., 2021), and other health behaviors (Kim et al., 2015).

Our research also has broader policy implications regarding how changes to social networks can influence an electoral practice widely believed to undermine political accountability and limit the provision of public goods. Our findings suggest that interventions weakening social ties could potentially reduce the effectiveness of vote buying, paving the way for more programmatic electoral strategies.

For example, Banerjee et al. (2024) demonstrate that the entry of formal financial institutions reduces households' social ties by 22%. If we simulate the effects of randomly removing 22% of the links in each village network, we predict a 9% reduction in vote buying and an associated 8% decrease in votes for the broker's party.[37] Furthermore, as an upper bound, if we set *hearing* to zero, effectively eliminating the role that social networks play in diffusing information from citizens to brokers, there would be a 27% reduction in vote buying and an associated 24% decrease in votes for the broker's party.

Given the welfare loss associated with vote buying, these estimates of the decrease in vote buying are inherently significant. Additionally, we make a back of the envelope estimate of the effect of reduced vote buying on electoral outcomes. Assuming our sample is representative of village voters and that broker support reflects voting behavior, removing 22% of links would result in a change in the identity of the winning mayor in one municipality out of the ten we study. If we instead set *hearing* to zero, the identity of the winning mayor would change in four municipalities.

While a reduction in social ties may decrease the incidence of vote buying and mitigate the associated harms, it is important to consider the potential costs of weakened social ties. Social networks provide several key welfare-enhancing benefits, such as informal insurance and the facilitation of technology adoption. Exploring the various welfare effects of social networks is an exciting area for future research.

References

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016). Network structure and the aggregation of information: Theory and evidence from Indonesia. *American Economic Review 106*(7), 1663–1704.

[37]We conducted 100 simulations, each time randomly removing 22% of the links in each village network and then recalculating *hearing*. Using the estimates in panel B of Table 2 (column 2 and column 5), we then predicted the likelihood that a broker would offer something to or approach a citizen and the likelihood that the citizen would support the broker's party. We then computed the average difference between the propensities predicted using the original *hearing* measure and the newly simulated *hearing* measures.

Alix-Garcia, J. M., K. R. Sims, and L. Costica (2021). Better to be indirect? Testing the accuracy and cost-savings of indirect surveys. *World Development 142*.

Alt, J., A. Jensen, H. Larreguy, D. Lassen, and J. Marshall (2022). Diffusing political concerns: How unemployment information passed between social ties influences Danish voters. *The Journal of Politics 84*(1), 383–404.

Arias, E., P. E. Balán, H. A. Larreguy, J. Marshall, and P. Querubín (2019). Information provision, voter coordination, and electoral accountability: Evidence from Mexican social networks. *American Political Science Review 113*(2), 475–498.

Auerbach, A. M. and T. Thachil (2020). Cultivating clients: Reputation, responsiveness, and ethnic indifference in India's slums. *American Journal of Political Science 64*(3), 471–487.

Auerbach, E. (2022). Identification and estimation of a partially linear regression model using network data. *Econometrica 90*(1), 347–365.

Baland, J.-M. and J. A. Robinson (2008). Land and power: Theory and evidence from Chile. *American Economic Review 98*(5), 1737–65.

Banerjee, A., E. Breza, A. G. Chandrasekhar, E. Duflo, M. O. Jackson, and C. Kinnan (2024). Changes in social network structure in response to exposure to formal credit markets. *Review of Economic Studies 91*(3), 1331–1372.

Banerjee, A., A. G. Chandrasekhar, S. Dalpath, E. Duflo, J. Floretta, M. O. Jackson, H. Kannan, F. Loza, A. Sankar, A. Schrimpf, and M. Shrestha (2021). Selecting the most effective nudge: Evidence from a large-scale experiment on immunization. Unpublished working paper.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144), 1236498.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *Review of Economic Studies 86*(6), 2453–2490.

Barsbai, T., V. Licuanan, A. Steinmayr, E. Tiongson, and D. Yang (2021). Information and immigrant settlement. Unpublished working paper.

Beaman, L., A. BenYishay, J. Magruder, and M. Mobarak (2021). Can network theory-based targeting increase technology adoption? *American Economic Review 111*(6), 1918–1943.

Blattman, C., H. Larreguy, B. Marx, and O. Reid (2024). Eat widely, vote wisely? Lessons from a campaign against vote buying in Uganda. Unpublished working paper.

Bloch, F., G. Genicot, and D. Ray (2008). Informal insurance in social networks. *Journal of Economic Theory 143*(1), 36–58.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler (2012). A 61-million-person experiment in social influence and political mobilization. *Nature 489*(7415), 295–298.

Bramoullé, Y. and G. Genicot (2020). Diffusion centrality: Foundations and extensions. Unpublished working paper.

Breza, E., A. G. Chandrasekhar, and H. Larreguy (2019). Network centrality and informal institutions: Evidence from a lab experiment in the field. Unpublished working paper.

Brierley, S. and N. L. Nathan (2021). The connections of party brokers: Which brokers do parties select? *Journal of Politics 83*(3), 884–901.

Cai, Y. (2023). Linear regression with centrality measures. Unpublished working paper.

Casas, A. (2018). Distributive politics with vote and turnout buying. *American Political Science Review 112*(4), 1111–1119.

Chandrasekhar, A. G., C. Kinnan, and H. Larreguy (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics 10*(4), 43–78.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica 88*(1), 1–32.

Chandrasekhar, A. G. and R. Lewis (2016). Econometrics of sampled networks. Unpublished working paper.

Chuang, Y. and L. Schechter (2015). Social networks in developing countries. *Annual Review of Resource Economics 7*(1), 451–472.

Collier, P. and P. C. Vicente (2014). Votes and violence: Evidence from a field experiment in Nigeria. *Economic Journal 574*(124), 327–355.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review 100*(1), 35–69.

Cox, G. W. (2015). Electoral rules, mobilization, and turnout. *Annual Review of Political Science 18*, 49–68.

Cox, G. W. and M. D. McCubbins (1986). Electoral politics as a redistributive game. *Journal of Politics 48*(2), 370–389.

Cruz, C. (2019). Social networks and the targeting of vote buying. *Comparative Political Studies 52*(3), 382–411.

Cruz, C., J. Labonne, and P. Querubín (2017). Politician family networks and electoral outcomes: Evidence from the Philippines. *American Economic Review 107*(10), 3006–3037.

Dixit, A. and J. Londregan (1996). The determinants of success of special interests in redistributive politics. *Journal of Politics 58*(4), 1132–1155.

Dosek, T. (2019). La persistencia de caudillos locales en América Latina. Doctoral Thesis. Santiago de Chile: Pontificia Universidad Católica del Chile.

Enríquez, J. R., H. Larreguy, J. Marshall, and A. Simpser (2024). Mass political information on social media: Facebook ads, electorate saturation, and electoral accountability in Mexico. *Journal of the European Economic Association*.

Fafchamps, M. and J. Labonne (2020). Family networks and distributive politics. *Journal of the European Economic Association 18*(4), 1697–1725.

Fafchamps, M., A. Vaz, and P. C. Vicente (2020). Voting and peer effects: Experimental evidence from Mozambique. *Economic Development and Cultural Change 68*(2), 567–605.

Fafchamps, M. and P. C. Vicente (2013). Political violence and social networks: Experimental evidence from a Nigerian election. *Journal of Development Economics 101*, 27–48.

Fellows, I. E. and M. S. Handcock (2023). Modeling of networked populations when data is sampled or missing. *METRON 81*, 21—35.

Finan, F. and L. Schechter (2012). Vote-buying and reciprocity. *Econometrica 80*(2), 863–881.

Freeman, L. C. (1978). Segregation in social networks. *Sociological Methods & Research 6*(4), 411–429.

Gans-Morse, J., S. Mazzuca, and S. Nichter (2014). Varieties of clientelism: Machine politics during elections. *American Journal of Political Science 58*(2), 415–432.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business and Economic Statistics 31*, 253–264.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica 85*(4), 1033–1063.

Griffith, A. and J. Kim (2023). The impact of missing links on linear reduced-form network-based peer effects estimates. Unpublished working paper.

Hsieh, C.-S., Y.-C. Hsu, S. I. M. Ko, J. Kovářík, and T. D. Logan (2024). Non-representative sampled networks: Estimation of network structural properties by weighting. *Journal of Econometrics 240*(1).

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives 28*(4), 3–22.

Jackson, M. O., T. Rodriguez-Barraquer, and X. Tan (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review 102*(5), 1857–97.

Johnsson, I. and H. R. Moon (2021). Estimation of peer effects in endogenous social networks: Control function approach. *Review of Economics and Statistics 103*(2), 328—345.

Kim, D. A., A. R. Hwong, D. Stafford, D. A. Hughes, A. J. O'Malley, J. H. Fowler, and N. A. Christakis (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet 386*(9989), 145–153.

Lachi, M. (2009). Recolección de fondos y gastos electorales en las elecciones municipales. Documentos de Trabajo del Centro de Estudios y Educación Popular Germinal.

Lachi, M. and R. Rojas-Schaffer (2018). *Correligionarios: Actitudes y Prácticas Políticas del Electorado Paraguayo*. Asunción: Arandura.

Larreguy, H. (2013). Monitoring political brokers: Evidence from clientelistic networks in Mexico. Unpublished working paper.

Larreguy, H., J. Marshall, and P. Querubín (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.

Lawson, C. and K. F. Greene (2014). Making clientelism work: How norms of reciprocity increase voter compliance. *Comparative Politics 47*(1), 61–85.

Lehoucq, F. (2007). When does a market for votes emerge? In F. C. Schaffer (Ed.), *Elections for Sale: The Causes and Consequences of Vote Buying*, Chapter 3, pp. 33–45. Boulder, CO: Lynne Rienner Publishers.

Ligon, E. and L. Schechter (2012). Motives for sharing in social networks. *Journal of Development Economics 99*(1), 13–26.

Lindbeck, A. and J. W. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice 52*(3), 273–297.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics 22*(4), 773–793.

Montes, R. (2018). Uno de cada 3 Mexicanos recibió una oferta de compra de voto, según encuesta. *Milenio*.

Munshi, K. (2014). Community networks and the process of development. *Journal of Economic Perspectives 28*(4), 49–76.

New York Times (2023). One secret to a Latin American party's dominance: Buying votes. *New York Times*.

Nichter, S. (2008). Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review 102*(1), 19–31.

Paraguay (2018). Detallan el sistema de la compra y venta de votos. *Paraguay.com*.

Parks, K. (2018). South America's star economy set to vote to keep it that way. *Bloomberg*.

Ravanilla, N., D. Haim, and A. Hicken (2021). Brokers, social networks, reciprocity, and clientelism. *American Journal of Political Science*.

Robinson, J. A. and T. Verdier (2013). The political economy of clientelism. *Scandinavian Journal of Economics 115*(2), 260–291.

Schaffer, F. C. (Ed.) (2007). *Elections for Sale: The Causes and Consequences of Vote Buying*. Boulder, CO: Lynne Rienner Publishers.

Schaffer, J. and A. Baker (2015). Clientelism as persuasion-buying: Evidence from Latin America. *Comparative Political Studies 48*(9), 1093–1126.

Schechter, L. and A. Yuskavage (2012). Inequality, reciprocity, and credit in social networks. *American Journal of Agricultural Economics 94*(2), 402–410.

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature 43*, 392–436.

Sparks, D. (2010). Birds of a feather tweet together: Partisan structure in online social networks. Unpublished working paper.

Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from Argentina. *American Political Science Review 99*(3), 315–325.

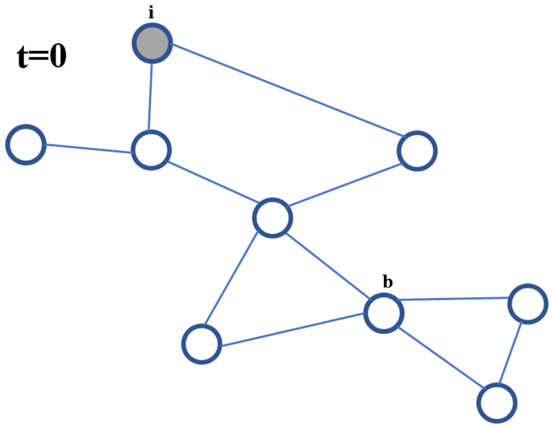
Stokes, S. C., T. Dunning, M. Nazareno, and V. Brusco (2013). *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge University Press.

Szwarcberg, M. (2012). Revisiting clientelism: A network analysis of problem-solving networks in Argentina. *Social Networks 34*(2), 230–240.

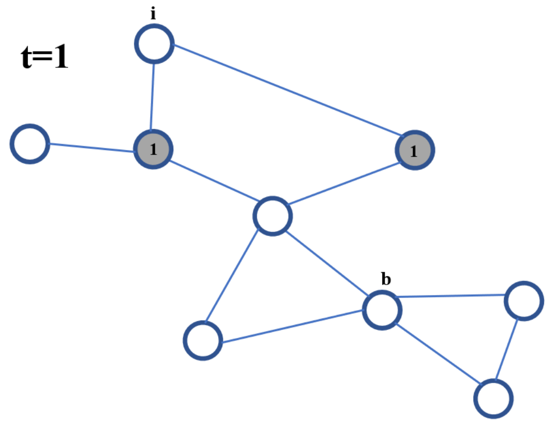
Takasaki, Y., B. L. Barham, and O. T. Coomes (2000). Rapid rural appraisal in humid tropical forests: An asset possession-based approach and validation methods for wealth assessment among forest peasant households. *World Development 28*(11), 1961–1977.

Tables and Figures

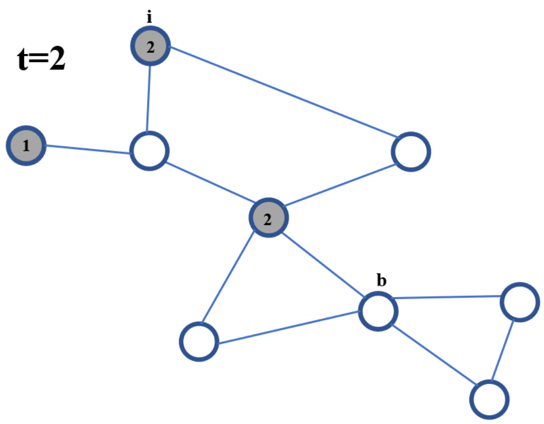
Figure 1: Illustration of the computation of *hearing* between citizen i and broker b



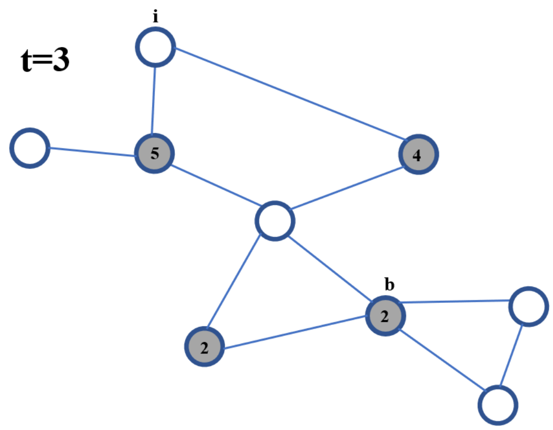
(a) $H_{ib}(0) = 0$



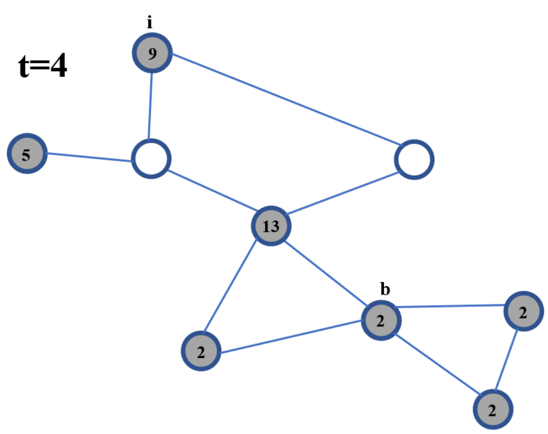
(b) $H_{ib}(1) = 0$



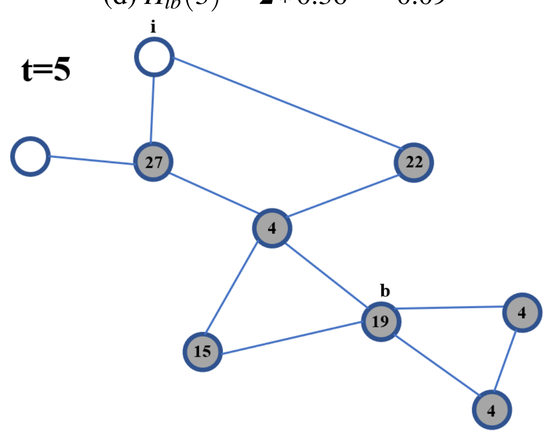
(c) $H_{ib}(2) = 0$



(d) $H_{ib}(3) = \mathbf{2} * 0.36^3 = 0.09$



(e) $H_{ib}(4) = \mathbf{2} * 0.36^3 + \mathbf{2} * 0.36^4 = 0.12$



(f) $H_{ib}(5) =$
$\mathbf{2} * 0.36^3 + \mathbf{2} * 0.36^4 + \mathbf{19} * 0.36^5 = 0.23$

Notes: $H_{ib}(t)$ is *hearing* between citizen i and broker b in period t. We set $p = 0.36$ (the inverse of the largest eigenvalue of the adjacency matrix). Numbers in nodes indicate the maximum number of times information about citizen i might have been transmitted to that node in that period.

Table 1: Summary statistics

	Observations	Mean	Median	Standard Deviation
Non-standardized Network Measures				
Hearing	932	0.120	0.077	0.128
Social distance	932	2.442	2	1.071
Citizen's degree centrality	295	8.875	8	4.508
Citizen's clustering coefficient	295	0.255	0.222	0.190
Citizen's betweenness centrality	295	0.044	0.031	0.049
Citizen's eigenvector centrality	295	0.122	0.110	0.082
Citizen's diffusion centrality	295	8.742	8.328	5.302
Citizen's diffusion centrality among copartisans	245	3.093	2.417	2.536
Citizen's diffusion centrality among non-copartisans	245	2.556	1.825	2.476
Absolute difference in degree centrality	932	4.921	4	4.141
Absolute difference in clustering coefficient	932	0.202	0.118	0.223
Absolute difference in betweenness centrality	932	0.042	0.028	0.046
Absolute difference in eigenvector centrality	932	0.081	0.073	0.061
Absolute difference in diffusion centrality	932	5.456	4.807	4.279
Existence of a support pair	932	0.529	1	0.499
Number of support pairs	932	1.351	1	1.852
Transaction tie	932	0.128	0	0.334
Mediating Measures				
Experimental reciprocity	85	0.044	0	0.076
Not registered to the broker's party	932	0.568	1	0.496
Additional Controls				
Absolute age difference	932	16	14	12
Broker and citizen have the same gender	932	0.580	1	0.494
Distance in kilometers between the broker's and citizen's residences	932	1.409	1.203	0.901
Outcome Measures				
Vote-buying Measures				
Citizen supports the broker's party	932	0.461	0	0.499
Components of the vote-buying targeting index:				
Broker approached citizen during electoral campaign	932	0.477	0	0.500
Broker offered citizen something	932	0.273	0	0.446
Knowledge Measures				
Components of the covariates index:				
Broker knows citizen	932	0.887	1	0.316
Broker knows citizen's spouse	932	0.773	1	0.419
Broker knows citizen's years of education	932	0.807	1	0.395
Broker knows citizen's amount of land	932	0.421	0	0.494
Components of the political index:				
Broker knows strength of citizen's party preference	932	0.593	1	0.491
Components of the social preferences index:				
Broker knows the frequency with which citizen would retaliate	932	0.586	1	0.493
Broker knows whether the citizen generally trusts others in the village	932	0.656	1	0.475

Table 2: Relationship between *hearing* and brokers' knowledge about citizens and vote-buying targeting

	Overall knowledge index (1)	Overall knowledge index (2)	Covariates index (3)	Political index (4)	Social preferences index (5)
Panel A: Brokers' knowledge about citizens					
Hearing	0.2864***	0.2111***	0.1703***	0.1975***	0.1401***
	(0.0416)	(0.0355)	(0.0411)	(0.0714)	(0.0442)
Mean of Dependent Variable	-0.0000	-0.0000	0.0000	-0.0000	0.0000
Broker FE	X	X	X	X	X
Citizen FE		X	X	X	X
Observations	932	932	932	932	932
R^2	0.1546	0.6395	0.6237	0.5362	0.6942

	Vote-buying targeting index (1)	Vote-buying targeting index (2)	Broker offered citizen (3)	Broker approached citizen (4)	Support the same party (5)
Panel B: Brokers' vote-buying targeting					
Hearing	0.2114***	0.3215***	0.0720**	0.1833***	0.1174***
	(0.0436)	(0.0574)	(0.0270)	(0.0309)	(0.0388)
Mean of Dependent Variable	-0.0000	-0.0000	0.2725	0.4775	0.4614
Broker FE	X	X	X	X	X
Citizen FE		X	X	X	X
Observations	932	932	932	932	932
R^2	0.3983	0.6129	0.6688	0.5181	0.4167

Notes: All specifications include broker fixed effects. The dependent variables in Panel A are standardized indices (with mean 0 and standard deviation of 1) that aggregate what the broker knows about the citizen. The covariates index aggregates indicators for whether the broker knows the citizen, whether the broker knows the citizen's spouse's name, whether the broker knows how much land the citizen owns, and whether the broker knows the citizen's years of education. The political index corresponds to an indicator for whether the broker knows the strength of the citizen's party preference. The social preferences index aggregates indicators for whether the broker knows whether the citizen generally trusts others in the village, and whether the broker knows the frequency with which the citizen would retaliate wrongdoing. The overall knowledge index aggregates indicators from all three knowledge categories. The vote-buying targeting index is a standardized sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. "Support the same party" is an indicator that the citizen claims to support the broker's party shortly after the election. *Hearing* is standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 3: Relationship between *hearing* and vote-buying targeting, testing for strategic interactions

	Vote-buying targeting index	
	(1)	(2)
Hearing	0.4149***	0.0642**
	(0.1036)	(0.0252)
Mean same-party brokers' *hearing*	0.1499*	
	(0.0748)	
Mean same-party brokers' targeting		-0.6736***
		(0.0660)
Mean other-party brokers' *hearing*	0.0606	
	(0.1291)	
Mean other-party brokers' targeting		-1.2420***
		(0.0777)
Mean of Dependent Variable	-0.0000	-0.0000
Observations	932	932
R^2	0.6218	0.9268

Notes: All specifications include broker and citizen fixed effects. The vote-buying targeting index is a standardized index (with mean 0 and standard deviation of 1) that takes the sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. *Hearing* is standardized. "Mean same-party brokers' *hearing* (targeting)" represents the standardized average value of *hearing* (the vote-buying targeting index) of other brokers from the same party with the same citizen. "Mean other-party brokers' *hearing* (targeting)" represents the standardized average value of *hearing* (the vote-buying targeting index) of brokers from another party with the same citizen. In case there is no data on same-party or other-party brokers, either because there is no such broker or because we don't have their data, we set mean *hearing* and targeting to zero and include indicators for those cases as controls. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 4: Relationship between *hearing* and vote-buying targeting, with robustness tests for network endogeneity

	Vote-buying targeting index									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hearing	0.3127***	0.2085*	0.2456**	0.2979***	0.2590*	0.3656***	0.3709***	0.1991***	0.3092***	0.3318**
	(0.0758)	(0.1218)	(0.1180)	(0.0749)	(0.1298)	(0.0740)	(0.0767)	(0.0425)	(0.0773)	(0.1517)
Direct link between broker and citizen		0.2224								0.1024
		(0.2001)								(0.2955)
Social distance				-0.0275						-0.0234
				(0.0669)						(0.0883)
Number of support pairs						-0.0761				-0.0896
						(0.0853)				(0.1049)
Transaction tie									0.0349	-0.1434
									(0.1399)	(0.1929)
Mean of Dependent Variable	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
Broker-citizen controls	X									X
Instrumenting for *hearing*			X							
Social distance controls					X					
Support pair controls							X			
Excluding transaction ties								X		
Observations	932	932	932	932	932	932	932	932	932	932
R^2	0.7335	0.6140		0.6130	0.6168	0.6135	0.6195	0.6095	0.6129	0.7345

Notes: All specifications include broker and citizen fixed effects. The dependent variable is a standardized index (with mean 0 and standard deviation of 1) that takes the sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. Columns (1) and (10) control for broker-citizen controls: a) an indicator for the broker and citizen having the same gender; b) an indicator for the citizen being registered to the broker's party; c) standardized geographical distance between the broker's and citizen's residences; d) standardized absolute difference in age, degree centrality, clustering coefficient, betweenness centrality, diffusion centrality, and eigenvector centrality; and e) broker fixed effects interacted with citizen's degree centrality, clustering coefficient, betweenness centrality, diffusion centrality, and eigenvector centrality. "Direct link between broker and citizen" is an indicator for observations which are directly linked. In column (3), we instrument for the original *hearing* measure with one that excludes information flows from directly connected citizen-broker dyads. We do not show the R^2 for the instrumental variable regression in column (3). Social distance controls include indicators for each value of social distance. Support pair controls include indicators for each value of the number of support pairs. All network measures (except the indicator variables) are standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 5: Relationship between *hearing* and vote-buying targeting (sequentially dropping least partially sampled remaining village)

	Vote-buying targeting index									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hearing	0.3215***	0.3347***	0.4051***	0.4371***	0.4238***	0.4190***	0.4197***	0.4241**	0.3456*	0.4735
	(0.0574)	(0.0579)	(0.0611)	(0.0606)	(0.0925)	(0.0950)	(0.1154)	(0.1376)	(0.1442)	(0.3668)
Mean of Dependent Variable	-0.0000	0.0227	0.0676	0.1484	0.1977	0.2056	0.2435	0.1717	0.1470	0.6491
Observations	932	846	698	608	524	466	348	258	171	84
R^2	0.6129	0.6092	0.6280	0.6322	0.6407	0.6344	0.6500	0.6442	0.6844	0.5064

Notes: All specifications include broker and citizen fixed effects. The vote-buying targeting index is a standardized (with mean 0 and standard deviation of 1) sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. Each column starting from column (2) drops the least partially sampled remaining village, until column (10) which contains only the most completely sampled village. *Hearing* is standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 6: Relationship between *hearing* and vote-buying targeting, by party registration and experimental reciprocity

	Vote-buying targeting index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hearing	0.2275***	0.2757***	0.2283***	0.2643**	0.1179*	0.1221	0.1135*	0.1497*
	(0.0597)	(0.0985)	(0.0550)	(0.1059)	(0.0662)	(0.0860)	(0.0611)	(0.0797)
Experimental reciprocity			-0.0595				0.1300	
			(0.0659)				(0.1462)	
Not reg. to broker's party					-0.6944***	-0.8182***	-0.7192***	-0.8451***
					(0.1314)	(0.1735)	(0.1410)	(0.1809)
Experimental reciprocity × *Hearing*			0.1346**	0.1243			-0.0623	-0.0549
			(0.0653)	(0.1417)			(0.0854)	(0.1321)
Experimental reciprocity × Not reg. to broker's party							-0.1707	-0.2424
							(0.1710)	(0.2283)
Not reg. to broker's party × *Hearing*					0.0949	0.0286	0.1028	0.0558
					(0.0972)	(0.1326)	(0.0925)	(0.1303)
Not reg. to broker's party × *Hearing* × Experimental reciprocity							0.2674**	0.2775*
							(0.1089)	(0.1531)
Mean of Dependent Variable	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Broker FE	X	X	X	X	X	X	X	X
Citizen FE		X		X		X		X
Observations	271	271	271	271	271	271	271	271
R^2	0.4839	0.6253	0.5002	0.6310	0.5769	0.6984	0.5995	0.7153

Notes: All specifications include broker fixed effects. The dependent variable is a standardized index (with mean 0 and standard deviation of 1) that takes the sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. "Not reg. to broker's party" indicates the citizen is not officially registered to the broker's party. "Experimental reciprocity" is the experimental measure of reciprocity used in Finan and Schechter (2012). The coefficients for experimental reciprocity in columns (4) and (8) are absorbed by the citizen fixed effects. *Hearing* and experimental reciprocity are standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 7: Relationship between *hearing* and citizen support for the broker's party, by party registration and experimental reciprocity

	Support same party							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hearing	0.1290***	0.1779**	0.1323***	0.1737**	0.0493	0.0413	0.0408	0.0505
	(0.0349)	(0.0695)	(0.0330)	(0.0714)	(0.0406)	(0.0516)	(0.0362)	(0.0520)
Experimental reciprocity			-0.0620				0.0770**	
			(0.0398)				(0.0357)	
Not reg. to broker's party					-0.6324***	-0.8742***	-0.6405***	-0.8824***
					(0.0712)	(0.0732)	(0.0719)	(0.0715)
Experimental reciprocity × *Hearing*			0.0681*	0.0453			-0.0192	-0.0175
			(0.0359)	(0.0544)			(0.0206)	(0.0165)
Experimental reciprocity × Not reg. to broker's party							-0.1163**	-0.0627
							(0.0497)	(0.0549)
Not reg. to broker's party × *Hearing*					0.0444	-0.0459	0.0514	-0.0389
					(0.0556)	(0.0608)	(0.0546)	(0.0589)
Not reg. to broker's party × *Hearing* × Experimental reciprocity							0.0817*	0.0871***
							(0.0447)	(0.0314)
Mean of Dependent Variable	0.4280	0.4280	0.4280	0.4280	0.4280	0.4280	0.4280	0.4280
Broker FE	X	X	X	X	X	X	X	X
Citizen FE		X		X		X		X
Observations	271	271	271	271	271	271	271	271
R^2	0.2408	0.4464	0.2645	0.4495	0.5482	0.7773	0.5617	0.7835

Notes: All specifications include broker fixed effects. The dependent variable is is an indicator that the citizen claims to support the broker's party shortly after the election. "Not reg. to broker's party" indicates the citizen is not officially registered to the broker's party. "Experimental reciprocity" is the experimental measure of reciprocity used in Finan and Schechter (2012). The coefficients for experimental reciprocity in columns (4) and (8) are absorbed by the citizen fixed effects. *Hearing* and experimental reciprocity are standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table 8: Brokers' differential relative position within their village social network

	Betweenness centrality	Eigenvector centrality	Diffusion centrality			Degree centrality		
			Overall	Among copartisans	Among non-copartisans	Overall	Among copartisans	Among non-copartisans
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Overall standardized centrality measures								
Broker	0.1046	0.6021***	0.5739***	-0.1868	2.1281***	0.3844***	-0.0195	1.8767***
	(0.1538)	(0.1428)	(0.1457)	(0.1740)	(0.1644)	(0.1233)	(0.1602)	(0.1777)
Mean of Dependent Variable	-0.0000	0.0000	0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000
R^2	0.2850	0.3840	0.3581	0.5096	0.5618	0.5406	0.5842	0.4882
Panel B: Within-village percentiles of centrality measures								
Broker	7.8281**	21.4197***	19.8667***	-3.2361	39.9104***	13.4666***	0.5356	33.5084***
	(3.6530)	(4.6490)	(4.4525)	(7.1397)	(6.5212)	(3.5972)	(6.7388)	(6.4322)
Mean of Dependent Variable	50.0964	50.0160	49.9884	48.6286	48.6286	49.9816	49.6653	50.0184
Village FE	X	X	X	X	X	X	X	X
Directly surveyed FE	X	X	X	X	X	X	X	X
Observations	1,032	1,032	1,032	245	245	1,032	245	245
R^2	0.4722	0.2173	0.2827	0.0090	0.1733	0.4998	0.0018	0.1125

Notes: All specifications include village fixed effects and control for whether the household was surveyed directly about their network ties. Sample for columns (1) to (3) and (6) includes all households directly and indirectly sampled. Columns (4), (5), (7), and (8) include all households for which we have party registration data. Column (3) presents diffusion centrality (the sum of *hearing*) across all households, column (4) presents diffusion centrality only among copartisans of the household, and column (5) presents diffusion centrality only among non-copartisans of the household. Column (6) presents degree centrality across all households, column (7) presents degree centrality only among copartisans of the household, and column (8) presents degree centrality only among non-copartisans of the household. Copartisanship is determined by whether individuals were registered to the same party. The dependent variables in Panel A are standardized (mean 0, s.d. 1), while those in Panel B are the within-village percentiles (ranging from 1 to 100) for each corresponding outcome. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

For Online Publication

A Model

In this section, we present a simple vote-buying model which fleshes out the conceptual framework in Section 3. The model is in the spirit of Gans-Morse et al. (2014).[38] The model provides a simple stylized theory of how brokers decide which citizens to target. Broker targeting depends on the extent to which the broker believes the citizen to be reciprocal, which is shaped by their relative positions in the network and the citizen's reciprocity, and whether the broker and citizen are copartisans. The model also highlights a party's interest in recruiting brokers who are centrally located among non-copartisans.

Citizens and parties

Consider an environment with N citizens and two political parties, an incumbent party and an opposition party. Each party has a fixed opposing platform on a uni-dimensional ideological spectrum and maximizes its vote share. To attract citizens, parties can offer targeted payments to citizens, subject to a limited budget B. As in Baland and Robinson (2008) and Gans-Morse et al. (2014), we assume for simplicity that the budget of the incumbent party relative to the opposition is sufficiently large that the incumbent can act as a monopolist in its targeting decisions. While this assumption might not hold across all settings, we conjecture that the main predictions of the model would continue to hold if we relaxed the assumption, as long as brokers working for different parties know sufficiently little about the information each other have.

Each citizen i receives expressive utility, $X_{ij}, j \in \{I, O\}$, for voting for the incumbent party I or the opposition party O. We can think of this utility as originating from the party's fixed opposing platform. We normalize expressive utility such that $X_{iI} + X_{iO} = 1$ and $X_{ij} \in \{0, 1\}$. This implies that each citizen gets expressive utility of 1 from voting for one of the parties and 0 from voting for the other. Citizens incur cost c_i to vote. For convenience, we assume that costs are distributed uniformly over the interval $[0, \bar{c}]$.

Citizens also receive utility from party transfers. Parties can offer two types of transfers, contractible and non-contractible. Contractible transfers, T_i^c, are payments conditional on an observ-

[38]Our model differs from Gans-Morse et al. (2014) in that it: a) introduces contractible and non-contractible transfers, b) includes citizens with heterogeneous reciprocity levels, c) includes brokers who can learn about citizen reciprocity through the social network, d) abstracts from abstention buying, and e) analyzes broker recruitment by the incumbent party.

able action, which in this setting is turning out to vote. Non-contractible transfers, T_i^{nc}, are monetary gifts that engender warm glow towards the party if the person is reciprocal, where reciprocity $\rho_i = \rho \in (0,1]$ if the person is reciprocal and zero otherwise.

A non-reciprocal citizen who is offered transfers T_i^{nc} and T_i^c by party I will receive the following utility:

$$\begin{cases} X_{ij} + T_i^{nc} + T_i^c - c_i, & \text{if they vote, and} \\ T_i^{nc}, & \text{if they do not vote.} \end{cases}$$

A reciprocal citizen who is offered transfers by party I will receive the following utility:

$$\begin{cases} X_{iI} + (1+\rho)T_i^{nc} + T_i^c - c_i, & \text{if they vote for party } I \text{ ,} \\ X_{iO} + T_i^{nc} + T_i^c - c_i, & \text{if they vote for party } O\text{, and} \\ T_i^{nc}, & \text{if they do not vote.} \end{cases}$$

Citizens are heterogeneous in their cost of voting (c_i) and in whether they are reciprocal (ρ_i). This model distinguishes between four groups of citizens, N_g with $g \in \{iv, in, ov, on\}$. These groups are defined based on whether, in the absence of transfers, the citizen would turn out to vote and for whom.

- Incumbent Voter (N_{iv}): $X_{iI} - c_i \geq 0$ where $X_{iI} = 1$ (so $1 \geq c$);
- Incumbent Non-Voter (N_{in}): $X_{iI} - c_i < 0$ where $X_{iI} = 1$ (so $1 < c$);
- Opposition Voter (N_{ov}): $X_{iO} - c_i \geq 0$ and $X_{iO} = 1$ (so $1 \geq c$);
- Opposition Non-Voter (N_{on}): $X_{iO} - c_i < 0$ and $X_{iO} = 1$ (so $1 < c$).

Brokers

The incumbent party relies on brokers to target their transfers. We assume that brokers know the voting cost of each citizen, c_i. The voting cost usually depends on variables such as the distance to the polling station and the opportunity cost due to occupation, all of which are relatively easy to observe. Given the Paraguayan context, where party registration is publicly available, we assume that brokers know whether each citizen is a copartisan, X_{iI} and X_{iO}. Brokers do not, however, know with certainty whether a citizen is reciprocal. They do know that a proportion ϕ of citizens are reciprocal, and they may receive signals about each citizen's reciprocity.

The number of signals that a broker receives about whether a specific citizen is reciprocal depends on their relative network positions. We define *hearing* H_{ib} as the number of times broker

b receives a signal about citizen i's reciprocity ρ_i. The signal is correct with probability $\theta \neq .5$. Denote $\widehat{\phi}_{ib}(\rho, H_{ib})$ as the expected belief held by broker b that citizen i is reciprocal (has $\rho_i = \rho$) when that citizen is actually reciprocal given the level of *hearing* H_{ib} between the two. This expectation (with derivation show in Section A.1) is:

$$\hat{\phi}(\rho, H) = \sum_{h=0}^{H} \frac{H!}{h!(H-h)!} \theta^h (1-\theta)^{H-h} \left(\frac{\theta^h (1-\theta)^{H-h} \phi}{\theta^h (1-\theta)^{H-h} \phi + \theta^{H-h} (1-\theta)^h (1-\phi)} \right) \tag{6}$$

Note that if the broker receives no signals about the citizen ($H_{ib} = 0$), then $\widehat{\phi}_{ib}(\rho, H_{ib}) = \phi$ which is the share of reciprocal citizens in the population. If $H_{ib} > 0$ and the signal is perfectly accurate ($\theta = 1$), then $\widehat{\phi}_{ib}(\rho, H_{ib}) = 1$ and the broker is sure that a reciprocal citizen is in fact reciprocal. This is more generally captured by the following lemma.

Lemma 1 $\frac{\Delta \widehat{\phi}_{ib}(\rho, H_{ib})}{\Delta H_{ib}} \geq 0$.

In words, in broker-citizen dyads with higher *hearing*, the broker's posterior about the likelihood that a reciprocal citizen is indeed reciprocal is more accurate. When $\theta = 0.5$, the signal is uninformative as the signal is as good as random, being accurate half the time and inaccurate half the time and the posterior beliefs don't change as the number of signals H increases. When $\theta = 0$ or $\theta = 1$, then the signal is completely informative, hearing it once is enough to learn the truth, and $\widehat{\phi}_{ib}(\rho, H_{ib})$ only increases when the number of signals increases from 0 to 1, but stays constant after that. It is more complicated to find a closed form solution for changes in $\hat{\phi}(\rho, H)$ as H increases for other values of θ. We evaluate how the function changes with H numerically and find that the difference is non-decreasing in H for all values of θ.

Targeting Strategies

The incumbent party and its brokers want to maximize their expected number of votes by offering targeted payments to citizens subject to a budget constraint B.

$$\begin{aligned} \max_{T_i^{nc}, T_i^c} \sum_{i \in N_{iv}} 1 \quad & + \sum_{i \in N_{in}} [\widehat{\phi}_{ib} \mathbb{1}\{1 + \rho T_i^{nc} + T_i^c \geq c_i\} + (1 - \widehat{\phi}_{ib}) \mathbb{1}\{1 + T_i^c \geq c_i\}] \\ + \sum_{i \in N_{ov}} \widehat{\phi}_{ib} \mathbb{1}\{\rho T_i^{nc} \geq 1\} \quad & + \sum_{i \in N_{on}} \widehat{\phi}_{ib} \mathbb{1}\{\rho T_i^{nc} \geq 1\} \mathbb{1}\{\rho T_i^{nc} + T_i^c \geq c_i\}. \end{aligned} \tag{7}$$

The first summation corresponds to the votes for the incumbent of incumbent voters. These citizens turn out and vote for the incumbent without receiving any payments. The second summation consists of votes for the incumbent of incumbent non-voters. These citizens vote for the incumbent as long as $1+T_i^c \geq c_i$. The third summation consists of votes of opposition voters. These citizens turn out to vote regardless of payment, but only vote for the incumbent if they are reciprocal and $\rho T_i^{nc} \geq 1$. The last summation corresponds to votes of opposition non-voters. To convince opposition non-voters to vote for the incumbent, the broker has to target reciprocal citizens and offer them a contractible payment to turn out plus a non-contractible payment to change their vote.

To solve the problem in Equation (7), note that $T_i^{nc} = 0$ for all $i \in N_{in}$. To get an incumbent non-voter to support the incumbent party, the broker needs to pay them their cost of voting. The broker would rather make this payment in the form of a contractible payment, which delivers a vote with certainty, versus a non-contractible payment that only matters if the citizen is reciprocal. Furthermore, given that $T_i^{nc} = 0$ for all $i \in N_{in}$, the second summation reduces to $\sum_{i \in N_{in}} 1\{1+T_i^c \geq c_i\}$, implying that reciprocity does not affect a broker's decision to buy the turnout of incumbent non-voters.

Denote the cost for attempting to buy a vote for the incumbent party $C(c_i,\rho)$ as the sum of the contractible cost $C^c(c_i)$ and the non-contractible cost $C^{nc}(\rho)$. For $i \in \{N_{iv}, N_{ov}\}$, $C^c(c_i) = 0$, whereas for $i \in \{N_{in}, N_{on}\}$, $C^c(c_i) = c_i - 1$. Similarly, let $C^{nc}(\rho)$ denote a vector containing the cost per vote based on a non-contractible payment. For $i \in \{N_{ov}, N_{on}\}$, $C^{nc}(\rho) = \frac{1}{\rho}$.

For given c_i, ρ, and $\widehat{\phi}_i$, the cost (C) of attempting to buy the vote of each citizen type and the probability (π) that they turn out to vote for the incumbent are:

- Incumbent Voter: $C_{iv}(c_i,\rho) = 0$ and $\pi_{iv} = 1$;
- Incumbent Non-Voter: $C_{in}(c_i,\rho) = c_i - 1$ and $\pi_{in} = 1$;
- Opposition Voter: $C_{ov}(c_i,\rho) = \frac{1}{\rho}$ and $\pi_{ov} = \widehat{\phi}_{ib}$;
- Opposition Non-Voter: $C_{on}(c_i,\rho) = c_i - 1 + \frac{1}{\rho}$ and $\pi_{on} = \widehat{\phi}_{ib}$.

An optimal allocation is a vector of payments $T(c_i, \widehat{\phi}_{ib})$ such that no other allocation produces a greater number of expected votes. Given the linearity of the problem, to achieve this optimal allocation under the model's assumptions, the broker simply ranks each citizen from highest to lowest in terms of the ratio of the probability that they turn out to vote for the incumbent and the cost of buying their vote, $\frac{\pi_i}{C(c_i,\rho)}$. The broker provides the required transfers to the first N^* voters such that

$$\sum_{i=1}^{N^*} T(c_i, \widehat{\phi}_{ib}) \leq B.$$

It is optimal for brokers to buy the votes of all incumbent non-voters with low voting costs before buying the votes of any opposition voter. When engaging in the latter, brokers should prioritize opposition voters with the highest expected reciprocity ($\widehat{\phi}_{jb}$).

Without loss of generality, assume that, among opposition voters and opposition non-voters, an opposition voter has the largest $\frac{\pi_i}{C(c_i,\rho)}$. This opposition voter then must have the largest $\widehat{\phi}_{jb}$, which we denote $\widehat{\phi}^{max}$. The next proposition characterizes the most empirically relevant equilibrium.

Proposition 1 *Suppose* $\frac{1-(\rho\widehat{\phi}^{max})^2}{2(\rho\widehat{\phi}^{max})^2\bar{c}} < B$. *Then there exists an equilibrium in which the incumbent party engages in both vote buying and turnout buying. Conditional on such an equilibrium, the incumbent targets incumbent non-voters with the lowest voting cost* c_i*, opposition voters with the highest expected reciprocity* $\widehat{\phi}_{jb}$*, and opposition non-voters with both the lowest* c_i *and highest* $\widehat{\phi}_{jb}$*.*

Proof The broker should first target incumbent non-voters with $c_i < \frac{1+\rho\widehat{\phi}^{max}}{\rho\widehat{\phi}^{max}}$, which has a cost of $\int_1^{\frac{1+\rho\widehat{\phi}^{max}}{\rho\widehat{\phi}^{max}}}(c_i-1)\frac{1}{\bar{c}}dc_i = \frac{1-(\rho\widehat{\phi}^{max})^2}{2(\rho\widehat{\phi}^{max})^2\bar{c}}$. If the budget is greater than such cost, the broker then targets opposition voters, starting with the one with $\widehat{\phi}_{jb} = \widehat{\phi}^{max}$. Then, the broker targets the voters with the largest $\frac{\pi_i}{C(c_i,\rho)}$ according to the following indifference curves $\frac{1}{c_i-1} = \rho\widehat{\phi}_{jb} = \frac{\widehat{\phi}_{kb}}{c_k-1+\frac{1}{\rho}}$ for incumbent non-voters i, opposition voters j, and opposition non-voters k. Given that $\frac{\partial\left(\frac{1}{c_i-1}\right)}{\partial c_i}, \frac{\partial\left(\frac{\widehat{\phi}_{kb}}{c_k-1+\frac{1}{\rho}}\right)}{\partial c_k} < 0$ and $\frac{\partial\left(\rho\widehat{\phi}_{jb}\right)}{\partial\widehat{\phi}_{jb}}, \frac{\partial\left(\frac{\widehat{\phi}_{kb}}{c_k-1+\frac{1}{\rho}}\right)}{\partial\widehat{\phi}_{kb}} > 0$, the broker targets incumbent non-voters and opposition non-voters with the smallest c_i and c_k, and opposition voters and opposition non-voters with the largest $\widehat{\phi}_{jb}$ and $\widehat{\phi}_{kb}$.

Corollary 1 follows from Lemma 1 and Proposition 1.

Corollary 1 *Brokers are more likely to target opposition voters and opposition non-voters who are both reciprocal and with whom they have higher hearing. However, their targeting of incumbent non-voters is independent of their hearing and reciprocity.*

From Lemma 1, for reciprocal citizens, the more signals the broker receives about voters and non-voters, the higher his perceived likelihood that reciprocal citizens are indeed reciprocal. Moreover, from Proposition 1, brokers are more likely to target reciprocal voters and non-voters who favor the opposition about whom they receive more signals. In turn, neither *hearing* nor reciprocity plays a role in the targeting of citizens who favor the incumbent.

Broker Selection

As a result of Proposition 1 and Lemma 1, the expected return to vote buying by the incumbent is increasing in the number of signals a broker receives about citizens who favor the opposition.

In contrast, the number of signals a broker receives about citizens that favor the incumbent is irrelevant to the expected return from vote buying since partisanship is observable and the broker can buy the turnout of incumbent non-voters through contractible transfers, which do not rely on reciprocity for enforcement. The incumbent party will then recruit brokers with the highest *hearing* among non-incumbent supporters.

In sum, this simple model provides three predictions that we test in the available data. First, neither a citizen's reciprocity level nor the broker's *hearing* with the citizen will affect the targeting of copartisans. Second, conditional on an equilibrium in which brokers have the incentive to target non-copartisans, brokers will target reciprocal non-copartisans with whom they are connected in the network in such a way that they can receive more signals. Third, conditional on the same equilibrium, the party will recruit brokers who, on average, have higher *hearing* among non-copartisans.[39]

A.1 Derivation of broker's posterior

The share of reciprocal people in the population is ϕ. Citizen i's reciprocity level is either $\rho_i = \rho$ if he is reciprocal, or $\rho_i = 0$ if he is not. Broker b receives binary signals regarding whether the citizen is reciprocal, and each signal is correct with known probability θ. The signal is r if it signals the citizen is reciprocal and the signal is 0 if it signals the citizen is not reciprocal. The number of signals broker b receives about citizen i is H_{ib}.

According to Bayes' rule, the broker's belief that a citizen is reciprocal ($\rho_i = \rho$) conditional on receiving one signal that the citizen is reciprocal (a signal of r rather than 0) is:

$$P(\rho_i = \rho | r) = \frac{P(r|\rho_i = \rho)P(\rho_i = \rho)}{P(r)} = \frac{P(r|\rho_i = \rho)P(\rho_i = \rho)}{P(r|\rho_i = \rho)P(\rho_i = \rho) + P(r|\rho_i = 0)(1 - P(\rho_i = \rho))}.$$

Because the true share of reciprocal people is ϕ, we know that $P(\rho_i = \rho) = \phi$. The probability of a reciprocal citizen giving a reciprocal signal is $P(r|\rho_i = \rho) = \theta$ and the probability of a non-reciprocal citizen giving a reciprocal signal is $P(r|\rho_i = 0) = 1 - \theta$. So, substituting in we get:

$$P(\rho_i = \rho | r) = \frac{\theta\phi}{\theta\phi + (1-\theta)(1-\phi)}.$$

[39] In an alternative model, parties may instead choose a broker who then invests in relationships that help him hear more about non-copartisans. However, we believe this alternative interpretation is less likely. *Hearing* does not just depend on the broker's direct connections with citizens, but on citizens' connections with one another. It seems unlikely that a broker can change the network architecture of the entire village.

Likewise, the broker's belief that a citizen is reciprocal conditional on receiving one signal of 0 is:

$$P(\rho_i = \rho|0) = \frac{(1-\theta)\phi}{(1-\theta)\phi + \theta(1-\phi)}.$$

The expected belief held by the broker that the citizen is reciprocal when that citizen is actually reciprocal after receiving one signal is $\hat{\phi}(\rho, 1)$. This is the weighted average of the two equations above. If the citizen is reciprocal, then with probability θ the broker receives a signal of r and with probability $(1-\theta)$ he receives a signal of 0. So, his expected posterior belief would be the following:

$$\hat{\phi}(\rho, 1) = \theta\left(\frac{\theta\phi}{\theta\phi + (1-\theta)(1-\phi)}\right) + (1-\theta)\left(\frac{(1-\theta)\phi}{(1-\theta)\phi + \theta(1-\phi)}\right).$$

We next extend this to the case where the broker receives H signals. If a broker receives H signals about a citizen who is actually reciprocal, the probability that h of the signals will be r signalling that the person is reciprocal, while the other $(H-h)$ signals will be 0 signalling that the person is not reciprocal, is:

$$p(h|\theta, H) = \frac{H!}{h!(H-h)!}\theta^h(1-\theta)^{H-h}$$

with $0! \equiv 1$. The broker's expected belief that a reciprocal citizen is in fact reciprocal after receiving H signals, is the weighted average of the belief he would have after receiving h signals of r (and $H-h$ signals of 0), weighted by the probability that he receives h out of H such signals.

$$\hat{\phi}(\rho, H) = \sum_{h=0}^{H} p(h|\theta, H)P(\rho_i = \rho|hr, (H-h)0).$$

The equation for the first term on the right hand side is given above. The equation for the second term on the right hand side is $P(\rho_i = \rho|hr, (H-h)0) = \left(\frac{\theta^h(1-\theta)^{H-h}\phi}{\theta^h(1-\theta)^{H-h}\phi + \theta^{H-h}(1-\theta)^h(1-\phi)}\right)$.

Putting this all together, the brokers' expected posterior belief that a reciprocal citizen is in fact reciprocal after receiving H signals is:

$$\hat{\phi}(\rho, H) = \sum_{h=0}^{H} \frac{H!}{h!(H-h)!}\theta^h(1-\theta)^{H-h}\left(\frac{\theta^h(1-\theta)^{H-h}\phi}{\theta^h(1-\theta)^{H-h}\phi + \theta^{H-h}(1-\theta)^h(1-\phi)}\right) \quad (8)$$

B Variable Construction

B.1 Variables using broker and citizen responses

Broker knows the spouse of the citizen: an indicator that the broker can correctly name the spouse of the citizen.

Broker knows the citizen's years of education: an indicator that the broker accurately states the citizen's years of education within a 3-year margin of error. To estimate this, we cross-checked the citizen's response regarding his years of education with the broker's corresponding estimate.

Broker knows the citizen's amount of land: an indicator variable that the broker correctly states how many hectares of land the citizen owns within a 25% or 1-hectare margin of error. To estimate this, we cross-checked the citizen's response regarding his land ownership with the broker's corresponding estimate.

Broker knows the strength of the citizen's party preference: an indicator that the broker is accurate when stating the strength of the citizen's party preference. The brokers were asked to indicate where they would situate citizens along "feeling thermometers" for both the Colorado and Liberal parties ranging from very cold (0) indicating strong opposition to very hot (40) indicating strong support. If the broker assigned a higher value to the Colorado party and the citizen stated after the election that he supports the Colorado party, we code the broker's response as accurate. We do the same for the Liberal party. If the broker assigns the same value to the citizen's feelings toward both parties and the citizen states after the 2006 election that he supports another party (UNACE or Patria Querida) or no party, we also indicate this as accurate. Note that this is different from official party registration, which is publicly available.

Broker knows the frequency with which the citizen would retaliate: an indicator that the broker accurately states the frequency with which the citizen would retaliate wrongdoing. In particular, citizens were asked the following question: "If someone puts you in a difficult situation, would you do the same to her/him?" Citizens could give answers of always, sometimes, or never. Similarly brokers were asked the same question about each citizen.

Broker knows if citizen generally trusts others in the village: an indicator that the broker accurately knows if the citizen would trust at least half of their village-mates. Citizens were asked what share of their village-mates they trust from 1 (nobody), to 3 (half), to 5 (everyone). Similarly brokers were asked the same question about each citizen.

Absolute age difference: the absolute value of the difference between the broker's and the citizen's ages.

Broker and citizen have the same gender: an indicator that the citizen and broker have the same gender.

Geographical distance between broker's and citizen's residences: geographical distance in kilometers between the broker's and citizen's residences.

B.2 Variables only using broker responses

Broker knows citizen: an indicator variable that the broker states that he knows the citizen.

Broker offered the citizen something: an indicator that the broker states his candidate or his campaign offered the citizen at least one of the following: food, medicines, paying their bills, free plowing of land, or money during the electoral campaign.

Broker approached the citizen: an indicator that the broker states that he approached the citizen during the electoral campaign to speak about the election.

B.3 Variables only using citizen responses

Citizen supports the broker's party: an indicator that, in 2007 a few months after the 2006 elections, the citizen states he supports the party that the broker works for.

Not registered to the broker's party: an indicator that the citizen is not registered to the party for which the broker works, including when the citizen is not registered to any political party.

Experimental reciprocity: computed using data from trust games citizens played in 2002 as described in Finan and Schechter (2012) and as described in Section 4.

B.4 Network variables

***Hearing*:** the expected number of times that broker b hears a piece of information originating from citizen i if information is diffused with probability p for T periods. Following Banerjee et al. (2013), we set T equal to 7, which is the maximum social distance between any citizen and any broker in all the village networks in our sample, and p equal to the inverse of the largest eigenvalue of the adjacency matrix representing the social network in each village.

Diffusion centrality: the sum of *hearing*. That is, the sum of the number of times that an individual would expect to hear pieces of information originating from each of the individuals in the network if information diffuses for T periods. We set T equal to 10, which is the diameter, or maximum social distance between any two individuals in all the village networks in our sample.

Existence of a support pair: an indicator for whether a broker and a citizen share a common friend. Jackson et al. (2012) propose support as a measure of enforcement ability.

Number of support pairs: the number of friends in common between a broker and a citizen.

Transaction tie: an indicator for whether one or more of the following non-political informal transaction ties exist between the broker's and citizen's households: 1) a member of one individual's household provided assistance when a member of the other individual's household fell sick in the year surrounding the election, 2) a member of one individual's household provided a member of the other individual's household with a monetary or in-kind transfer in the year surrounding the election, 3) a member of one individual's household lent money to a member of the other individual's household in the year surrounding the election, or 4) a member of one household states they would go to the other for monetary assistance in times of need.

Degree centrality: the number of individuals to which a broker/citizen is directly connected.

Betweenness centrality: the proportion of shortest paths between any two individuals that pass through the broker/citizen.

Eigenvector centrality: a recursively defined measure such that the centrality of a broker/citizen is proportional to the centrality of the individuals with whom the broker/citizen is directly connected.

Clustering coefficient: the number of actual connections between the nodes within a broker/citizen's neighborhood divided by the total possible number of connections between them.

Absolute difference in diffusion centrality: the absolute difference in the diffusion centrality of the broker and citizen.

Absolute difference in degree centrality: the absolute difference in the degree centrality of the broker and citizen.

Absolute difference in betweenness centrality: the absolute difference in the betweenness centrality of the broker and citizen.

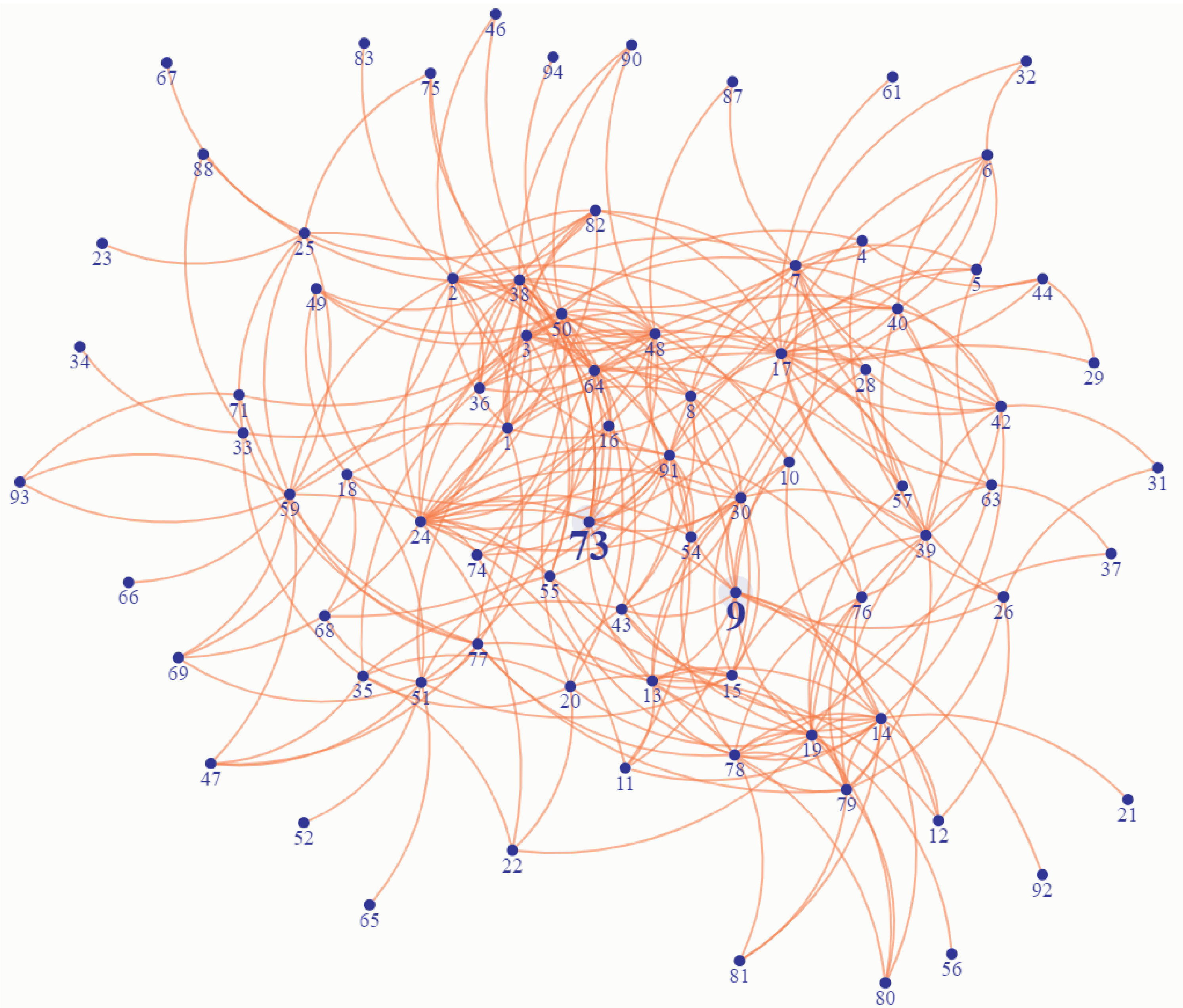
Absolute difference in eigenvector centrality: the absolute difference in the eigenvector centrality of the broker and citizen.

Absolute difference in clustering: the absolute difference in the clustering coefficient of the broker and citizen.

Freeman segregation index (FSI): given individuals affiliated to one of two political parties, $\text{FSI} = 1 - \frac{p}{\pi}$, where p is the observed proportion of between-party connections and π is the expected proportion if connections were generated randomly. It ranges between 0 for a random network and 1 for a network with fully segregated partisan groups.

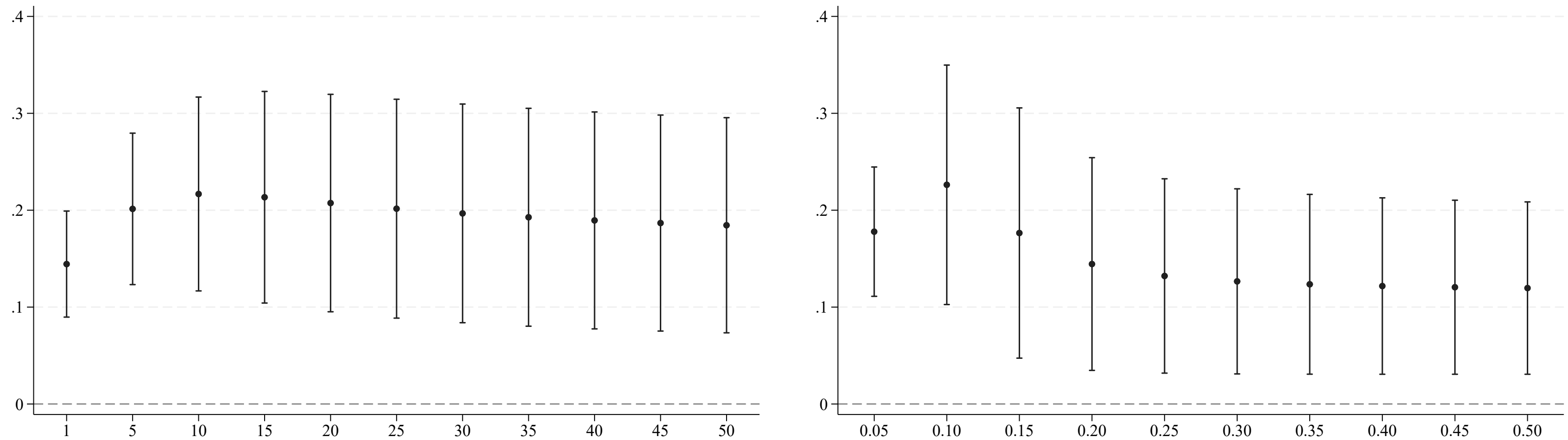
C Additional Tables and Figures

Figure C1: Social network mapping of households in one village



Notes: The graph represents the social network in one of the villages in our sample. It shows all connections between households (brokers and citizens) directly or indirectly sampled within the village. The two brokers live in households 9 and 73, which are labeled in larger bold font.

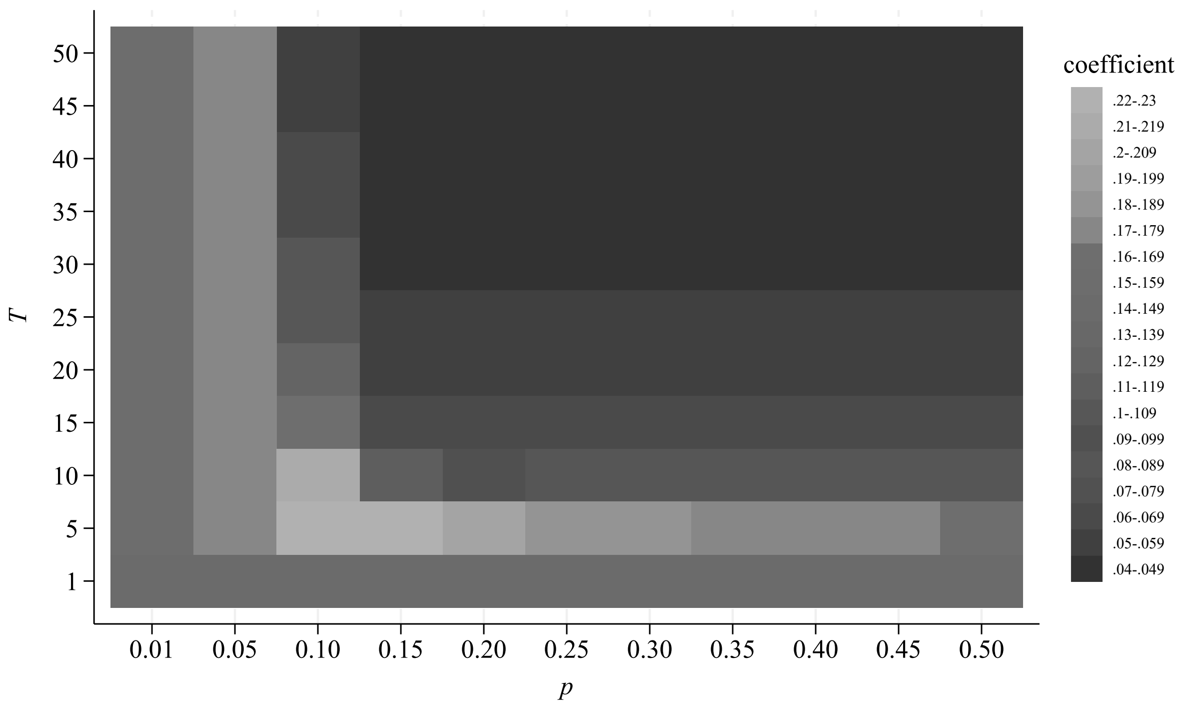
Figure C2: Relationship between *hearing* and vote-buying targeting by choice of T and p (with 95% confidence intervals)



(a) Choice of T (b) Choice of p

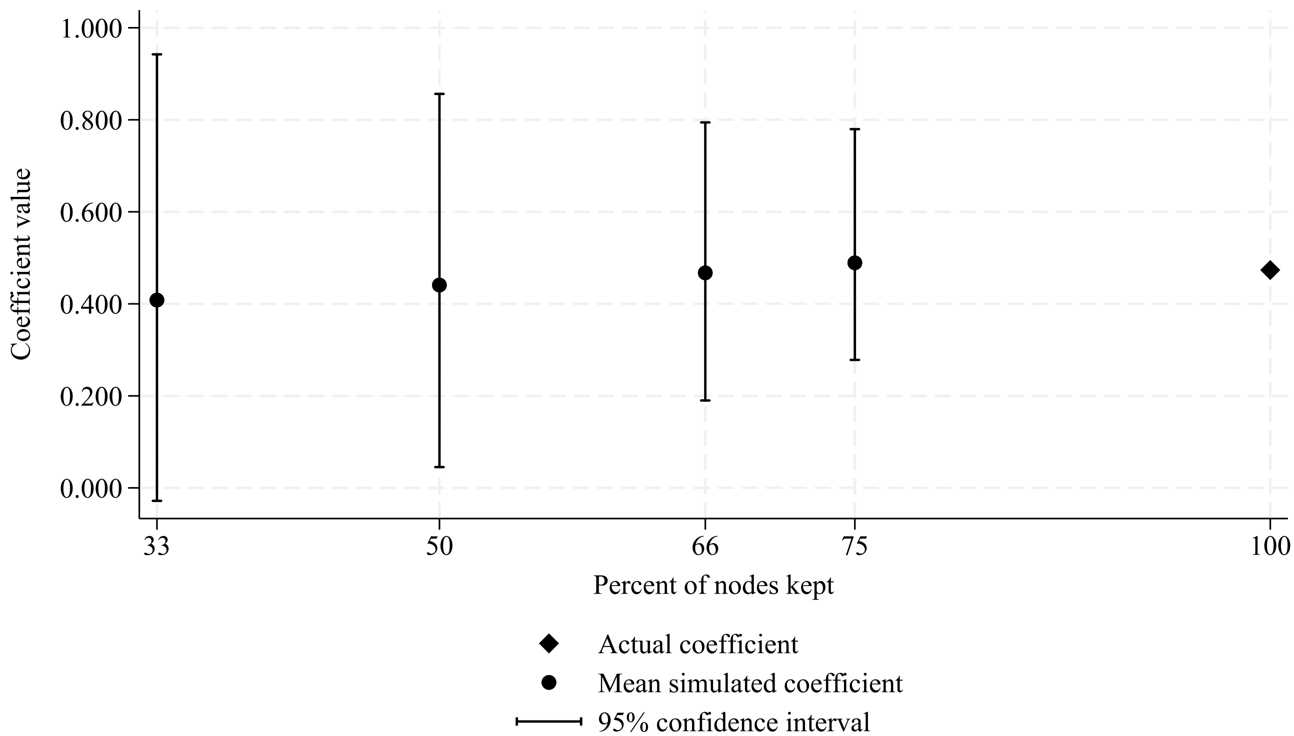
Notes: Estimates in Panel (a) are the coefficient on *hearing* from eleven regressions with a specification equivalent to the one in column (1) of Panel B of Table 2, with the exception that *hearing* is calculated varying T from 1 to 50 at intervals of 5. Estimates in Panel (b) are the coefficient on *hearing* from ten regressions with a specification equivalent to the one in column (1) of Panel B of Table 2, with the exception that *hearing* is calculated varying p from 0.05 to 0.50 at intervals of 0.05. Standard errors are computed using two-way clustering at the broker and citizen levels.

Figure C3: Heat map of *hearing* coefficient magnitude by choice of T and p



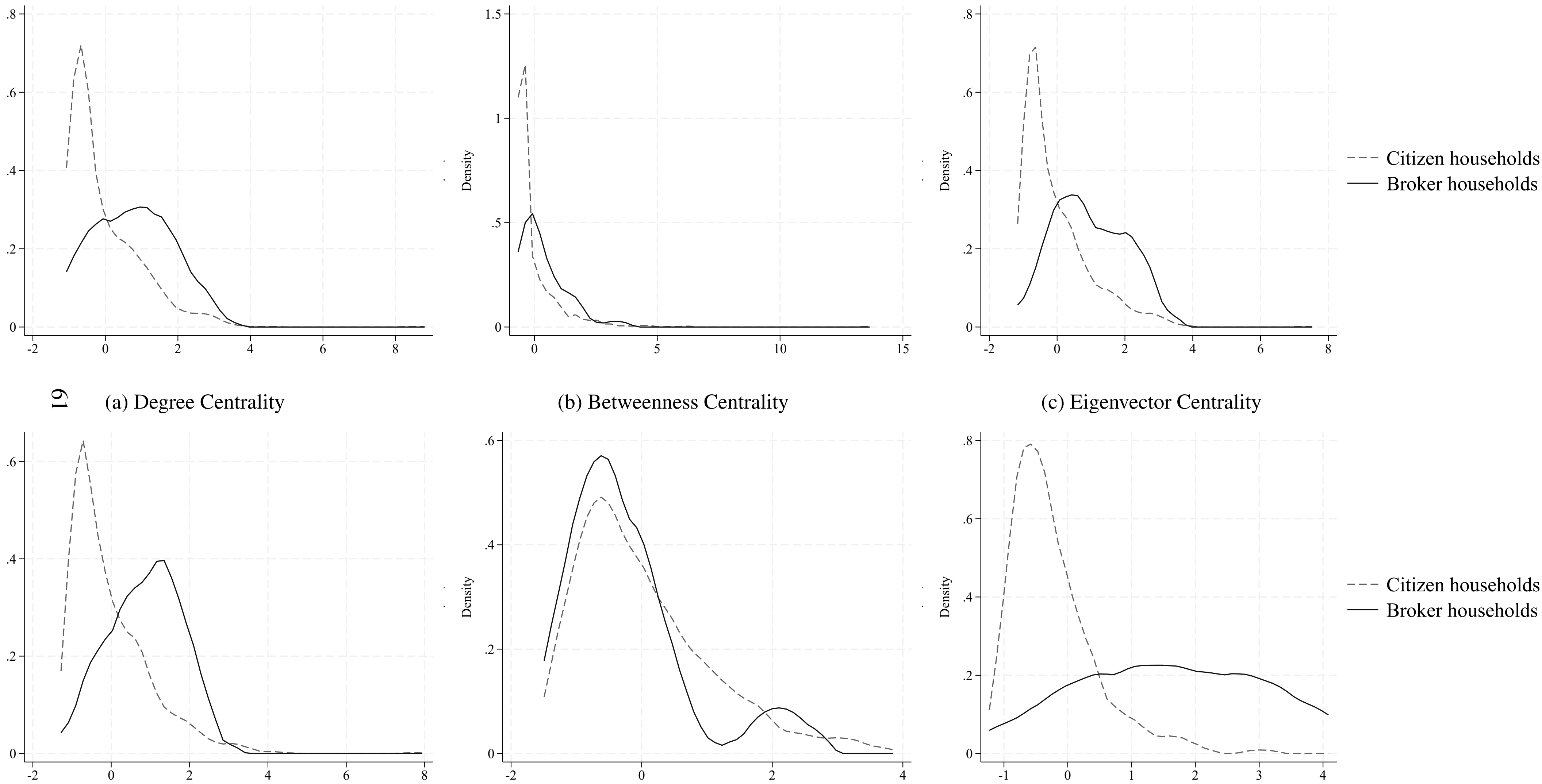
Notes: Coefficient estimates are the coefficient on *hearing* from regressions with a specification equivalent to the one in column (1) of Panel B of Table 2, with the exception that *hearing* is calculated using differing values of T and p. Lighter colors denote larger coefficients.

Figure C4: Coefficient estimates from partial sampling simulation



Notes: The figure shows the actual coefficient on *hearing* in the village with the highest direct sampling rate (91%). It also shows means and 95% confidence intervals of simulated coefficients when keeping only 33, 50, 66, and 75% of the nodes.

Figure C5: Kernel density estimates of household centrality measures



(a) Degree Centrality

(b) Betweenness Centrality

(c) Eigenvector Centrality

(d) Diffusion Centrality

(e) Diffusion Centrality among copartisans

(f) Diffusion Centrality among non-copartisans

s: Each plot shows the Epanechnikov kernel density estimates of standardized (with mean 0 and standard deviation of 1) household centrality measures separately for citizens and brokers.

Table C1: Relationship between *hearing* and constituent components of brokers' knowledge about citizens

	Knows citizen (1)	Knows citizen's spouse (2)	Knows citizen's amount of land (3)	Knows citizen's years of education (4)
Panel A: Covariates index component measures				
Hearing	0.0344***	0.0784***	0.0603**	0.0404**
	(0.0103)	(0.0219)	(0.0281)	(0.0186)
Mean of Dependent Variable	0.8873	0.7725	0.4206	0.8069
Broker FE	X	X	X	X
Citizen FE	X	X	X	X
Observations	932	932	932	932
R^2	0.6359	0.5789	0.5524	0.6090

	Knows strength of citizen's party preference (1)	Knows the frequency with which the citizen would retaliate (2)	Knows whether the citizen generally trusts others in the village (3)
Panel B: Political and Social Preferences index component measures			
Hearing	0.0971***	0.0500*	0.0554***
	(0.0351)	(0.0272)	(0.0138)
Mean of Dependent Variable	0.5933	0.5858	0.6556
Broker FE	X	X	X
Citizen FE	X	X	X
Observations	932	932	932
R^2	0.5362	0.6524	0.7760

Notes: All specifications include broker and citizen fixed effects. *Hearing* is standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table C2: Relationship between *hearing* and brokers' overall knowledge and vote-buying targeting, with T set to the maximum broker-citizen social distance for each village, or restricting to broker observations with direct network data

	Overall knowledge index	Vote-buying targeting index
	(1)	(2)
Panel A: T set to the maximum broker-citizen social distance for each village		
Hearing	0.1459***	0.2379***
	(0.0259)	(0.0449)
Mean of Dependent Variable	-0.0000	-0.0000
Broker FE	X	X
Citizen FE	X	X
Observations	932	932
R^2	0.6384	0.6129

	(1)	(2)
Panel B: Excluding brokers without direct network data		
Hearing	0.2233***	0.4374***
	(0.0501)	(0.0718)
Mean of Dependent Variable	-0.0000	-0.0000
Broker FE	X	X
Citizen FE	X	X
Observations	488	488
R^2	0.7042	0.6411

Notes: All specifications include broker and citizen fixed effects. The outcome variables and *hearing* are standardized (with mean 0 and standard deviation of 1) within each sample. All panels run similar specifications to those in column (2) of Table 2. In Panel A, *hearing* is calculated setting T equal to the maximum social distance between any citizen and broker in the village network. In Panel B, we keep only observations for brokers whose network connections were directly surveyed. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table C3: Placebo test of relationship between *hearing* and non-political transaction ties, by party registration and experimental reciprocity

	Transaction tie							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hearing	0.2601***	0.3899***	0.2630***	0.3889***	0.2457***	0.3780***	0.2535***	0.3780***
	(0.0420)	(0.0809)	(0.0428)	(0.0815)	(0.0517)	(0.0812)	(0.0595)	(0.0850)
Experimental reciprocity			-0.0248*				-0.0431	
			(0.0123)				(0.0293)	
Not reg. to broker's party					0.0036	0.0752	0.0101	0.0748
					(0.0265)	(0.0518)	(0.0275)	(0.0532)
Experimental reciprocity × *Hearing*			-0.0174	0.0100			-0.0175	0.0061
			(0.0204)	(0.0338)			(0.0328)	(0.0500)
Experimental reciprocity × Not reg. to broker's party							0.0259	-0.0089
							(0.0335)	(0.0289)
Not reg. to broker's party × *Hearing*					0.0307	0.0692	0.0268	0.0700
					(0.0396)	(0.0608)	(0.0420)	(0.0621)
Not reg. to broker's party × *Hearing* × Experimental reciprocity							0.0029	0.0065
							(0.0345)	(0.0469)
Mean of Dependent Variable	0.1439	0.1439	0.1439	0.1439	0.1439	0.1439	0.1439	0.1439
Broker FE	X	X	X	X	X	X	X	X
Citizen FE		X		X		X		X
Observations	271	271	271	271	271	271	271	271
R^2	0.4990	0.7440	0.5074	0.7443	0.5005	0.7522	0.5098	0.7526

Notes: All specifications include broker fixed effects. The dependent variable is an indicator for whether a non-political transaction tie exists between the broker's and citizen's households. "Not reg. to broker's party" indicates the citizen is not officially registered to the broker's party. "Experimental reciprocity" is the experimental measure of reciprocity used in Finan and Schechter (2012). The coefficients for experimental reciprocity in columns (4) and (8) are absorbed by the citizen fixed effects. *Hearing* and experimental reciprocity are standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table C4: Relationship between *hearing* and vote-buying targeting, by citizen-diffusion measures

	Vote-buying targeting index			
	(1)	(2)	(3)	(4)
Hearing	0.3569***	0.3598***	0.3353***	0.3324***
	(0.0622)	(0.0729)	(0.0701)	(0.0557)
Citizen's degree centrality × *Hearing*	-0.0696**			
	(0.0311)			
Citizen's diffusion centrality (with $T = 10$) × *Hearing*		-0.0563		
		(0.0409)		
Citizen's eigenvector centrality × *Hearing*			-0.0237	
			(0.0466)	
Citizen's betweenness centrality × *Hearing*				-0.0514*
				(0.0258)
Mean of Dependent Variable	-0.0000	-0.0000	-0.0000	-0.0000
Broker FE	X	X	X	X
Citizen FE	X	X	X	X
Observations	932	932	932	932
R^2	0.6144	0.6137	0.6130	0.6139

Notes: All specifications include broker and citizen fixed effects. The dependent variable is a standardized index (with mean 0 and standard deviation of 1) that takes the sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. The coefficients for citizen's degree centrality, eigenvector centrality, and diffusion centrality are absorbed by the citizen fixed effects. Network measures are standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.

Table C5: Summary statistics of network measures

	Observations	Mean	Median	Standard Deviation
Non-standardized Network Measures				
Broker Network Measures				
Degree centrality	32	7.969	8	4.540
Betweenness centrality	32	0.031	0.022	0.032
Eigenvector centrality	32	0.136	0.120	0.073
Diffusion centrality	32	9.351	9.781	4.247
Diffusion centrality among copartisans	31	2.524	2.033	2.305
Diffusion centrality among non-copartisans	31	6.565	6.507	3.488
Citizen Network Measures				
Degree centrality	1,000	4.585	3	4.304
Betweenness centrality	1,000	0.019	0.004	0.036
Eigenvector centrality	1,000	0.066	0.041	0.070
Diffusion centrality	1,000	5.111	3.642	4.749
Diffusion centrality among copartisans	214	3.175	2.464	2.562
Diffusion centrality among non-copartisans	214	1.975	1.559	1.622
Broker and Citizen Network Measures				
Degree centrality	1,032	4.690	3	4.349
Betweenness centrality	1,032	0.019	0.004	0.036
Eigenvector centrality	1,032	0.068	0.043	0.071
Diffusion centrality	1,032	5.243	3.799	4.789
Diffusion centrality among copartisans	245	3.093	2.417	2.536
Diffusion centrality among non-copartisans	245	2.556	1.825	2.476

Notes: Broker diffusion centrality among copartisans and non-copartisans has one fewer observation because one broker belongs to the UNACE party.

D Multiple Hypothesis Adjustments

Table D1: Multiple hypothesis adjustment for the relationship between *hearing* and brokers' knowledge about citizens and vote-buying targeting

Outcome	**Coeff.**	***p*-values**			
		Unadjusted	Multiplicity adjusted		
		Remark 3.1	Theorem 3.1	Bonferroni	Holm
Overall knowledge index	0.2111	0.0003***	0.0003***	0.0027***	0.0020***
Covariates index	0.1703	0.0007***	0.0017***	0.0053***	0.0027***
Political index	0.1975	0.0003***	0.0003***	0.0027***	0.0017***
Social preferences index	0.1401	0.0027***	0.0073***	0.0213**	0.0080***
Vote-buying targeting index	0.3215	0.0003***	0.0003***	0.0027***	0.0023***
Broker offered citizen	0.0720	0.0040***	0.0040***	0.0320**	0.0040***
Broker approached citizen	0.1833	0.0003***	0.0003***	0.0027***	0.0027***
Support the same party	0.1174	0.0030***	0.0057***	0.0240**	0.0060***

Notes: This table replicates the regressions from Table 2 which include both broker and citizen fixed effects with the multiple hypothesis testing correction methodology developed by List et al. (2019) using the mhtreg Stata command introduced in Barsbai et al. (2021). The coefficients displayed in the "coeff." column are that on *hearing*. The first four outcomes are standardized indices (with mean 0 and standard deviation of 1) that aggregate what the broker knows about the citizen in three categories. The Covariates index aggregates indicators for whether the broker knows the citizen, whether the broker knows the citizen's spouse's name, whether the broker knows how much land the citizen owns, and whether the broker knows the citizen's years of education. The Political index corresponds to an indicator for whether the broker knows the strength of the citizen's party preference. The Social preferences index aggregates indicators for whether the broker knows whether the citizen generally trusts others in the village, and whether the broker knows the frequency with which the citizen would retaliate wrongdoing. The Overall knowledge index aggregates indicators from all three knowledge categories. The vote-buying targeting index is a standardized sum of indicators for whether the broker offered the citizen something during the electoral campaign and whether the broker approached the citizen to talk about the electoral campaign. "Support the same party" is an indicator that the citizen claims to support the broker's party shortly after the election. *Hearing* is standardized. Standard errors use two-way clustering at the broker and citizen levels. *, **, and *** indicate that the corresponding *p*-values are less than 10%, 5%, and 1%, respectively.

Table D2: Multiple hypothesis adjustment for brokers' differential relative position within their social network

Outcome	**Coeff.**	***p*-values**			
		Unadjusted	Multiplicity adjusted		
		Remark 3.1	Theorem 3.1	Bonferroni	Holm
Panel A: Columns (1)-(3) and (6) of Table 8					
Betweenness centrality	0.1046	0.4820	0.4820	1.0000	0.4820
Eigenvector centrality	0.6021	0.0003***	0.0003***	0.0027***	0.0027***
Diffusion centrality	0.5739	0.0003***	0.0003***	0.0027***	0.0023***
Degree centrality	0.3844	0.0020***	0.0037***	0.0160**	0.006**
Betweenness centrality (percentile)	7.8281	0.0573*	0.0907*	0.4587	0.1147
Eigenvector centrality (percentile)	21.4197	0.0003***	0.0003***	0.0027***	0.0017***
Diffusion centrality (percentile)	19.8667	0.0003***	0.0003***	0.0027***	0.0020***
Degree centrality (percentile)	13.4666	0.0003***	0.0003***	0.0027***	0.0013***
Panel B: Columns (4)-(5) of Table 8					
Diffusion centrality among copartisans	-0.1868	0.4257	0.5190	1.0000	0.8513
Diffusion centrality among non-copartisans	2.1281	0.0003***	0.0003***	0.0013***	0.0010***
Diffusion centrality among copartisans (percentile)	-3.2361	0.6817	0.6817	1.0000	0.6817
Diffusion centrality among non-copartisans (percentile)	39.9104	0.0003***	0.0003***	0.0013***	0.0013***
Panel C: Columns (7)-(8) of Table 8					
Degree centrality among copartisans	-0.0195	0.9433	0.9917	1.0000	1.0000
Degree centrality among non-copartisans	1.8767	0.0003***	0.0003***	0.0013***	0.0010***
Degree centrality among copartisans (percentile)	0.5356	0.9493	0.9493	1.0000	0.9493
Degree centrality among non-copartisans (percentile)	33.5084	0.0003***	0.0003***	0.0013***	0.0013***

Notes: This table replicates the regressions from Table 8 with the multiple hypothesis testing correction methodology developed by List et al. (2019) using the mhtreg Stata command introduced in Barsbai et al. (2021). The coefficients displayed in the "coeff." column are that on the broker indicator. The sample for Panel A includes all households directly and indirectly sampled. The sample for Panel B includes all households for which we have party registration data. Copartisanship is determined by whether villagers were registered to the same party. All specifications include village fixed effects and control for whether the household was surveyed directly about their network ties. The first outcomes in each panel are standardized (mean 0, s.d. 1), while the later outcomes in each panel are the within-village percentiles (ranging from 1 to 100) for each corresponding outcome. *, **, and *** indicate that the corresponding p-values are less than 10%, 5%, and 1%, respectively.