Village Social Network Structures and Electoral Competition*

Cesi Cruz†1, Julien Labonne‡2, and Pablo Querubin§3

1University of British Columbia
2University of Oxford
3New York University

September 13, 2017

Abstract.

In this paper, we test whether the structure of village social networks affects political competition in local elections. We use data on 20 million individuals in 15,000 villages of the Philippines. We take advantage of naming conventions to assess intermarriage links between families and reconstruct the family networks in all of those villages. Using data from the 2010 local elections we show that there is less political competition in villages with dense social networks, a result that is robust to controlling for a large number of village and candidate characteristics and to alternative estimation techniques. We then explore the mechanisms behind this effect and present evidence, from a detailed dataset collected in 284 villages after the 2013 local elections, that political influence is more concentrated in denser villages and that this is associated with higher turnout and higher vote margins for the winning candidate.

*We thank Nick Eubank, James Fowler, Yana Gorokhovskaia, Steph Haggard, Alex Hughes, Franziska Keller, and Sarah Shair-Rosenfeld for helpful suggestions. We are also grateful to participants at the UCSD Human Nature Group Workshop and UBC Comparative-Canadian Workshop for feedback. All remaining errors are ours.

†cesi.cruz@ubc.ca
‡julien.labonne@bsg.ox.ac.uk
§pq5@nyu.edu
1 Introduction

Social network analysis has contributed to the literature on electoral competition by linking political outcomes to the ties between politicians and their core bases of electoral support. Politician networks are especially important because they provide access to voters through the ability to monitor and control political brokers at each level (Cruz, Labonne and Querubin 2015; Kitschelt and Wilkinson 2007; Szwarzberg 2011; Stokes et al. 2013; Brusco, Nazareno and Stokes 2004). Accounting for politician networks has also led to more nuanced approaches to understanding the targeting of clientelist benefits (Fafchamps and Labonne 2012; Stokes et al. 2013).

At the same time, there has been considerably less emphasis on the structure of social networks in villages and communities. This is partly because of data limitations: network data is often collected through surveys, which are well-suited for analyses of the position of the individual in the network or the social features of the individual, but considerably more difficult for studies at the community level. Network structures may be important for understanding political behavior as they affect the extent to which individuals in a given community are connected to one another. This can influence information flows, social norms and social sanctions (that can influence people’s political preferences and voting behavior such as turnout). Social network structures may also facilitate elite capture. For example, a candidate can get the votes of many individuals in a community simply by securing the support of a handful of influential leaders who can influence others political views, elicit compliance and diffuse rewards throughout the network (Banerjee et al. 2013).

To estimate the effect of family networks on political competition we use a unique dataset on all 20 million individuals in over 15,000 villages of the Philippines, combined with detailed precinct-level results for the 2010 mayoral elections. The dataset includes information on family names and we use naming conventions in the Philippines to establish ties between families through inter-marriages. Following Padgett and McLean (2006, 2011), a tie between two families exists whenever we observe at least one marriage between members of the two families. We are able to graph the full family network in all 15,000 villages, allowing us to compute network density for all villages in our dataset.

We find that densely connected villages—that is, villages where the ratio of actual ties to potential ties in the family network is high—experience less electoral competition. We argue that this is a result of the concentration of political influence, which allows for political capture in clientelistic political environments.

To explore the specific mechanisms behind this effect, we demonstrate other testable implications of political capture. First, we show that densely connected villages also exhibit higher turnout. Second, we take advantage of a detailed survey implemented shortly after the 2013 local elections in the provinces of Ilocos Norte and Ilocos Sur to show that political influence is more concentrated in denser villages.
We also rule out that our results simply capture the fact that voter’s preferences are more aligned in denser villages. Together those results are consistent with the view that candidates find it easier to negotiate with political elites in denser villages where they exert influence over a larger share of voters.

The remainder of the paper is organized as follows. Section 2 provides some background on family networks and elections. Section 3 discusses the data. Section 4 presents the estimation strategy and results. Finally, Section 5 concludes.

2 Social Networks and Electoral Competition

There is strong theoretical and empirical evidence that voting has a social dimension (see, e.g., Sheingold, 1973; Beck et al., 2002; McClurg, 2003; DellaVigna et al., 2014). Most of the literature has focused on the role of individual social networks or attributes in political participation (see, e.g., Cox, Rosenbluth and Thies, 1998), rather than the overall structure of the network. Two notable exceptions include Fowler (2005)’s study of turnout cascades and Rolfe (2012)’s book on social voting. Even then, because the structure of social networks encompasses a broad range of interactions, it is difficult to establish the precise mechanisms behind the observed relationships between features of the social network structure and outcomes for voting behavior. For example, Rolfe (2012) develops a theory of voter turnout rooted in social context, emphasizing that the features and structure of networks exercise a positive effect on turnout distinct from the aggregate individual characteristics of the people that make up the networks. By contrast, using the proximity of Saint’s Day fiestas in Mexico, Atkinson and Fowler (2014) demonstrate that the increase in social capital associated with the fiestas decreased turnout. In their framework, the fiestas impose obligations on the time and resources of communities and can serve as a substitute means of civic engagement. Similarly, Fowler (2005)’s model of turnout implies that increasing the density of social networks only improves turnout to a certain point; networks that are too dense may lead to declines in turnout.

Studies at the community level tend to focus on the role of social groups in generating social capital, which is linked in the literature to both economic (Knack and Keefer, 1997; Dasgupta and Serageldin, 2001; Feigenberg, Field and Pande, 2013) and political benefits (Putnam, 1993; Tsai, 2007). The literature on social capital suggests that social networks and the relationships between individuals can have important implications for democracy. In a study of Italy, Putnam (1993) links the effectiveness of subnational governments to social networks and civic culture. According to Putnam (1995), social capital is comprised of the “features of social life—networks, norms, and trust—that enable participants to act together more effectively to pursue shared interests.” However, one disadvantage is that social capital can encompass a broad range of attitudes and interactions, making it difficult to establish the precise mechanisms behind the observed relationships between social
capital and political outcomes.

We extend the literature by focusing on the structure of social networks and the potential channels through which village social structures can affect electoral outcomes. Although studies of social capital tend to point to a largely positive effect of network structures, we argue that in clientelistic political environments, dense social network structures in communities can lead to less competitive elections.

Denser social networks can decrease electoral competition through a number of mechanisms. First, we argue that the density of social networks have implications for the concentration of political influence in the village. In their work on social influence, Katz and Lazarsfeld (1955) found that face-to-face interactions were more influential than messages from the media or elites, and that the individuals were most influenced by peers of similar social status. Similarly, research from surveys in the United States identify discussion networks (Huckfeldt and Sprague, 1991) and social cohesion (Huckfeldt et al., 1995) as important factors for the transmission of political information to citizens.

Related to this, dense social networks make it easier for politicians to mobilize voters. Previous research linking social pressure to turnout have found that concerns with social image increase turnout (Gerber, Green and Larimer, 2008; DellaVigna et al., 2014), and turnout can spread through social networks (Nickerson, 2008). These effects also apply to campaigning: voter mobilization efforts are more effective when the canvassers are volunteers (Nickerson, 2008) or members of the neighborhood (Sinclair, McConnell and Michelson, 2013) instead of professionals or outsiders. Dense social ties in a village make it easier for politically influential individuals to consolidate their position within the network.

Second, electoral strategies are often conducted at the group level. Cruz (2013) finds that politicians disproportionately buy votes from individuals that are part of larger social networks, because it allows for group monitoring: politicians can observe the vote choices of the group. Rueda (2014) notes that politicians can use vote tallies at the precinct level to monitor groups of voters for vote buying. In his model, politicians and brokers can ensure cooperation by conditioning future bribes on threshold levels of votes within the group. Similarly, in coercive political environments, different configurations of social networks can affect the costs of monitoring and enforcing voter intimidation. For example, using laboratory experiments in India, Breza, Chandrasekhar and Larreguy (2013) find that adding the ability to punish in a joint investment game improves outcomes, but only when the person with the ability to punish is central in the social network. This suggests that position in the network affects the ability to effectively threaten and punish.

Third, to the extent that politicians can more easily observe the voting patterns of an entire group, being a member of a dense social network can change individual perceptions regarding their choice of vote. Smith, Bueno de Mesquita and LaGatta (2011) model the decision to vote as rational when considering voters as members
of groups that are competing for benefits from politicians. Being part of a social network creates incentives for compliance, to the extent that individuals are concerned about the welfare of the group (Cruz, 2013; Rueda, 2014; Smith, Bueno de Mesquita and LaGatta, 2011). Individuals with a larger network are more concerned with obtaining group-level benefits from politicians (and conversely, avoiding group-level ramifications of falling out of favor with politicians). The literature on reciprocity relies on a much more general concept of reciprocity that can obligate voters to fulfill their end of a vote buying bargain—after all, voters may not have direct ties with the politician or the politician’s brokers. By contrast, an approach focusing on voter social networks relies only on the assumption that individuals will perceive a sense of obligation within their social circle, creating incentives to comply with clientelist mechanisms of political exchange.

3 Families, Social Networks, and Politics in the Philippine Context

Political competition in Philippine municipalities is characterized by strong clientelistic practices organized around family units. As a result, electoral strategies tend to focus on contingent political exchange, such as patronage (Lande, 1996) and vote buying (Cruz, 2013; Khemani, 2011). Since the passage of the 1991 Local Government Code, each municipality in the Philippines is governed by a mayor, a vice-mayor and eight municipal councillors; all elected at-large every three years. Candidates often form coalitions (mayoral and vice-mayoral candidates plus eight municipal council candidates) but voters have to select their choice for each office individually as there is no straight-ticket voting.

The nature of political competition revolves around family alliances (Lande, 1964; Hutchcroft and Rocamora, 2003). A number of municipalities are tightly controlled by so-called political dynasties (Querubin, 2010). Family members often hold office at different levels of government during the same electoral cycle and/or circumvent the three-term limit for local office by taking turns holding the same office (Querubin, 2011). Fegan (2009) argues that the family is a more effective political unit than an individual because its reputation, loyalties, and alliances are transferable from members who die or retire to the younger generations. Corpuz (1965, p 83) also makes reference to the importance of norms of behavior within families: behavior in the family is regulated by ethics and norms that are unwritten and informal, depending for their effectiveness upon internalized sanctions. In particular, one feature of Filipino culture is the concept of utang na loob (literally, “inner debt”), which refers to a debt of gratitude that fosters reciprocity and feelings of social obligation. These norms of loyalty and reciprocity often extend beyond immediate relatives. Thus, if individual A did a favor to me, I feel indebted not only to individual A but also to relatives of individual A and all others to whom individual A expresses loyalty.

Consistent with the previous discussion, there is also qualitative evidence from the Philippines that candidates rely on family members to run their campaigns. For
example, Kerkvliet (2002, p 221) in his ethnography of local politics in a Philippine municipality, argued that being from a larger, well-established extended families gave Tomas Gregorio a reservoir of supporters and advocates or campaigners (mga lider).

Recalling how he organized his campaigns in 1960 and 1964, Tomas pointed towards the areas of San Ricardo lining the road leading to the village of Bagong Sikat and named several cousins, nieces and nephews, in addition to his brothers and sisters, whom he had asked to campaign for him. Corpuz (1965, p 84) also makes explicit reference to this: The network of family ties offers a ready-made political machine in support of electoral candidacies.

Finally, the high levels of cohesion and hierarchy within families often imply that vote-buying transactions can be made directly with family heads who commit to delivering all the votes of their extended families, rather than with individual voters. This facilitates the process of vote-buying.

While our empirical analysis is limited to the Philippines, we believe that many features of the social and political environment are shared by other countries. On the one hand, the family is still the most important social institution in many contemporary societies. Also, strong political dynasties play an important role in other countries such as India, Ireland and Japan amongst others.

4 Data

In this Section we present our various data sources and explain how our main independent variables are computed. The data are from Philippine municipal elections.

4.1 Data Sources

We leverage two main data sources. First, we use data collected for the National Household Targeting System for Poverty Reduction (NHTS-PR). The large-scale household-survey, implemented between 2008 and 2010, collected information on assets, residence characteristics, access to public utilities, and participation in government programs. In addition, the survey reports the gender, age, educational attainment and occupational category of every household member. We have access to the full dataset but focus on the 709 municipalities where full enumeration took place. This leaves us with information on 20 million individuals in about 15,000 villages in 709 municipalities. Importantly, we secured access to the non-
anonymized version of the dataset and have two family names (the middle and last name) for every individual.\footnote{Fernandez (2012) describes the data in more detail.}

Second, we use results from the 2010 municipal elections that were collected from the Commission of Elections (COMELEC) official election results website. For each precinct, we have data on the number of registered voters, the number of individuals who voted and the number of votes received by each candidate. Our two main outcome variables are win margin (vote share of the candidates that received the most votes in that precinct minus vote share of the runner-up in that precinct) and turnout (number of effective voters as a fraction of registered voters). We combine this information with data from the Project of Precincts to match each precinct to a specific village (in the Philippines there is at least one precinct per village).

4.2 Independent Variables

We are able to measure large scale family networks in the Philippines due to naming conventions with three convenient features: (i) within a municipality, a shared family name implies family connections; (ii) each individual carries two family names: their father’s last name and their mother’s maiden name for men and unmarried women and their husband’s last name and their father’s last name for married women; and (iii) names are difficult to change.\footnote{As indicated by Fafchamps and Labonne (2014), there are strict legal constraints on name changes in the Philippines which reduce concerns about strategic name changes. For example, in the majority decision in the case Wang v. Cebu City Civil Registrar (G.R. No. 159966, 30 March 2005, 454 SCRA 155.), that reached the Supreme Court, Justice Tinga wrote: a change of name is a privilege and not a right, so that before a person can be authorized to change his name given him either in his certificate of birth or civil registry, he must show proper or reasonable cause, or any compelling reason which may justify such change. Otherwise, the request should be denied.}

A natural concern with this matching procedure is that individuals from the same province who share a middle name or last name may not necessarily be related to each other.\footnote{This matching procedure will identify almost all existing relatives in the dataset, with the exception of sons-in-law. The main concern is the existence of false positives, or matches that do not correspond to actual relatives.}

The naming structure and distribution of family names in the Philippines can be traced back to the 19th century. In 1849, Governor Narciso Claveria y Zaldúa instructed municipal officials throughout the country to allocate one name (from a catalog with a list of 61,000 surnames) to each family in order to facilitate tax collection.\footnote{Claveria complained that the natives “arbitrarily adopt the names of saints and this practice has resulted in the existence of thousands of individuals having the same surname.” He added: “I saw the resultant confusion with regard to the administration of justice, government, finance, and public order, and the far-reaching moral, civil and religious consequences to which this might lead, because the family names are not transmitted from the parents to their children, so that it is sometimes...} A different set of surnames (often starting with the same letter).
was assigned to each town, and local officials had to assign a different surname to the different family heads. Since then, names have been transmitted through generations according to well-established and enforced naming conventions. As a consequence very common family names are not as prevalent in the Philippines and thus, sharing a family name is very strongly correlated with an actual family tie. This is especially the case within a municipality.

Given the full names of all individuals in an area, we are able to reconstruct all of the ties (edges) in the family network by examining the joint occurrences of middle and last names. As noted above, each individual maintains two family names: their father’s name and either their mother’s maiden name or their husband’s name, in the case of married women. Thus each individual’s set of family names indicates an intermarriage between the two families—either in their generation (in the case of married women) or their parents’ generation (in the case of men and single women). As a result, we are able to observe ties between families merely by the occurrence of the names within an individual.

For example, Figure 1 below depicts the family network that can be drawn from a partial list of relatives of the current Philippine President, Benigno Cojuangco Aquino. His middle name is his mother’s maiden name, Cojuangco, and his last name is his father’s last name, Aquino. Just by observing his full name, we are able to infer a tie between his mother’s family, the Cojuangcos, and his father’s family, the Aquinos. To use one example from his sisters, Aurora Aquino Abellada is married, so we can draw a tie between the Aquino family and the family of her husband, as indicated by Aurora’s last name. Similarly, we can show a tie between the Aguirre and Aquino families by adding the name of President Aquino’s cousin, Bam Aguirre Aquino. Last, the names of President Aquino’s cousin Gilberto Cojuangco Teodoro and uncle Jose Sumulong Cojuangco show ties between the Cojuangco family and the Teodoro and Sumulong families, as well as an indirect tie to the Prieto family through Gilberto’s wife Monica Prieto Teodoro.

impossible to prove the degrees of consanguinity for purpose of marriage, rendering useless the parochial books which in Catholic countries are used for all kinds of transactions.” See National Archives of the Philippines (1973).

Fafchamps and Labonne (2015) compute a Herfindhal Index of name heterogeneity for a large sample of municipalities in the Philippines. A value of 0 indicates that there is only one family name in the province, while a value very close to 1 suggests a very low concentration of family names. The overall Herfindhal Index for the municipalities in their sample is greater than 0.999. The most common surname in their data, De La Cruz, is used by only 0.32 percent of individuals. By contrast, they show that the prevalence of common names is much higher in other countries in the region. The percentage of individuals that uses the most common surname is 7.25% in China, 5.5% in India, 11% in Taiwan and 38% in Vietnam.
Our measure of network structure is the density of family ties in the barangay, which is the number of marriage ties as a share of the total potential ties in the barangay. Given the family network $f$, where $t$ indicates the number of marriages between families and $n$ is the total number of families in the network, the density of $f$ is given by the following equation:

$$\text{Density}_f = \frac{t}{n(n-1)/2}$$ (1)

Network density is a way of assessing the connectedness of a network [Jackson, 2010; Wasserman and Faust, 1994]. Figure 2 is a fully connected network in which there are intermarriages linking all of the five families, or 10 ties total. Using Equation 1 above, the density of this network can be calculated by dividing the total ties in the network by the total potential ties in the network, which in this case is also 10 (calculated using the number of families: $5(5-1)/2 = 20/2 = 10$), yielding a density of 1.
Figure 2: Fully Connected Family Network.

Figure 3 shows densities for different configurations of the network of five families. In Figure 3.1, there are 8 intermarriages linking the five families. Using Equation 1 above, the density of this network can be calculated by dividing the total ties in the network by the total potential ties in the network, which in this case is 10 (calculated using the number of families: $5(5-1)/2 = 20/2 = 10$), yielding a network density of .8 as indicated.

Figure 3.2 shows the same five families, but with only 4 ties among them. In this example, family A is linked by marriage to all of the other four families, but none of these four families have marriage ties among them. Because the number of families is the same as in 3.1, the total potential ties in this network is also 10, yielding a network density of .4 as indicated.

Last, Figure 3.3 shows the five families without any marriage ties among them. In this example, the actual ties are 0, making the network density 0 as well. This is presented just as an example: because an intermarriage tie can be assessed using the combination of middle and last names of an individual and there are no villages without people. By definition we do not observe any networks with zero density in our sample.
One property of social networks that applies especially to family networks is what Lazarsfeld and Merton (1954) termed homophily, which refers to the tendency for people to associate with individuals that are similar to them. This has been a well-established feature of social networks\[^{10}\] that has linked similarities based on characteristics such as race, educational background, profession, religion, and age to social network formation.\[^{11}\] This has also been observed in family networks through assortative marriage (see, e.g., Kalmijn, 1998). There is a substantial literature in sociology and anthropology detailing the tendency for individuals to choose marriage partners based on similar educational (Mare, 1991; Blossfeld and Timm, 2012), religious (Ernest W. Burgess, 1943), and ethnic (Qian, Glick and Batson, 2012) backgrounds. Dense family networks at the village level may therefore reflect homogeneity, making it important to distinguish between effects of similar education and occupational backgrounds with the effect of the family network structure itself. As a result, we use a comprehensive set of indicator variables for different educational and occupational characteristics.

Two extreme examples of family networks are presented in Figure 4. In the first example, Barangay Sapat in the municipality of Pasuquin is primarily a military base, where families tend not to stay in the barangay for more than the length of their posting, let alone for generations. As a result, we would expect Sapat to have fewer family ties. Sapat has 92 families in residence, but only 59 ties among them, yielding one of the lowest densities in the sample.

By contrast, Barangay Manarang in the municipality of Vintar is more typical of Filipino villages (it was chosen because of similar demographics and number of families residing in the village). In Manarang, there are 103 families with 164 ties

---

\[^{10}\]See, e.g., Kossinets and Watts (2009).

\[^{11}\]See, e.g., Miller McPherson (2001) for a comprehensive review of the literature.
among them, yielding a value of .03 for the village density.

Figure 4: Barangay Examples

In Table 1 we present some descriptive statistics for our main variables of interest. The average density in our sample is .009, implying that on average a family is married to less than 1% of the potential families in a village with which a marriage tie could have been established. However, notice that we do have in our sample fully connected villages in which every family has at least one marriage to every other family in the village. Votes in a precinct are quite concentrated in the winning candidate, with a winning margin of 34 percentage points. Finally, turnout is on average 73% though we also observe substantial variation in this variable across the different precincts (note that voting is not compulsory in the Philippines).

Table 1: Village Density, Margin of Victory, and Turnout

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.009</td>
<td>0.018</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Win Margin</td>
<td>0.336</td>
<td>0.268</td>
<td>0.001</td>
<td>1.000</td>
</tr>
<tr>
<td>Turnout</td>
<td>0.726</td>
<td>0.131</td>
<td>0.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

5 Results and Discussion

In this section we explore empirically the relationship between village network density and electoral competition. To do so we estimate OLS regressions of the form:
\[ \text{WinMargin}_{pvm} = \alpha + \beta \text{Density}_{vm} + \gamma X_{vm} + \eta_m + \epsilon_{pvm} \quad (2) \]

where \( \text{WinMargin}_{pvm} \) is the difference in vote share between the top two candidates in the precinct in the 2010 mayoral elections in precinct \( p \), located in village \( v \) and municipality \( m \). Our main independent variable of interest is \( \text{Density}_{vm} \) which measures the density of the family network in village \( v \). \( X_{vm} \) is a set of village-specific characteristics and \( \eta_m \) corresponds to a full set of municipality fixed effects included in all specifications. Finally, \( \epsilon_{pvm} \) is the usual idiosyncratic error term and standard errors account for potential correlation within municipalities.

The baseline estimates based on equation (2) are reported in Table 2. In column 1 we present the estimate of \( \beta \) without any additional village controls (only including the municipality fixed effects). The estimate is positive and statistically significant. This implies that a standard deviation increase in network density leads to an increase in the win margin of 1.7 percentage points (or a 5% increase relative to the mean win margin of 33%).

<table>
<thead>
<tr>
<th>Density</th>
<th>(1) 0.87***</th>
<th>(2) 0.61**</th>
<th>(3) 0.46***</th>
<th>(4) 0.38**</th>
<th>(5) 0.40**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.24)</td>
<td>(0.18)</td>
<td>(0.19)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Observations</td>
<td>16,858</td>
<td>16,858</td>
<td>16,858</td>
<td>16,858</td>
<td>16,858</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.503</td>
<td>0.507</td>
<td>0.510</td>
<td>0.511</td>
<td>0.511</td>
</tr>
</tbody>
</table>

Notes: Results from precinct-level regressions with municipal fixed-effects. The dependent variable is win margin in the 2010 mayoral elections. Regressions control for village-level average age, average length of stay in the village, gender ratio and whether the village is classified as rural (Columns 2-5), education levels in the village (Columns 3-5), occupation in the village (Columns 4-5) and average per capita income and poverty incidence (Column 5). The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

One natural concern with the specification in column 1 is that the competitiveness of elections may be correlated with other village characteristics that also affect political behavior. To address this concern in columns 2-5 we include a series of village characteristics. In column 2 we include basic demographic characteristics such as average age of residents of the village, average number of years living in the village, gender ratio as well as an indicator variable for whether the village is rural. This addresses the concern that villages with older citizens, or citizens that have been living for a long time in the village may be denser, and their citizens more interested in local politics and thus more likely to vote for specific candidates.
In columns 3-5 we control for various socio-economic characteristics of the village, more concretely their educational and occupational background as well as different measures of income. For example, if more highly educated villages tend to be denser, and educated people are more likely to vote for the same candidate, this could bias our estimate of $\beta$. More specifically we control for the number of village residents in each of 17 educational categories (columns 3-5), 11 occupational categories (columns 4-5), and for their income per capita and poverty incidence (column 5). Reassuringly, while the size of the point estimates decreases, they remain statistically significant. This suggests that our density measure is not simply confounding other socio-economic characteristics of the village.

In Table A.1 in the appendix, we present additional robustness checks for our main result. All of the regressions include the full set of controls included in column 5 of Table 2, with the exception of the educational and occupational dummies for column 3. In column 1, we eliminate all municipalities in the Autonomous Region of Muslim Mindanao (ARMM) where the distribution of family names makes it more likely for two unrelated individuals to share the same family name. In addition, local politics in that region are less competitive and revolve around lineages whose power can be traced back to historical rulers, raising the concern that our results are merely capturing the effects of network density there. However, our results hold even when we exclude ARMM municipalities from our sample.

In column 2 we control linearly for the number of registered voters in the precinct, while in column 3 we do so flexibly by including almost 100 indicator variables for whether the number of registered voters in the precinct is lower than 10, lower than 20, and so on in intervals of 10 until 1,000. This allows us to address the concern that our estimates simply capture a correlation between precinct size and density, and precinct size and margin of victory. This is especially important when using network density as our measure, as it is very sensitive to the size of the networks. While our point estimates in this case decrease in magnitude they nonetheless remain statistically significant, a result that is robust to different ways of flexibly controlling for population.

In column 4 we include a measure of network distance between the incumbent mayor and the barangay captain in that village. Barangay captains are the most important elected authority at the village level. Moreover, they are usually influential members of their community and often play the role of political brokers who mobilize voters in support of candidates in local and national elections. In places in which the mayor and barangay captain are very closely connected, barangay captains may have access to more resources and higher incentives to mobilize village

---

12 The different educational categories correspond to different years of education, from zero (no grade completed) to 17 (having a graduate degree).

13 Examples of occupational categories are Government Officials, Professionals, Farmers, Clerks, Laborers and Unskilled Workers, amongst others.

14 Distance here is calculated as the number of edges (links) in the shortest path between the incumbent mayor’s family and the barangay (village) captain’s family in the village.
residents to turn out to vote for the incumbent mayor (or one of her/his relatives). However, our estimate in column 4 is larger than our baseline estimates in Table 3 suggesting that the positive correlation between density and margin of victory does not simply confound the role of distance between the barangay captain and the mayor.

We also control for several village-specific characteristics of the barangay captain (column 5), the incumbent mayor (column 6), and other challengers in the mayoral race (column 7). In column 8 we include the characteristics of all of these politicians simultaneously. More concretely we control for different measures of each of these politician’s families in that village such as their centrality in the village’s network (eigenvector centrality), number of individuals with the family name, the number of women with the family name, education levels in the family (number of members of the family in each of the 17 educational categories) and occupational background of the family (number of members of the family in each of the 14 occupational categories). Intuitively, in places in which the barangay captain, incumbent mayor or challengers are important members of the community (either because their family is numerous or prestigious in that village) we expect a larger number of people voting for them (Cruz, Labonne and Querubin, 2015). If villages where these politicians are prominent also tend to be more dense, this could explain our baseline results. Reassuringly, our estimates remain relatively similar and become, if anything, larger once we account for politician’s characteristics in the village.

In sum, the estimates reported in Table 2 suggest that family network density has a negative effect on electoral competition in the village. This effect does not seem to confound other demographic or socio-economic characteristics of the village as a whole, or the prominence and importance of barangay and municipality politicians in the village.

5.1 Evidence on Mechanisms

In addition to establishing the importance of network density for electoral competition, it is equally important to identify the mechanisms behind the relationship. We argue that denser networks facilitate clientelistic practices at the village level. Influential members of the community, connected to multiple families in villages with dense social networks, could deliver many votes to candidates aligned with them. This could result in less competitive political environments that reflect the influence and mobilization efforts of few individuals. This is consistent with literature showing that voter mobilization is more effective under social pressure (Gerber, Green and Larimer, 2008; DellaVigna et al., 2014), and when canvassers are volunteers (Nickerson, 2008) or members of the neighborhood (Sinclair, McConnell and Michelson, 2013). In this section we present evidence consistent with

15 For more details on the eigenvector centrality measures see Cruz, Labonne and Querubin (2015).
this interpretation.

To do so, we first estimate the effect of density on voter turnout. We report estimates of $\beta$ in Table 3 where the specifications in columns 1-5 (set of village characteristics used as controls) are analogous to those of Table 2. These estimates suggest that village social network density is positively associated with voter turnout. The magnitude of the coefficient suggests that a one standard deviation increase in village density leads to an increase in turnout of roughly 1 percentage point. This suggests that villages with denser family networks have higher rates of turnout that yield votes concentrated in a single candidate. This result is robust across specifications and does not seem to confound demographic or socio-economic characteristics of the village. The point estimates are remarkably stable across specifications.

Table 3: Village Density and Turnout

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.39***</td>
<td>0.40***</td>
<td>0.42***</td>
<td>0.38***</td>
<td>0.38***</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Observations</td>
<td>18,079</td>
<td>18,079</td>
<td>18,079</td>
<td>18,079</td>
<td>18,079</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.515</td>
<td>0.517</td>
<td>0.518</td>
<td>0.520</td>
<td>0.520</td>
</tr>
</tbody>
</table>

Notes: Results from precinct-level regressions with municipal fixed-effects. The dependent variable is turnout in the 2010 mayoral elections. Regressions control for village-level average age, average length of stay in the village, gender ratio and whether the village is classified as rural (Columns 2-5), education levels in the village (Columns 3-5), occupation in the village (Columns 4-5) and average per capita income and poverty incidence (Column 5). The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

In Table A.2 in the appendix, we present some additional robustness checks for the results on turnout, using the same specifications as the robustness checks for the main results in Table A.1. In all of these additional estimations, our results for turnout hold.

Second, in order to link village social network structures with the social mechanisms for mobilizing voters, we explore the relationship between network density and the concentration of political influence. Because detailed information on political influence is not available for our main sample, we use data from a survey implemented shortly after the 2013 mayoral elections. The sample covers 3,408 households in 284 villages in 12 municipalities in the provinces of Ilocos Norte and Ilocos Sur. Importantly, the survey collected detailed information on sources of

---

16 More information on the survey is available in Cruz, Keefer and Labonne (2014).
political influence in the village, including allowing respondents to name up to five individuals whose political views and opinions they respected.

Using this dataset, we graph the network of political influence, by drawing connections between survey respondents and the individuals they reported as influential in the village network. We then measure concentration of political influence by assessing differences in political influence among individuals in the network (see equation 6 below). Low concentration scores imply network structures in which individuals are similarly influential, while high scores imply network structures in which few individuals are much more connected than all of the other individuals. These measures allow us to distinguish villages with a concentrated source of political influence from those with multiple sources of influence or those in which influential individuals are fairly equally positioned.

Given the network in village $v$, where $c^*$ indicates the individual with the most political influence (calculated by counting the number of respondents who named this individual as a source of political influence), take the sum of the differences between $c^*$ and all other individuals, divided by the theoretical maximum value of this sum for the network (i.e. a situation in which only one individual is reported as influential by all others).

$$Concentration_{v} = \frac{\sum_i [c^* - c_i]}{\max_{c}(\sum_i [c^* - c_i])}$$

In Table 4 we report estimates of equation (2) on the Ilocos sample where we use $Concentration_{v}$ as the dependent variable. Consistent with our interpretation we find that in villages with denser family networks there is a higher concentration of political influence on relatively few individuals. The point estimate in column 1 is positive and statistically significant at the 5% level. The point estimate remains essentially unchanged in column 2 when we control for whether the village is classified as rural, average education, age, household size, and length of residence, as well as the share of population that is female, receives remittances from abroad, and benefits from a CCT program.

17More details about these networks of political influence are available in the appendix. Note that unlike the family networks that we use in the main analysis, these networks are directed: individuals in the village are influential if other people named them as sources of influence and not related to how many people they named as influential themselves. By contrast, the family networks are undirected in that a marriage tie that exists between two families contributes as a tie for both families.

18While we do not report our results here, we can replicate the regression of density on turnout in the Ilocos sample as well.
Table 4: Density and Concentration of Political Influence

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>6.83**</td>
<td>7.00**</td>
</tr>
<tr>
<td></td>
<td>(2.91)</td>
<td>(2.89)</td>
</tr>
<tr>
<td>Additional Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>284</td>
<td>284</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.185</td>
<td>0.234</td>
</tr>
</tbody>
</table>

Notes: Results from village-level regressions with municipal fixed-effects. Dependent variable is the concentration of political influence (the difference in the number of respondents reporting being influenced by the most influential individual compared to all other influential individuals, accounting for the total number of potential people influenced). Column 1 is a bivariate regression and Column 2 controls for whether the village received the PPCRV experiment, is classified as rural, average education, age, household size, and length of residence, as well as the share of population that is female, receives remittances from abroad, and benefits from a CCT program. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Finally, we rule out one alternative interpretation of our results, namely that voters in denser villages have more aligned preferences and thus are more inclined to vote for the same candidate, leading to less competitive races. In the Ilocos survey, respondents were asked to report the share of the Local Development Fund that they would like the municipal government to spend on different sectors such as health, education, emergencies, water, roads, community facilities, economic programs, agriculture, peace and festivals. This gives us an individual level measure of respondent preferences on these different policy items. We then take the standard deviation of individual responses for each policy item, to generate a village-level measure of preference alignment (with a lower standard deviation indicating more agreement/alignment within the village on the importance of that specific policy item).

In Table 5, we report regressions of the standard deviation of each policy item, on our network density measure. The sign of the coefficients varies across different policy items and effects are only statistically significant at the 10% level for roads (where density is associated with less policy alignment) and for peace (where density is associated with more policy alignment). Thus, the evidence available does not suggest that our estimates simply reflect the effect of policy alignment in voters in denser villages, but rather the role of the concentration of political influence.

19However, it is not obvious that more aligned preferences should lead to less competitive races as all candidates may tend to converge to such set of preferences, making elections potentially more, and not less competitive.
20Every year, each municipality receives transfers from the central government and mayors are encouraged to allocate 20 percent of the transfers to development projects (i.e. the so-called LDF).
Table 5: Density and Alignment of Policy Preferences

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Health</td>
<td>Education</td>
<td>Emergencies</td>
<td>Water</td>
<td>Road</td>
<td>Com. Faci</td>
<td>Econ. Prog.</td>
<td>Agriculture</td>
<td>Peace</td>
<td>Festivals</td>
</tr>
<tr>
<td>Density</td>
<td>-54.89</td>
<td>-34.30</td>
<td>-41.08</td>
<td>57.46</td>
<td>90.97*</td>
<td>-28.94</td>
<td>-106.76</td>
<td>99.67</td>
<td>-40.12*</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>(44.41)</td>
<td>(68.21)</td>
<td>(49.52)</td>
<td>(60.05)</td>
<td>(46.64)</td>
<td>(27.98)</td>
<td>(76.94)</td>
<td>(132.21)</td>
<td>(18.90)</td>
<td>(33.10)</td>
</tr>
<tr>
<td>Observations</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
<td>284</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.146</td>
<td>0.082</td>
<td>0.164</td>
<td>0.119</td>
<td>0.110</td>
<td>0.114</td>
<td>0.127</td>
<td>0.130</td>
<td>0.191</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Notes: Results from village-level regressions with municipal fixed-effects. Dependent variable is the standard deviation of the share of the LDF that village respondents would allocate to the respective policy item. We control in all columns for whether the village received the PPCRV experiment, is classified as rural, average education, age, household size, and length of residence, as well as the share of population that is female, receives remittances from abroad, and benefits from a CCT program. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.
6 Conclusion

In this paper we explore the effect of social network structure, in particular, network density, on electoral competition in local elections in the Philippines. We find that network density is robustly associated with less political competition. We further show that network density is associated with higher turnout. Our evidence suggests that this may be driven by a higher concentration of political influence in relatively few individuals whose political interests play an important role in shaping people’s voting decision.

Our findings have a number of important implications. First, they highlight the importance of studying social network structures, and not only the network position or connections of specific individuals in the community. Using a unique dataset and approach, we are able to execute the first large-scale study of village network structures than has been previously possible using traditional survey-based methods of collecting social network data. We find compelling evidence that network structures shape the interaction and flow of information amongst citizens. Social pressure and social sanctions in denser networks may affect people’s likelihood to vote as well as their vote choice. More importantly however, influential leaders in the community may take advantage of dense social networks to gather widespread support for candidates aligned to them. This may facilitate clientelistic practices that are often subject to agency problems such as monitoring that are easier to enforce in tight social networks. The positive correlations between network density, concentration of votes, turnout, and concentration of political influence we document in this paper is highly suggestive of a higher prevalence of elite dominance in villages with dense social networks.

Social networks are associated with a host of positive outcomes in established democracies, such as voter turnout, political interest and knowledge, and broader civic engagement. In emerging or consolidating democracies, however, politicians can use the same social network structures to circumvent normal channels of political competition by mobilizing voters through networks of political influence. Our paper provides first evidence of how local social network structures can hamper electoral competition across the developing world.
1 Appendix

1.1 Degree Centrality

Degree centrality is the simplest measure of individual position, counting the number of inward ties that an influential individual has (number of times that someone reported the individual as influential). This is a raw measure of the total number of times the individual was indicated as influential.

\[ \text{Degree}_i(f) = \sum F_{ij} \] (4)

where \( F \) is the adjacency matrix of the village political influence network \( f \), such that \( F_{ij} = 1 \) if there is a tie between nodes \( i \) and \( j \), and 0 otherwise.

1.2 Eigenvector Centrality

Eigenvector centrality accounts not only for the number of ties, but also whether these ties are themselves well connected (Bonacich, 1972, 1987). Eigenvector centrality is computed recursively such that the centrality of an individual in the village influence network is proportional to the sum of centrality scores of the people who indicated him or her as influential:

\[ \text{Eigenvector}_i(f) \propto \sum F_{ij} \cdot \text{Eigenvector}_j(f) \] (5)

where \( F \) is the adjacency matrix of graph \( f \), such that \( F_{ij} = 1 \) if there is a tie between nodes \( i \) and \( j \) and 0 otherwise. This weights all of the ties to \( i \) by the connectedness of the tie (Bonacich, 1972, 1987).

1.3 Centralization Measures

Centralization measures assess the structure of the family network as a whole by accounting for the differences in centrality among families. These measures can be calculated using any standard centrality measure. We use two common forms: degree centralization and eigenvector centralization (described above).

Given the village influence network \( f \), where \( c^* \) indicates the individual with the highest centrality value (calculated using either degree or eigenvector centrality as described), take the sum of the differences between \( c^* \) and all other individuals, divided by the theoretical maximum value of this sum for the network (i.e. a situation in which one person is reported influential by everyone else).
Centralization$_f = \frac{\sum_i [c^* - c_i]}{\max_j(\sum_i [c^* - c_i])}$

A centralization score of 1, for example, indicates a star network, in which one person is reported as influential by all other residents in the village, and no other individuals are reported as influential.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.78***</td>
<td>0.39**</td>
<td>0.27*</td>
<td>1.76***</td>
<td>0.83**</td>
<td>0.59**</td>
<td>0.41**</td>
<td>0.88**</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.19)</td>
<td>(0.15)</td>
<td>(0.55)</td>
<td>(0.38)</td>
<td>(0.25)</td>
<td>(0.19)</td>
<td>(0.44)</td>
</tr>
<tr>
<td>ARMM Included</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Barangay Population</td>
<td>No</td>
<td>Linearly</td>
<td>Flexibly</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Distance Inc. &amp; Captain</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Captain characteristics</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Incumbent characteristics</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Challenger characteristics</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>14,953</td>
<td>16,858</td>
<td>16,858</td>
<td>8,908</td>
<td>14,883</td>
<td>12,482</td>
<td>16,552</td>
<td>10,780</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.499</td>
<td>0.511</td>
<td>0.516</td>
<td>0.504</td>
<td>0.528</td>
<td>0.511</td>
<td>0.507</td>
<td>0.522</td>
</tr>
</tbody>
</table>

Notes: Results from precinct-level regressions with municipal fixed-effects. The dependent variable is the margin of victory in the 2010 mayoral elections. Regressions control village-level average age, average length of stay in the village, gender ratio and whether the village is classified as rural, average per capita income and poverty incidence. Columns 1, 2, and 4-6 add controls for education levels in the village and occupation in the village. The characteristics included for captain, incumbent and challenger (cols. 5-8) are: betweenness and eigenvector centrality, number of individuals with the family name, the number of female with the family name, education levels in the family, occupation in the family. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.
### Table A.2: Robustness Checks for Turnout I [Alternative samples and additional controls]

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Density</strong></td>
<td>0.60***</td>
<td>0.34***</td>
<td>0.23***</td>
<td>1.19***</td>
<td>1.14***</td>
<td>0.49***</td>
<td>0.38***</td>
<td>1.14***</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.11)</td>
<td>(0.07)</td>
<td>(0.33)</td>
<td>(0.25)</td>
<td>(0.14)</td>
<td>(0.11)</td>
<td>(0.32)</td>
</tr>
<tr>
<td><strong>ARMM Included</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Barangay Population</strong></td>
<td>No</td>
<td>Linearly</td>
<td>Flexibly</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Distance Inc. &amp; Captain</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Captain characteristics</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Incumbent characteristics</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Challenger characteristics</strong></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>16,021</td>
<td>18,079</td>
<td>18,079</td>
<td>9,590</td>
<td>15,964</td>
<td>13,403</td>
<td>16,866</td>
<td>10,797</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.563</td>
<td>0.522</td>
<td>0.527</td>
<td>0.543</td>
<td>0.532</td>
<td>0.504</td>
<td>0.490</td>
<td>0.472</td>
</tr>
</tbody>
</table>

Notes: Results from precinct-level regressions with municipal fixed-effects. The dependent variable is turnout in the 2010 mayoral elections. Regressions control village-level average age, average length of stay in the village, gender ratio and whether the village is classified as rural, average per capita income and poverty incidence. Columns 1, 2, and 4-6 add controls for education levels in the village and occupation in the village. The characteristics included for captain, incumbent and challenger (cols. 5-8) are: betweenness and eigenvector centrality, number of individuals with the family name, the number of female with the family name, education levels in the family, occupation in the family. The standard errors (in parentheses) account for potential correlation within municipality. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.
References


Cruz, Cesi. 2013. “Social Networks and the Targeting of Vote Buying.” *Annual Meeting of the American Political Science Association* .


Cruz, Cesi, Philip Keefer and Julien Labonne. 2014. “Incumbent advantage, voter information, and vote buying.” *mimeo* .
URL: https://books.google.ca/books?id=6PZ8bvQQmxEC

URL: http://www.nber.org/papers/w19832


URL: http://restud.oxfordjournals.org/content/80/4/1459.abstract


URL: http://www.jstor.org/stable/2111668

Princeton University Press.


Rueda, Miguel. 2014. “Buying Votes with Imperfect Local Knowledge and a Secret Ballot.” *Journal of Theoretical Politics*.


