Politician Family Networks and Electoral Outcomes: Evidence from the Philippines*

Cesi Cruz†
Julien Labonne‡
Pablo Querubin§

March, 2015

Abstract

We demonstrate the electoral importance of politician family networks and provide evidence of the mechanisms behind the relationship. We use a 20 million person dataset, allowing us to reconstruct intermarriage networks for over 15,000 villages in 709 municipalities in the Philippines. We show that politicians are disproportionately drawn from more central families and that, controlling for candidate fixed effects, candidates receive a higher vote share in villages where their families are more central. We present evidence that centrality confers organizational and logistical advantages that facilitate clientelistic transactions such as vote buying and do not operate through popularity, name recognition or through the choice of policies more aligned with their constituents' preferences.

*We thank Daron Acemoglu, Michael Davidson, Fred Finan, James Fowler, Horacio Larreguy, Clare Leaver, Simon Quinn, Nico Ravanilla, Byunghwan Son, and conference and seminar participants at APSA 2014, Berkeley "Brokering Votes" Conference, UBC (Comparative-Canadian Workshop), UCSD (Human Nature Group) and Oxford University for feedback. All remaining errors are ours.

†University of British Columbia: cesi.cruz@ubc.ca
‡OxfordUniversity: julien.labonne@bsg.ox.ac.uk
§New York University: pablo.querubin@nyu.edu

1 Introduction

Electoral incentives shape how politicians campaign and allocate resources when in office. For instance, public goods will tend to be under-provided as politicians prefer targeting specific voters with private goods (Lizzeri and Persico, 2001). In contexts where politicians are unable to commit to ex-post policies, electoral strategies may take the form of vote buying or patronage, often leading to inefficiencies and weakening electoral accountability (Acemoglu and Robinson, 2001; Keefer and Vlaicu, 2008; Robinson and Verdier, 2013; Stokes, 2005).

However, politicians do not operate in a vacuum–they are are embedded in social relations that shape their electoral strategies. In particular, elites can use extra-political tools to increase their political power. For example, Anderson et al. (forthcoming) and Baland and Robinson (2008) show how rural elites in Chile and India take advantage of their social status and control over land to win elections. Similarly, Dunning and Harrison (2010) convincingly argue that cross-cutting ties in Mali reduce the appeal of ethnic voting. Munshi and Rosenzweig (2013) show that caste networks may influence the type of politicians elected as well as the policies they choose once in office.

In this paper, we study how clientelistic relationships interact with existing social institutions, focusing on the family (Alesina and Giuliano, 2013). We argue that family ties can interact with the political process in ways that reduce political competition. First, politicians with ties to multiple families can exploit norms of loyalty to those connected to them in order to gain electoral support (Fafchamps and Labonne, 2014). Second, high levels of trust and reciprocity within extended families can reduce the relevance of agency concerns associated with monitoring of individual voter and broker behavior in clientelistic transactions such as vote buying (Finan and Schechter, 2012). Third, the hierarchical nature of families in many societies allows family leaders to deliver *en bloc* the votes of their members to a given set of candidates.

To estimate the effect of family networks on electoral outcomes we use a unique dataset on all 20 million individuals in 709 municipalities 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 graph the full family network in all 709 municipalities, allowing us to compute network centrality measures for all families in our sample. We

can compare network centrality of political and non-political families, and among the set of candidates, assess the relationship between family centrality and electoral prospects. To the best of our knowledge, this study is the first to analyze the effects of politician family networks on electoral outcomes.

First, we find that candidates for public office are disproportionately drawn from more central families. The average political family is in the 90th percentile of the distribution of centrality measures in their municipality. Second, we show that betweenness centrality is positively correlated with a higher aggregate vote share during the elections, a result that holds after controlling for various family-level characteristics.

Second, we introduce candidate fixed effects and exploit within-candidate variation in the distribution of their votes across different villages and find that candidates receive more votes in villages where their families are more central. Again, these results are robust to controlling for various family characteristics at the village-level, such as education levels or occupations. This allows us to rule out that network centrality is simply capturing the effect of other individual or family attributes correlated with electoral success. We also show that our results hold when we restrict the analysis to candidates from families that have never run for mayor in the past. This allows us to rule out reverse causality (i.e. politicians being able to arrange strategic marriages for their family members to become more central in family networks). We also confirm our main findings using an out-of-sample test.[1]

Third, we test alternative mechanisms that explain our benchmark results. We rule out a number of potential explanations for our findings. First, we control non-parametrically for the number of each candidate's relatives in each village. This rules out a simple story where network centrality is proxying for large families (in a non-linear way) and that our results simply capture the fact that individuals vote for their relatives.[2] Second, we show that candidates also receive more votes in villages where their party mates are more central. Given that there is no straight-ticket voting and voters have to select each candidate individually, this rules out that family centrality operates simply through higher name recognition. Similarly, we find that network centrality has no effect for candidates who are running unopposed. This again allows us to rule out name recognition as the sole driver of electoral success for highly central candidates and is consistent with the clientelistic role of family networks, which need to be activated for political purposes and do not operate mechanically. Third, using data from an

[1] We use network measures based on data collected separately from provinces that were not included in the original dataset and for the 2013 rather than the 2010 elections.

[2] We also show that our results are robust to controlling for *degree* centrality–the number of different families connected to the candidate's family by marriage.

in-depth survey conducted in two provinces following the 2013 elections, we do not find any evidence that voters' support for or alignment with a candidate's proposed policies are higher in villages where the candidate's family is more central. This rules out a story where centrality gives candidates an informational advantage that allows them to identify policies that better reflect voters' needs and priorities. Similarly, candidates are not perceived as being "better" (i.e. more honest, approachable or experienced) in villages in which they are more central. However, voters from villages in which the candidate is more central do perceive the candidate as being more politically connected. This is consistent with the idea that candidates use their central position in social networks for political purposes.

Finally, we find that network centrality allows candidates to mobilize voters through vote buying. Specifically, a one standard deviation increase in challenger's centrality leads to a 0.24 standard deviation increase in the incidence of vote buying at the village-level. This suggests that candidates from central families have an advantage in organizing clientelistic practices such as vote buying. We hypothesize that this is due to the ability of central families to exploit norms of loyalty in extended family networks, overcoming organizational and agency problems that hamper clientelistic transactions. Family members can also play the role of political brokers and assist with the conduct of vote buying by helping to distribute the money to voters in dispersed villages and monitoring voter compliance.

More generally, our results highlight the importance of social networks for the performance of political institutions. In many contemporary democracies, elections coexist alongside other traditional social institutions and organizations such as families, clans and religious groups, amongst others. Influential individuals within these social organizations can often take advantage of their position within these networks to take control over elections through clientelistic practices. This may undermine political accountability and the inclusive principle of democratic political institutions.

Our paper contributes to the literature on the effects of social networks in developing countries (Jackson, 2014; Munshi, 2014).[3] Networks have been shown to matter in various contexts. For example, Banerjee et al. (2013) study the diffusion of microfinance in India and find that in villages where first-informed individuals are more central, there is more diffusion. Similarly, Conley and Udry (2010) provide evidence that farmers in rural Ghana learn from their peers about a new technology. Finally, Alatas et al. (2014) analyze how network structure affects the aggregation of information in 600 Indonesian villages. More recently, and closely related to our paper, Naidu et al. (2015) study elite

[3]See Chuang and Schechter (2015) for a comprehensive review.

networks in Haiti and find that more central families are more likely to support a coup.

Our paper is also related to the other strands of the literature. First, it is connected to the literature documenting the value of political connections (Fisman, 2001; Khwaja and Mian, 2005; Faccio, 2006). Second, it complements the literature on the role of families on the functioning of democratic institutions and businesses (Alesina and Giuliano, 2013; Bertrand and Schoar, 2006).

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

2 Family Networks and Elections

There are multiple channels through which politicians may benefit electorally from a central position in municipal or village family networks. In this Section we outline some of these alternative channels that we explore empirically in Section 4.

One potential mechanism through which family centrality can affect electoral performance is name recognition. Central families tend to be better-known in their respective villages and municipalities. Voters may have heard of the family before and may even know members of the family personally, leading to a preference for candidates from known families than candidates from less popular families. There is indeed evidence that voters are more likely to vote for candidates from well-known families and that family names function as a "brand" that voters can identify (see, e.g., Kam and Zechmeister, 2013).

In addition, a central candidate's network of relatives, and relatives' relatives, will likely derive rents from having a family member in office and as such their incentives are closely aligned with the candidate's electoral objectives. For example, Fafchamps and Labonne (2014) show that being related to an elected local official increases the likelihood of having a better-paying job. Thus, a higher number of close or extended relatives may give central candidates a higher number of "core supporters."

Candidates from more central families may also enjoy an informational advantage. They can use their privileged position in the network to learn about the communities' preferences and tailor their campaign platforms and promises. This informational advantage may also enable them to learn about places where their support is relatively weak or strong and adjust their campaign strategies accordingly. Relatedly, strong family networks may also allow politicians to solve commitment problems regarding

the policies they will put in place once in office (Munshi and Rosenzweig, 2013).

Finally, centrality in family networks may be particularly important in many new democracies without meaningful differences between parties and programs, where politics is dominated by clientelistic political exchange (Keefer and Vlaicu, 2008). Forms of political exchange include vote buying, a one-time exchange of money for the vote (see, e.g., Schaffer and Schedler, 2007), or patronage in the context of a patron-client relationship (see, e.g., Stokes, 2005; Fafchamps and Labonne, 2014). A recent strand of literature has emphasized the *informational* problems that candidates face in order to be competitive in elections (Stokes et al., 2013). According to Kitschelt and Wilkinson (2007), the informational costs in programmatic political contexts are primarily internal organizational costs (i.e. to ensure that political actors within a party have a coherent platform to present to voters). By contrast, clientelist contexts are much more demanding, requiring sufficient information for the identification of clients, delivery of benefits, and monitoring of political exchange (Calvo and Murillo, 2009; Kitschelt and Wilkinson, 2007).

Politicians in clientelist systems need detailed information about voter attributes and preferences so they can target their electoral strategies accordingly (Finan and Schechter, 2012; Nichter, 2008; Cruz, 2013). They also need to hire campaign workers–also known as brokers–who will be responsible for implementing those strategies in the ground and convincing voters (Lehoucq, 2007). Last, they need to ensure that all actors at each stage–from voters to brokers at various levels–are complying with the agreements associated with clientelistic political exchange (Szwarcberg, 2011; Stokes et al., 2013; Brusco et al., 2004). Candidates and their political brokers must make sure that the voters they target for vote buying comply and don't vote for a different candidate (or decide to stay at home) after payment has been made. At the same time, given that politicians don't have perfect information about their brokers' types and actions, they also need to exert effort monitoring them (see, e.g., Kitschelt and Wilkinson, 2007; Stokes et al., 2013; Larreguy, 2013; Larreguy et al., 2014). These compliance and monitoring problems are especially acute because the types of transactions associated with clientelism tend to be if not outright illegal (as in vote buying), certainly perceived as undesirable (as in patronage). There are no legal or institutional means of enforcing clientelistic agreements and politicians need help from people that they can trust to deliver (Kitschelt and Wilkinson, 2007; Stokes et al., 2013; Larreguy et al., 2013).

Family networks may be particularly effective in overcoming many of these informational and

agency problems in the clientelistic chain. Social norms of loyalty and reciprocity among extended relatives can substitute for more complex mechanisms of monitoring that are often necessary in other contexts. [4]

More central candidates may be less concerned about voters cheating and not delivering their vote for them after payments have been made. On the one hand, candidates who are well-embedded into local social networks are in a better position to identify voters who are more likely to be reciprocal. More generally, central candidates may be less concerned about voters cheating them as the social pressure and ties of loyalty between voters and members of the politician's extended family are often sufficient to ensure compliance.

Candidates from central families can also avoid or minimize agency problems between them and their brokers. For example, connected candidates can use their relatives as political brokers. The repeated, sustained interactions between family members increase incentives for compliance. In addition, because family ties extend beyond politics, candidates have additional means with which to reward compliance and punish shirking. Candidates can condition access to family resources on assistance with campaign activities. As a result, shirking during the campaign period could have broader repercussions for brokers with family ties to the politician. Last, even if the brokers in the village are not members of the politician's family, a central position in the network will improve the candidate's ability to broker agreements between and among the families in charge of delivering votes.

In sum, we hypothesize that a central position within family networks will lead to both an increase in the propensity to run as well as an improved chance of winning more votes. Moreover, we expect that family networks will confer significant logistical and operational advantages for candidates in conducting contingent political exchange. In particular, for direct voter interventions such as vote buying, centrality facilitates identification and access to voters with reduced monitoring requirements, as well as a providing a larger pool of potential brokers to implement the vote buying transaction. Consistent with this, we expect that family centrality will improve the candidate's ability to mobilize voters using vote buying.

[4]For example, politicians must often rely on disaggregated voter reports, or forms of technology that undermine the secrecy of the ballot–such as cellphone cameras–in order to monitor voter and broker behavior and ensure that clientelistic transactions are fulfilled.

3 Context and Data

3.1 The Setting

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.

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.[5]

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

[5]This was revealed by a public official in an interview with one of the authors in August of 2014.

play an important role in other countries such as India, Ireland and Japan amongst others.[6]

3.2 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.[7] This leaves us with information on 20 million individuals in about 15,000 villages in 709 municipalities.[8] 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.[9]

Second, we use results from the 2010 municipal elections that were collected from the Commission of Elections (COMELEC) official election results website. For each candidate we have data on their party affiliation and number of votes received in each precinct. 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) and to get data on the number of registered voters. Finally, we use official listings of candidates to get each candidate's full name, including last and middle names. Restricting the data to municipalities for which we have NHTS-PR data leaves us with data on about 1,920 candidates for the mayoral elections and 18,400 electoral precincts.

We use two additional data sources to test for mechanisms that could explain our main effects. First, we use data from a survey that was 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.[10] We secured access to precinct-level lists of voters prepared by COMELEC that include the full names of all voters registered in each precinct. Combining data from all precincts in each village, we can get data on all registered voters in all villages in the 12 municipalities. This will allow us to replicate our benchmark results using a different sample. Most importantly, the survey

[6] See for example Bohlken and Chandra (2014) and Smith (2012).

[7] In the remaining municipalities, only households in so-called *pockets of poverty* were interviewed.

[8] Villages in the Philippines, also known as *barangays* in the local language, are the lowest administrative unit. Municipalities are split into multiple barangays, that are administered by an elected barangay captain and a barangay council (*barangay kagawad*). Presently, there are roughly 42,000 barangays across the Philippines

[9] Fernandez (2012) describes the data in more detail.

[10] More information on the survey is available in Cruz et al. (2014).

collected detailed information on candidate's proposals regarding the allocation of the municipality's Local Development Fund (LDF) across 10 different sectors[11], voter's preferences over the allocation of the LDF and their subjective rating of candidate's proposals. Voters also had to report whether they associated candidates with different traits such as honesty, approachability, experience and political connectedness. Finally, voters were also asked to report whether vote-buying by any candidate took place in their village. This allows us to assess the extent to which more central candidates are better able to chose policies reflecting their constituents' preferences or to organize vote buying.

3.3 Explanatory Variables

Our primary explanatory variables are social network variables intended to measure the centrality of politician families within the larger family network in their locality, in this case either the village (over 15,000 networks) or the municipality (709 networks).[12] We use centrality because we are interested not only in a family's number of connections but also in how a family's *position* in social networks affects electoral outcomes. We construct networks (and network measures) both at the municipality and at the village level.

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, which enables us to establish that a marriage took place between members of those two families; (iii) names are difficult to change.[13]

More concretely, family names in the Philippines have the following structure:

firstname midname lastname

where *firstname* corresponds to the individual's first name, *midname* corresponds to the mother's maiden name (for men and single women) or the father's family name (for married women) and

[11] Every year, each municipality receive 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). The 10 sectors on which we have data are: public health services, public education services, cash or in-kind transfers (such as loans or job assistance), water and sanitation services, road construction and rehabilitation, construction of community facilities (such as multipurpose halls or basketball courts), business loans and other private economic development programs, agricultural assistance and irrigation systems and peace and security and community events and festivals.

[12] Importantly, we do not use a sampled network to generate our centrality measures and as such they do not suffer from the problems identified by Chandrasekhar and Lewis (2011).

[13] 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.*

lastname corresponds to the father's family name (for men and single women) or the husband's family name (for married women).

The naming structure and distribution of family names in the Philippines can be traced back to the 19th century. In 1849, concerned with the arbitrary way in which Filipinos chose their surnames (and the implications for tax collection), Governor Narciso Claveria y Zaldua created a catalog with a list of 61,000 different surnames. Municipal officials throughout the country then allocated one name to each family. 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 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.

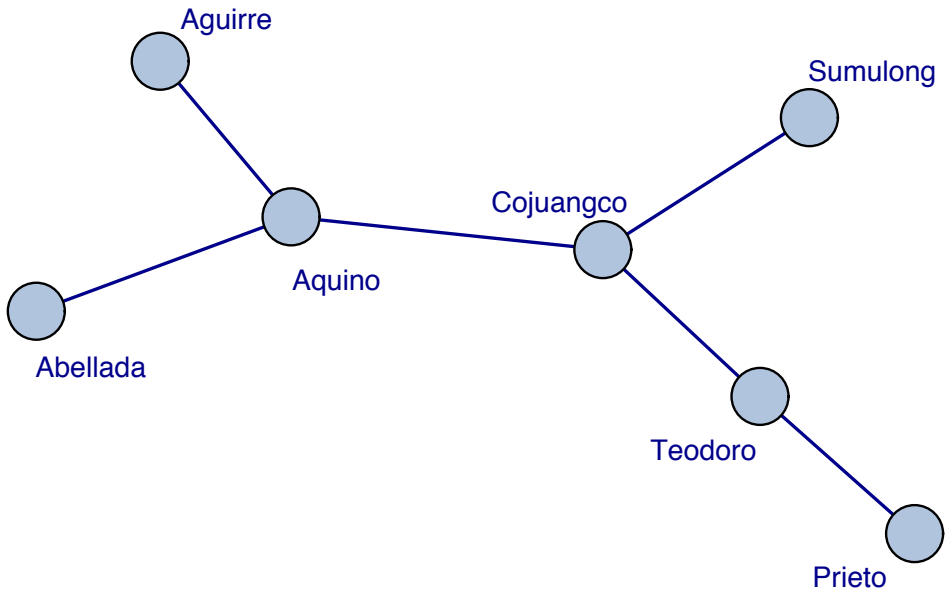


Figure 1: Family network for selected members of President Aquino's family.

Once the networks are constructed within the localities, we compute two different centrality measures for all families in the locality.

Eigenvector centrality is a measure of centrality that 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 by calculating the prestige of a family weighted by whether the others connected to the family are themselves influential (see equation 5 in the Appendix).[14] Families that would be considered central using this measure are those families that have many ties to other well-positioned families. This is one of the more intuitive measures of centrality and is often used to assess prestige and popularity.

Betweenness centrality is the extent to which the family serves as a link between different groups of families. It assesses centrality by looking at whether the family is an important hub in the paths traversing the network and is calculated using the number of shortest paths in the network that necessarily pass through the family (Freeman, 1977). Betweenness for any single family is calculated in terms of its position compared to all other pairs of families (equation 6 in the appendix).[15] Betweenness

[14]Since our ties represent intermarriages, they are undirected–that is, observing a tie from family A to family B implies that the same tie exists from family B to family A. As a result, we do not need to consider in-degree (inward) and out-degree (outward) ties.

[15]For example, given three families, i, j, and k, the betweenness of i would be the ratio of the number of shortest paths

centrality has implications for the ability of the family to broker relations between different groups (Padgett and Ansell, 1993). This measure is often used for thinking about diffusion through the network and the spread of information, contagion, or other effects.[16]

The two measures tend to be correlated, but there are also conceptual differences. High eigenvector centrality families are those that are likely at the center of their main group. By contrast, families with high betweenness centrality are positioned at strategic locations between groups of families.

In our benchmark specifications, for each centrality measure, we assign to each candidate the maximum centrality value associated with either their last name or middle name. The resulting measures are then normalized to be of mean zero and standard deviation one. However, all of our results remain essentially unchanged if we simply use the average of the centrality values for the last and middle names, or if we only use the centrality measure associated with either the last name or the middle name.

4 Empirical Analysis

In this section we present our empirical results. We begin with some descriptive municipal-level analysis of how politicians (winning and losing candidates to the mayoral election) differ from non-politicians in terms of their position within the municipal family networks. We also explore the extent to which a candidate's centrality in the municipality predicts his or her electoral success. Next, we move to the village level analysis where we study the extent to which a candidate's centrality in the village networks predicts the candidate's distribution of votes across the different villages in the municipality. This allows us to introduce candidate-fixed effects and thus rule out the possibility that other individual characteristics of candidates from central families confound our results. We also show that our results do not reflect reverse causality–namely, that politicians' families can strategically marry other families in order to become more central. Finally, we present some evidence on mechanisms that allows us to rule out alternative interpretations of our results and provides support for the clientelistic role of family networks. In all of our regression tables, both dependent and independent variables are normalized. Also, "number of family members" refers to the sum of family members traced by both

between j and k that pass through i to the total number of shortest paths between j and k. A ratio of 1 would imply that family i lies on all paths connecting k and j, while a ratio of 0 would imply that family i is not important to connecting k and j. The betweenness centrality of i could then be calculated for the entire network by averaging this ratio across all of the other node pairs.

[16] For example, if we were transmitting information from one family to another family in a different part of the network using word-of-mouth, then we would be very likely to have to pass through the family with high betweenness centrality.

last and middle name.

4.1 Municipal-Level Analysis: Family Networks, Selection into Politics and Electoral Success

In Table 1 we present some descriptive statistics on the different network measures across our sample of politicians (mayoral candidates) and non-politicians. For each of our two centrality measures, the top panel shows the average value while the bottom panel reports the average percentile in the municipal distributions. The results suggest that politicians' families are substantially more central within the municipal family networks than families of non-politicians. On average, all centrality measures are more than an order of magnitude larger in the sample of politicians than in the sample of non-politicians. The differences between the two samples are also statistically significant. Moreover, the bottom panel shows that politicians are on average in the 90th percentile of the distribution of centrality measures, providing further evidence that they tend to come from the most central and well-connected families in the municipality.

In Table 2 we explore more systematically the role of network measures in predicting selection into politics. To do so we estimate linear probability models of the form:

$$Y_{im} = \alpha C_{im} + \beta X_{im} + \rho_m + \epsilon_{im} \tag{1}$$

where Y_{im} is a dummy equal to one if at least one member of family i in municipality m ran in the 2010 mayoral election. C_{im} corresponds to one of the different centrality measures for family i in the municipality and thus α is the parameter of interest, X_{im} is a set of family*municipality-specific characteristics, ρ_m is a full set of municipality fixed effects that we include in some specifications and ϵ_{im} is the usual idiosyncratic error term. Standard errors account for potential correlations within municipalities.[17]

The results confirm the patterns revealed by the descriptive statistics in Table 1. Both centrality measures are positively correlated with the probability of a member of that family running for office. The point estimates are larger for the betweenness centrality measure for which a one-standard deviation increase is associated with an increase in 0.15 standard deviations in the probability of running for office. One natural concern with these regressions is that estimates may confound the effect of network centrality with other characteristics of the family also correlated with the decision to run for office. For

[17]The sample covers 709 municipalities and so we are not concerned about having too few clusters.

example, more central families may also be larger, or have a higher socio-economic status, which can also be correlated with the decision to enter politics. In order to address this concern, we control for additional family characteristics in Columns (2)-(3). In Column (2) we control for the total number of individuals who belong to the family (carry either of the family names), number of female members of the family and for the number of villages in the municipality where at least one family member lives. Point estimates for betweenness and eigenvector centrality remain relatively unchanged. In Column (3) we control for socio-economic characteristics of the family captured by educational attainment and occupation. In particular, we control for the number of family members in each of the 17 educational categories[18] and 11 occupational categories included in the NHTS-PR.[19] Again, the point estimates remain relatively unchanged, which suggests that our centrality measures do not simply capture other family characteristics associated with economic status. Finally, in Column 4 we include municipality fixed effects. Our point estimates remain stable.

The evidence presented in Tables 1 and 2 should be interpreted cautiously and is used primarily for descriptive purposes. At the same time, these patterns are consistent with the widely held belief that politicians come from highly connected families that play a central role within their respective networks. This effect captures the family's *position* in the municipal network and is not driven simply by how large the family is. The betweenness and eigenvector centrality measures suggest that a family's status and capacity to mediate and potentially broker deals with other prestigious families in the municipality play an important role in the decision to enter politics. To our knowledge, ours is the first paper to provide quantitative evidence on the central network position of those who seek public office. In this sense, we contribute to the nascent literature on the underlying attributes and characteristics of leaders.[20] The patterns we document suggest that a strategic position within social networks may be an important attribute of those who seek elected positions.

Next, we focus on the sample of families that run for office and explore the extent to which family centrality predicts electoral success in the municipality. We estimate regressions of the form:

$$VS_{im} = \alpha C_{im} + \beta X_{im} + \rho_m + \epsilon_{im} \tag{2}$$

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

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

[20]See Ahlquist and Levi (2011) for a review of this literature.

where VS_{im} is candidate i's vote share in the 2010 mayoral elections in municipality m. As above, C_{im} is a measure of how central family i is in municipality m, X_{im} is a set of municipality*family-specific characteristics and ρ_m is a full set of municipality fixed effects ϵ_{im} is the usual idiosyncratic error term. α is the main parameter of interest. Standard errors account for potential correlations within municipalities. In order to isolate the effect of candidate's centrality on turnout, VS_{im} is computed as a fraction of registered voters rather than as a fraction of those who actually voted. However, results are similar when using vote shares as fraction of actual voters (results available upon request).

Results suggest that a candidate's family centrality is positively correlated with electoral success. However, only the estimates for the betweenness centrality measure are large and statistically significant. The point estimates for betweenness centrality remain relatively stable when we control for number of family members, and education and occupational categories in Columns 2 and 3, respectively. The inclusion of municipal fixed effects in Column 4 does not affect the point estimate, but the standard error increases noticeably and the coefficient is no longer statistically significant.

In order to illustrate our results, we show a graphical representation of an actual municipal family network in Figure 2. The blue node denotes the family of the winning mayoral candidate and the red node denotes the family of the losing candidate. Light gray nodes indicate families without a candidate for mayor. The families of the candidates clearly occupy a central position within the municipal family network, and the family of the winning candidate is noticeably more central than the family of the losing candidate.

Overall, the results in Tables 1-3 are suggestive of the importance of network centrality for political success (both for running for public office in the first place and for being successful once in the race). We are confident that these results do not capture the effects of other observable family-level confounders since we control for a wide range of family characteristics. However, the municipal-level analysis in Table 3 should be interpreted with caution as we cannot control for *individual* candidate characteristics that may be correlated with family centrality and electoral success.[21] This may be a concern particularly given the fact that in Table 3 we are conditioning on the sample of families who ran for office. Thus, candidates from families that are not very central may have nonetheless decided to run against very central candidates due to other individual characteristics that we cannot observe. To address this concern in the remaining analysis we include candidate fixed effects in all of our regressions and focus

[21]Our NHTS-PR dataset does not include first names and thus we cannot measure socio-economic characteristics of individual candidates. Furthermore, candidates provide very limited data when they register their candidacy and those data are, to the best of our knowledge, not recorded consistently by the COMELEC.

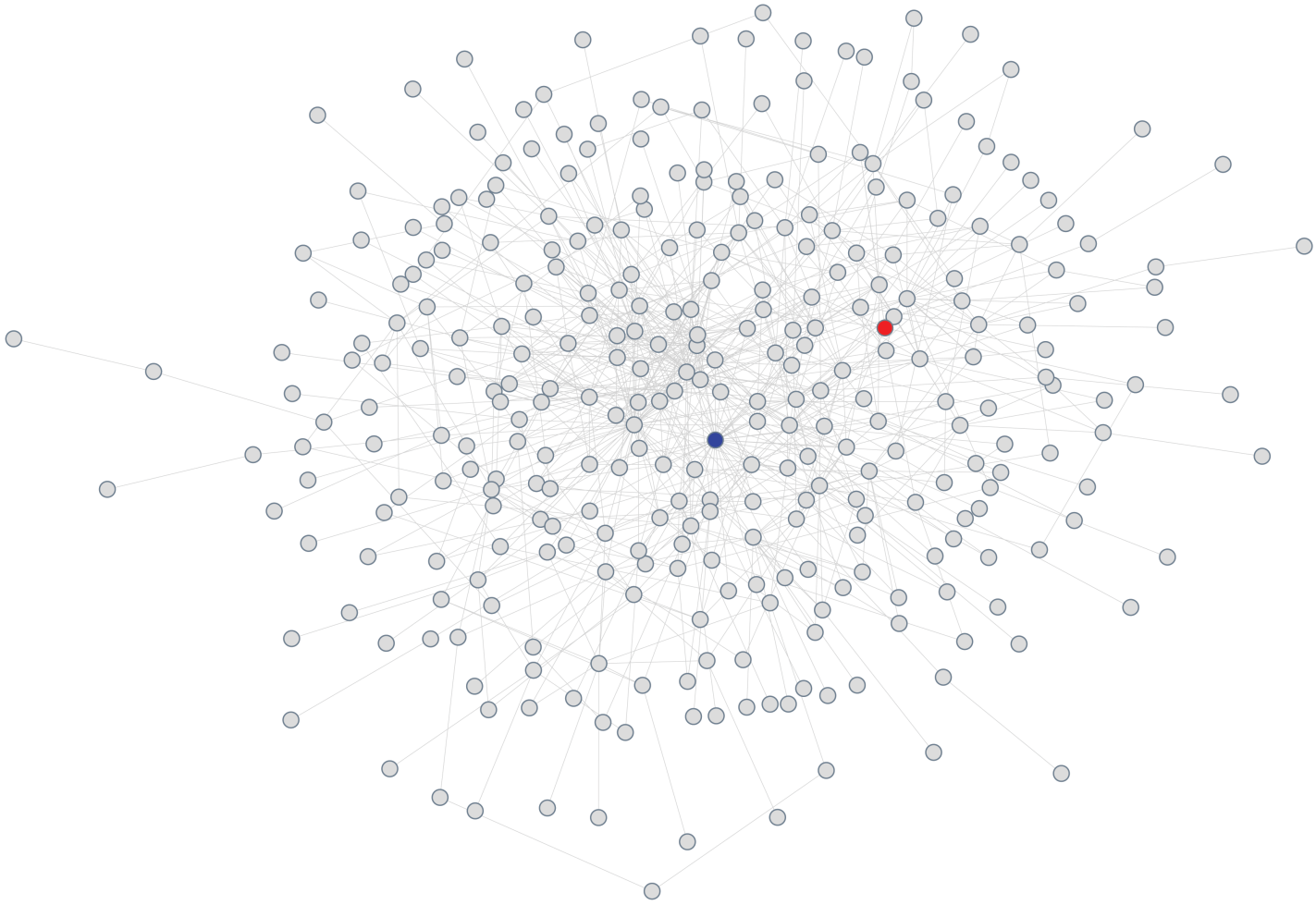


Figure 2: Family network for a municipality. The blue and red nodes are families of the winning and losing mayoral candidates, respectively.

on variation in a candidate's family centrality across the different village networks in the municipality.

4.2 Village-Level Analysis: Family Networks and Spatial Distribution of Votes

In the previous discussion we focused on the different candidates' position in the broader network of families in each municipality and the correlation with overall electoral success. However, a family's presence and influence is not necessarily evenly spread across the municipality. Candidates may be more connected or influential within the smaller networks of some villages than others. Moreover, the village is the relevant unit of analysis whenever studying clientelistic strategies such as vote buying in the Philippines. Elected officials at the village level, such as village captains (*punong barangay*) and village councilors (*barangay kagawad*), often serve as *brokers* who mobilize voters in their villages for candidates in mayoral, provincial or congressional electoral races. We have access to electoral data at

the *precinct* level (recall that there is at least one precinct per village), and thus we can observe voting patterns at very disaggregated levels.

To test the role of family centrality in explaining the spatial distribution of a candidate's electoral support, we estimate regressions of the form:

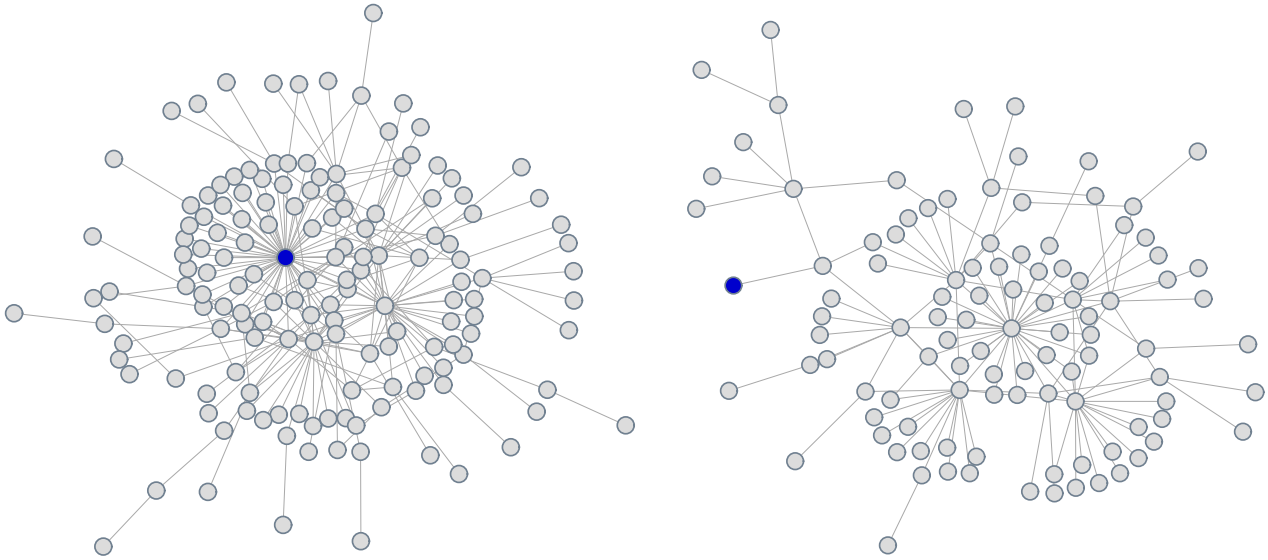
$$VS_{ipv} = \alpha C_{iv} + \beta X_{iv} + \delta_v + \eta_i + \epsilon_{ipv} \tag{3}$$

where VS_{ipv} is candidate i's vote share in the 2010 mayoral elections in precinct p in village v. As above, we measure vote share as a percentage of the registered population and not as a percentage of the actual voting population. C_{iv} is a measure of how central family i is in village v, X_{iv} is a set of village*family-specific characteristics and δ_v is a set of village fixed effects. The term η_i corresponds to candidate fixed effects included in all specifications. Finally, ϵ_{ipv} is the usual idiosyncratic error term and standard errors account for potential correlation within municipalities.

The estimates based on equation (3) are reported in Table 4. Estimates in Column 1 suggest that candidates receive more votes in villages where their families are more central. The coefficients for the two centrality measures are positive and statistically significant and suggest that a one standard deviation increase in family centrality leads to a 0.06 standard deviation increase in the candidate's vote share in the precinct. All of our estimates remain relatively unchanged when we control for number of total and female family members (Column 2), when we control for number of family members in the different educational and occupational categories (Column 3) and when we include village fixed effects (Column 4). This specification improves upon the municipal-level regressions since we can now account for all individual candidate characteristics (that are absorbed by the candidate fixed effects) that may have confounded the effect of family centrality in the regressions in Table 3.

These results are illustrated in Figure 3, which compares the village-level centrality across two villages for the same winning candidate's family (the blue dot) depicted in Figure 2. The first picture is an example of a village in which the winning candidate received approximately 60 percent of the vote, while the second picture is a village in which the candidate received only 20 percent of the vote. The winning candidate is noticeably more central in the first village than in the second. A comparison of Figures 2 and 3 allows us to illustrate the advantage of including candidate fixed effects in our specification. While estimates from the municipal-level regressions illustrated in Figure 2 (and reported in Table 3) rely on variation across *different* candidates within the same municipality (i.e. red

dot vs. blue dot), the estimates from regressions with candidate fixed effects illustrated in Figure 3 rely on variation in the centrality and electoral success of the *same* candidate (blue dot) across different villages.



Candidate received 60% of the vote

Candidate received 20% of the vote

Figure 3: Family networks in two villages in the same municipality. The blue dot represents the winning candidate's family.

For ease of interpretation in Tables 2-4 we run separate regressions for each centrality measure. However, in Table A.1 in the Appendix we report the results when we regress a candidate's vote share on both centrality measures simultaneously. All specifications are analogous to the ones reported in Table 4. The estimates for both betweenness and eigenvector centrality remain positive and statistically significant (point estimates are slightly smaller which is expected given the positive correlation between both network measures).

In Table 5 we assess the robustness of the benchmark estimates reported in Table 4. In Columns 1-3 we eliminate outliers, more concretely, all observations with values in the top and bottom 1, 5 and 10 percentiles (respectively), of the distribution of the different centrality measures. The point estimates for betweenness and eigenvector centrality remain stable. In Column 4 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 family networks there. However, even when we exclude ARMM municipalities from our sample, the point estimates remain similar. In Column 5 we show that our results are robust to weighting each observation by the number of registered voters in the precinct.

In Column 6 we control non-parametrically for all family attributes. In particular, we include over 1,500 dummy variables; one for each possible value of (i) total number of family members (a total of 221 dummies); (ii) number of female family members (138 dummies); (iii) number of family members in each of the educational categories (511 dummies); and (iv) number of family members in each of the occupational categories (695 dummies). This is a very demanding specification in that we don't assume any functional form for the family characteristics. The point estimates for betweenness and eigenvector remain relatively unchanged and statistically significant. Relatedly, to deal with concerns that our results are simply capturing the number of ties to other families, rather than the family's position in the network, in Column 7 we control for the candidate's family degree centrality. Again, the point estimates are stable and remain statistically significant.[22] Finally in column 8 we rule out that our results simply reflect a "hometown" effect (candidates may be particularly central and electorally successful in their hometown village). We do not know the exact village of residence or birth of each candidate. Thus, for each candidate, we identify the village where the candidate has the most relatives. We then create a dummy for precincts located in that village. We then reproduce results reported in Column 4 of Table 4 but control for the hometown dummy. The estimates remain essentially unchanged.

We also perform a placebo exercise where we randomly reallocate the betweenness and eigenvector centrality measures across candidates and villages within the same municipality. We perform 10,000 iterations of this exercise. In Figure A.1 we plot the distribution (density) of point estimates of α from estimating regression 3 in the "placebo" dataset generated in each iteration. The top panel shows the distribution of estimates for betweenness centrality and the bottom panel for eigenvector centrality. The distribution of the placebo estimates is centered around zero. Moreover, none of the placebo estimates from the 10,000 iterations comes even close to the actual estimate reported in Column 3 of Table 4 and depicted with a red vertical line in each plot.

A final potential concern with our estimates is reverse causality. Rather than capturing the extent

[22]The likelihood that two individuals sharing a middle or last name are not related might vary according to municipal population. To check that our results are not affected by this, we estimate our regressions on samples excluding outlying municipalities in the population distribution. Results, available in Table A.6, are robust to excluding municipalities in the bottom and top 1, 5, 10 and 25 percent of the population distribution. Again, the point estimates for betweenness and eigenvector remain relatively unchanged and statistically significant.

to which more central families enjoy an electoral advantage, our estimates may simply reflect the fact that a politician can use political power to marry her family members strategically to other families and become more central as a result. In order to address this possibility, in Table 6 we report estimates of the specifications reported in Table 4 but where we allow a different coefficient of the centrality measures for old and new political families. New political families are those that never had a family member running for mayor in their municipality prior to 2010 (more concretely during the 1988-2007 elections). The point estimates show that the effect of family centrality is, if anything, *stronger* for new families. This gives us further confidence that our results do not capture reverse causality but rather the effect of centrality on electoral success. It also suggests that social networks may be particularly important for challengers or new candidates that do not have access to public resources or to the exposure that incumbency provides. These new families may rely strictly on their capacity to mobilize voters through clientelistic transactions. We return to this issue below.

For completeness, in Tables A.2-A.4 in the Appendix we show the municipal-level results reported in Tables 1-3 but differentiating between old and new families. The patterns are very similar. Even when focusing on new families we observe that their centrality measures are in the 85th percentile of the distribution. The relationship between the different centrality measures and the decision to run for office is positive and statistically significant in all specifications for both new and old families. The relation between eigenvector and betweenness centrality in the municipal networks and vote share is positive and statistically significant *only* for new families suggesting that our baseline results are not driven by reverse causality and that if anything, social networks are more important for new families without any previous control of public office.

4.3 Why Does Network Centrality Matter? Alternative Channels and Mechanisms

In Tables 1-6 we have documented a positive and robust relationship between an individual's family network centrality and the decision to run for office and, among those who run, electoral success in the municipality as a whole and across villages. We are confident that these results do not confound the effect of other family characteristics plausibly correlated with network centrality (we control flexibly for family size and socio-economic characteristics) or of other individual characteristics of candidates from central families (we include candidate fixed effects in our regressions). We have also ruled out the possibility that our relationship simply captures the effect of previous political power on current network centrality as the effect of network centrality is *stronger* for candidates from new families

without any prior political experience.

In this Section we explore the empirical relevance of some of the mechanisms we introduced in Section 2, through which centrality may affect a candidate's political success.

4.3.1 Substitutes for Other Candidate Attributes

One possibility is that family centrality only matters for relatively weak candidates who must rely more on their family networks due to lack of exposure, political experience or access to resources. For example, one could hypothesize that incumbents rely less on their family network connections than challengers since they have experience with policy making, can run on their track record and can have higher exposure, create connections while in power and use public resources to guarantee their electoral success. We test this possibility in Panel A of Table 7 where we report estimates of equation (3) but we allow for a different coefficient on the centrality measures for incumbents and challengers. All regressions include candidate fixed effects. Columns 2 and 4 control for total number of families, number of female relatives and number of family members in the different educational and occupational categories. The results provide some suggestive evidence that family networks are slightly more important for challengers than for incumbents. The interaction of the centrality measures with an incumbent dummy is negative and statistically significant for betweenness centrality. The effect of centrality is nonetheless positive for incumbents, but the effect is roughly 33% smaller than the effect for challengers. This result is consistent with the estimates reported in Table 6 where we found that the effect of network centrality is stronger for candidates from new families without any prior political experience (that by definition will be coded as challengers). In Panel B we perform another exercise to assess the extent to which centrality matters only for very weak candidates who get a handful of votes (but have no real chance of winning the election). In this specification, we drop all candidates who place 3rd or worse in the mayoral race in the municipality and thus limit our analysis only to strong or "serious" competitors (winners and runner-ups). The point estimates are similar to those we found in the benchmark regressions in Table 4. In sum, estimates in Panels A and B of Table 7 suggest that the effect of network centrality is stronger for challengers than for incumbents, though the effect for the latter is not negligible. Moreover, our results are not driven by very weak candidates getting a handful of additional votes due to their network position. Our results remain identical when we focus only on winners and runner-ups.

4.3.2 Name Recognition

As discussed in Section 2, one potential mechanism through which family centrality can affect electoral performance is name recognition. In Panel C of Table 7 we explore this by looking at the effect of centrality on the vote share of candidates running in unopposed races.[23] If the underlying mechanism is simply name recognition, then we should observe a higher vote share for candidates in villages in which they are more central (and people are more familiar with the name) irrespective of the underlying level of competition of the race.[24] However, the estimated coefficients show that centrality plays no role in uncontested races. Thus, it seems that candidates need to deliberately *activate* their social networks for electoral purposes, only when it is necessary (i.e. when the race is competitive). Networks do not seem to operate mechanically through mechanisms such as name recognition. This is consistent with the potential clientelistic role of social networks that we will provide more direct evidence for later in the paper. Importantly, the point estimates on this sub-sample are much smaller than on the full sample which indicates that failure to reject the null is not merely a result of loss of statistical power due to lower sample size.

An alternative way to assess the relevance of the name recognition channel is to explore the effect of the family centrality of a mayoral candidate's party-mates running for the vice-mayorship and municipal council in the same municipality. While political parties in the Philippines do not have a very centralized structure or organization that coordinates policy proposals and exerts discipline and alliances in candidates across the different offices and provinces, within a municipality shared party affiliation does usually describe political alliances or coalitions.[25] However, voters in the Philippines cannot do straight-ticket voting through a single option in the ballot, but must always mark separately their preferred candidate in each office. Thus, if we observe a strong correlation in the electoral success of a mayoral candidate and its party-mates, this has to be explained by active and deliberate campaigning of individual candidates in support of their co-partisans and cannot be explained by mechanical straight ticket voting. In Table 8 we report regressions based on equation (3) but where we estimate the separate effect of a mayoral candidate's own family network centrality and the centrality

[23] Importantly, write-in candidacies are not allowed in the Philippines. Unopposed candidates only need one vote to be elected.

[24] Notice that since our vote share measure is normalized by the fraction of registered voters (and not by the number of those who effectively voted), unopposed candidates do not simply receive a vote share of 100% in our dataset. Vote shares in this case mostly reflect the decision of a candidate's supporters to turn out to vote. In fact, there is substantial variation in the vote share for candidates in unopposed races, with an average of 60.7 and a standard deviation of 18.1.

[25] For example, Sidel (1995, p 156) argues that *"ties of consanguinity and affinity allow a politician to activate a network of relatives, even as his coalition partners (the vice mayor, municipal councilors, barangay captains) do the same on his behalf."*

of her party-mates' families. In each village, we simply keep the centrality value of the vice-mayoral or council candidate from the same party with the highest centrality measure in that village. However, results are qualitatively similar if we simply take the average of the centrality measures amongst all the party mates.

The results reveal that the family centrality of a mayoral candidate's party mates in a given village has a positive effective on the vote share of the candidate in that respective village. This holds across both centrality measures and is robust to controlling for number of family members (total and female – Column 2), by education and occupation of family members (Column 3) and to the inclusion of village fixed effects (Column 4). Naturally, point estimates on own centrality are larger than those for party-mates centrality. At the same time, the positive and significant effect of party-mates centrality is inconsistent with name recognition being the only channel through which network centrality affects electoral success. A mayoral candidate's family may not be very well known in a village in which one of her party mates' family is very central. Thus, the fact that family centrality of party mates is also important suggests that candidates take advantage of their position within family networks in their village to actively seek the support for other party mates who may not necessarily be well connected in that village. This also anticipates the channel that we will emphasize in this paper, namely, the clientelistic role that family networks play in order to mobilize voters.

4.3.3 Family Centrality, Policy Choices and Candidate Traits

In this section we explore whether the electoral advantage enjoyed by candidates in villages where their families are more central is driven by other candidate attributes that make them appear as "better" candidates to voters in these villages. To do this, we use data from the survey described in Section 3 that was implemented in Ilocos Norte and Sur provinces shortly after the May 2013 mayoral elections. As an additional check, we replicate our baseline results reported in Table 4 on this sample of municipalities not included in our main sample and using network measures based on the family names of registered voters. Results are reported in Appendix Table A.5. Reassuringly, the main results are reproduced in this different sample and for a different election year. The point estimates for both network measures are positive and statistically significant. Unfortunately, in all regressions using data from the Ilocos Survey we cannot control for detailed family characteristics (these measures come from the NHTS-PR dataset). Nonetheless, in Column 2 we control for total number of family members (in this case, total number of registered voters who share either the candidate's last name or middle

name) and in Column 3 we add village fixed effects. The point estimates remain relatively unchanged. The fact that our main results hold on a different sample, with network measures based on a different enumeration of village residents, and with electoral results from a different election year (2013) gives us further confidence in our results.

First we test whether more central candidates choose policies more aligned with voter's preferences. We report estimates of equation (3) but use as dependent variables alternative measures of voters's support for or alignment with the candidate's proposals. All regressions reported in Table 9 include candidate and village fixed effects and control for the number of registered voters who share one of the candidate's family names. In Column 1 the dependent variable is the (normalized) average rating of voters in the village of the policies proposed by the candidate.[26] We find no evidence that policies and programs of mayoral candidates from more central families are better rated by voters. Point estimates for both centrality measures are very small and never statistically significant irrespective of the specification. In column 2 we use a normalized measure of congruence between a candidate's proposed policies and voter preferences. For each voter-candidate pair, we compute the fraction of the budget on which the candidate and the voter agree. Then, for each candidate, we average this congruence measure over all voters in the village and normalize by subtracting the mean and dividing by the standard deviation. The estimates for this congruence measure are small and are never statistically significant for either centrality measure. This provides suggestive evidence that the channel through which family centrality affects electoral outcomes is not through informational advantages that enable central candidates to craft more popular policies. Our results in Table 9 suggest that any informational advantages conferred by networks do not translate into policies more aligned with voter's preferences.

Next we estimate equation (3) on a set of dependent variables that measure candidate traits such as honesty (Column 3), approachability (Column 4), experience (Column 5) or political connectedness (column 6) as rated by voters.[27] Estimates for most traits are small and are not statistically significant.

[26]The exact question is as follows: "Candidates often propose policies or programs that they would like to implement after they are elected. We'd like to know how much you agree or disagree with the candidate's proposals and platform. We'll show you a worksheet with a scale of 0 to 4, where 0 is strongly disagree and 4 is strongly agree. Please place the candidates' names where they belong on the scale." We start by taking the average rating given by each individual to all candidates in the municipality. We remove the individual-specific average from the individual rating. We then take the village-level averages for each candidate and normalize the resulting variable to be mean zero and standard deviation of one.

[27]The exact question in the survey is as follows: "Now we are going to show you a set of worksheets one for each candidate as well as some flashcards containing some traits [Approachable/Friendly; Experienced in politics; Honest; and Politically well-connected] that candidates might have. For each of these traits, please place them on the worksheet of the candidate that you most associate with that trait. You may place the same trait on both worksheets or you may choose not to place a

One notable exception is that candidates are perceived as being more politically connected in villages in which their family has higher betweenness centrality. The estimate in column 6 for this outcome is positive and statistically significant. Interestingly, this is the only trait directly related with network position. Moreover, as previously argued, betweenness centrality captures the ability to intermediate and broker between different families in the village connected to each other through the candidate's family. This potential role for brokerage may reflect the stronger political connections reported by voters for more central candidates.

4.3.4 Family Networks as Clientelistic Networks

Finally, we assess whether network centrality allows candidates to engage in illegal clientelistic practices aimed at mobilizing voters, such as vote buying. We also rely on data from the Ilocos survey where respondents were asked whether they had received money in exchange for their vote. Voters do not report the specific candidate that offered money in exchange for their vote, simply whether *any* of the candidates engaged in vote buying.[28] Thus, since we only have one observation per village we cannot include village fixed effects in this specification. We estimate a regression similar to (3) but we use as dependent variable a normalized measure of vote buying incidence and we report a different coefficient for incumbents and challengers.[29] The results reported in Table 10 suggest that centrality of challengers is strongly associated with a higher incidence of vote buying. An increase in one standard deviation in network centrality of the challenger, for any of our two main centrality measures, leads to an increase in vote buying of roughly 0.25 standard deviations. Estimates remain unchanged when controlling for total number of family members (Column 2) or when we control for a broad set of village characteristics (Column 3)[30].

Coefficients on incumbent centrality on the other hand are very small and always statistically insignificant across all specifications. One possible interpretation of our results is that challengers

trait at all if you feel that it does not apply to any of the candidates." We start by taking the average response given by each individual to all candidates in the municipality. We remove the individual- specific average from the individual rating. We then take the village-level averages for each candidate and normalize the resulting variable to be mean zero and standard deviation of one.

[28]The exact question was: "As far as you know, are there cases of vote buying in this barangay during the recent elections?" [0 = No; 1 = Yes; 2 = Prefer not to answer]. If yes: "Did someone offer you money for your vote?" [0 = No; 1 = Yes; 2 = Prefer not to answer]. We generate village-level averages and normalize the resulting variable to be mean zero and standard deviation one.

[29]As the data was collected for an experiment that had an effect on vote buying (Cruz et al., 2014) we only use observations in the control group here.

[30]We control for village density, the relevant village-level network measure, a dummy equal to one if the village is classified as rural, average education, age, household size and length of residence as well as the share of the population that is female, that receives remittances from abroad and benefits from a CCT program.

are more reliant on vote buying than incumbents (Vicente, 2014). This is certainly plausible since incumbents have at their disposal other policy instruments or government resources and programs with which they can co-opt and mobilize voters and are less likely to be perceived as vote buying. On the other hand, challengers have no access to policy or government programs and have to engage in direct vote buying around election time. However, since our measure does not allow us to distinguish whether the vote buying reported by voters was made by challengers or incumbents, another possibility then is that the presence of highly central challengers creates a threat for the incumbent and as a consequence both incumbents and challengers need to engage in vote buying. Finally, while higher challenger centrality may be correlated with the village being more competitive, winning the mayoral elections requires winning the municipality as a whole. Since there is no advantage in winning more votes in any specific village, there is no reason why political competition by itself could explain higher rates of vote buying in villages in which challengers are more central. Thus, we believe this reflects the effect of network position which enable vote buying rather than some other characteristic of the village.

5 Conclusion

Our results show that politician family networks are strong predictors of candidacy and electoral success: candidates for public office are disproportionately drawn from more central families and family networks contribute to higher vote shares during the elections. We establish that family networks confer advantages for pursuing clientelist electoral strategies, by presenting evidence that politicians conduct vote buying in villages where their families are central. Furthermore, we rule out potential alternative explanations for our findings: the results are not driven by reverse causality, by name recognition or by the fact the individuals vote for their relatives.

Our findings have a number of implications. First, since family networks are relatively slow to change, this could explain why political power tends to be concentrated among a few families in a number of consolidating democracies (Querubin, 2010). Second, candidates might be able to circumvent having to closely monitor their brokers by hiring individuals to whom they are closely connected. This reduces the risk that they shirk on the job. Our findings contribute to the broader literature on clientelism and demonstrate the importance of the personal networks of politicians for explaining the puzzle of monitoring identified in the literature (Kitschelt and Wilkinson, 2007; Stokes et

al., 2013; Finan and Schechter, 2012; Cruz, 2013). Furthermore, many of the intuitions for why network centrality can facilitate clientelistic practices (i.e. extended networks of trust and loyalty) also apply to other types of social networks, such as castes and ethnic groups, and not just to family networks.

Although family networks represent the most basic and fundamental type of social relationship, limitations in data collection have made it difficult to empirically isolate the political importance of these networks. We are able to use a unique dataset to demonstrate that while we might have expected family networks to matter for politics, it's not for the reasons that we might have thought: family networks have less to do with status, popularity, or name recognition than with the organizational and logistical advantages that these ties can confer. We present evidence to suggest that politicians are able to leverage family networks to improve the effectiveness of vote buying and other clientelistic political strategies. The clientelistic relationships among brokers and politicians is situated in a richer social context that is often difficult to account for empirically, and while the literature has indicated the importance of personal relationships, ours is the first large-scale evidence of the substantial role that they play.

References

Acemoglu, Daron and James Robinson, "Inefficient Redistribution," *American Political Science Review*, 2001, *95* (3), 649–661.

Ahlquist, John S. and Margaret Levi, "Leadership: What It Means, What It Does, and What We Want to Know About It," *Annual Review of Political Science*, 2011, *14*, 1–24.

Alatas, Vivi, Abhijit Banerjee, Arun Chandrasekhar, Rema Hanna, and Benjamin Olken, "Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia," *mimeo, MIT*, 2014.

Alesina, Alberto and Paola Giuliano, "Family Ties," in Philippine Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, North Holland, 2013.

Anderson, Siwan, Patrick Francois, and Ashok Kotwal, "Clientelism in Indian Villages," *American Economic Review*, forthcoming.

Baland, Jean-Marie and James A. Robinson, "Land and Power: Theory and Evidence from Chile," *American Economic Review*, September 2008, *98* (5), 1737–65.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, "The diffusion of microfinance," *Science*, 2013, *341* (6144).

Bertrand, Marianne and Antoinette Schoar, "The Role of Family in Family Firms," *Journal of Economic Perspectives*, 2006, *20* (2), 73–96.

Bohlken, Anjali and Kanchan Chandra, "Dynastic Politics and Party Organizations: Why Family Ties Improve Electoral Performance in India," *mimeo*, 2014.

Bonacich, Philip, "Factoring and weighting approaches to clique identification," *Journal of Mathematical Sociology*, 1972, *2*, 113–120.

_ , "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 1987, *92* (5), 1170–1182.

Brusco, Valeria, Marcelo Nazareno, and Susan C. Stokes, "Vote Buying in Argentina," *Latin American Research Review*, 2004, *39* (2), 66–88.

Calvo, Ernesto and Maria Victoria Murillo, "Selecting Clients: Partisan Networks and Electoral Benefits of Targeted Distribution," 2009.

Chandrasekhar, Arun and Randall Lewis, "Econometrics of Sampled Networks," *mimeo*, 2011.

Chuang, Yating and Laura Schechter, "Social Networks in Developing Countries," *Annual Review of Resource Economics*, 2015, *7* (1).

Conley, Timothy G. and Christopher Udry, "Learning About a New Technology: Pineapple in Ghana," *American Economic Review*, 2010, *100* (1), 35–69.

Corpuz, Onofre D., *The Philippines*, Englewood Cliffs, New Jersey: Prentice-Hall, 1965.

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

_ **, Philip Keefer, and Julien Labonne**, "Incumbent advantage, voter information, and vote buying," *mimeo*, 2014.

Dunning, Thad and Lauren Harrison, "Cross-cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali," *American Political Science Review*, 2 2010, *104*, 21–39.

Faccio, Mara, "Politically Connected Firms," *American Economic Review*, 2006, *96* (1), 369–386.

Fafchamps, Marcel and Julien Labonne, "Do Politicians' Relatives Get Better Jobs? Evidence from Municipal Elections in the Philippines," *University of Oxford, CSAE Working Paper WPS/2014-37*, 2014.

Fegan, Brian, "Entrepreneurs in Votes and Violence: Three Generations of a Peasant Political Family," in Alfred McCoy, ed., *An Anarchy of Families: State & Family in the Philippines*, Madison, WI: University of Wisconsin Press, 2009, pp. 33–108.

Fernandez, Luisa, "Design and Implementation Features of the National Household Targeting System in the Philippines," *World Bank - Philippines Social Protection Note No 5*, 2012.

Finan, Frederico and Laura Schechter, "Vote Buying and Reciprocity," *Econometrica*, 2012, *80* (2), 863–881.

Fisman, Raymond, "Estimating the Value of Political Connections," *American Economic Review*, 2001, *91* (4), 1095–1102.

Freeman, L.C., "A Set of Measures of Centrality Based on Betweenness," *Sociometry*, 1977, *40*, 35–41.

Hutchcroft, Paul and Joel Rocamora, "Strong Demands and Weak Institutions: The Origins and Evolution of the Democratic Deficit in the Philippines," *Journal of East Asian Studies*, 2003, *3*, 259–292.

Jackson, Matthew O., *Social and Economic Networks* Princeton University Press, Princeton University Press, 2010.

_ , "Networks in the Understanding of Economic Behaviors," *Journal of Economic Perspectives*, 2014, *28* (4), 3–22.

Kam, Cindy D. and Elizabeth J. Zechmeister, "Name Recognition and Candidate Support," *American Journal of Political Science*, 2013, *57* (4), 971–986.

Keefer, Philip and Razvan Vlaicu, "Democracy, Credibility, and Clientelism," *Journal of Law, Economics, and Organization*, 2008, *24* (2), 371–406.

Khemani, Stuti, "Buying Votes Vs. Supplying Public Goods: Micro-evidence on the Implications of Political Clientelism," 2011.

Khwaja, Asim Ijaz and Atif Mian, "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market," *Quarterly Journal of Economics*, 2005, *120* (4), 1371–1411.

Kitschelt, Herbert and Steven Wilkinson, *Patrons, Clients, and Policies*, Cambridge: Cambridge University Press, 2007.

Lande, Carl H., *Leaders, Factions, and Parties: the Structure of Philippine Politics*, New Haven: Yale University Press, 1964.

_ , *Post-Marcos Politics : A Geographical and Statistical Analysis of the 1992 Presidential Election* Institute of Southeast Asian Studies, New York: St. Martin's Press, 1996.

Larreguy, Horacio, "Monitoring Political Brokers: Evidence from Clientelistic Networks in Mexico," 2013. Working paper.

_ **, Cesar Montiel, and Pablo Querubin**, "The Role of Labor Unions as Political Machines: Evidence from the Case of the Mexican Teachers' Union," 2013.

_ , **John Marshall, and Pablo Querubin**, "What is the Effect of Turnout Buying? Theory and Evidence from Mexico," 2014.

Lehoucq, Fabrice, "When Does a Market for Votes Emerge?," in Frederic Schaffer, ed., *Elections for Sale: The Causes and Consequences of Vote Buying*, Boulder, Colorado: Lynn Rienner, 2007.

Lizzeri, Alessandro and Nicola Persico, "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review*, 2001, *91* (1), 225–239.

Munshi, Kaivan, "Community Networks and the Process of Development," *Journal of Economic Perspectives*, 2014, *28* (4), 49–76.

_ **and Mark Rosenzweig**, "Networks, Commitment, and Competence: Caste in Indian Local Politics," *NBER Working Paper 19197*, 2013.

Naidu, Suresh, James A. Robinson, and Lauren Young, "Social Origins of Dictatorships: Elite Networks and Political Transitions in Haiti," *Working Paper*, 2015.

Nichter, Simeon, "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot," *American Political Science Review*, 2008, *102* (1), 19–31.

Padgett, John F. and Christopher K. Ansell, "Robust Action and the Rise of the Medici, 1400-1434," *American Journal of Sociology*, 1993, *98* (6), 1259–1319.

_ **and Paul D. McLean**, "Organizational Invention and Elite Transformation: The Birth of the Partnership Systems in Renaissance Florence," *American Journal of Sociology*, 2006, *111* (5), 1463–1568.

_ **and** _ , "Economic Credit in Renaissance Florence," *Journal of Modern History*, 2011, *83* (1), 1–47.

Querubin, Pablo, "Family and Politics: Dynastic Persistence in the Philippines," 2010.

_ , "Political Reform and Elite Persistence: Term Limits and Political Dynasties in the Philippines," 2011.

Robinson, James and Thierry Verdier, "The Political Economy of Clientelism," *Scandinavian Journal of Economics*, 2013, *115* (2), 260–291.

Schaffer, Frederic and Andreas Schedler, "What is Vote Buying," in Frederic Schaffer, ed., *Elections for Sale: The Causes and Consequences of Vote Buying*, Boulder, Colorado: Lynn Rienner, 2007.

Sidel, John, "The Philippines. The Languages of Legitimation," in Muthiah Alagappa, ed., *Political Legitimacy in Southeast Asia. The Quest for Moral Authority*, Stanford University Press, 1995.

Smith, Daniel, "Succeeding in Politics: Dynasties in Democracies," *UCSD, PhD Dissertation*, 2012.

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

_ **, Thad Dunning, Marcelo Nazareno, and Valeria Brusco**, *Brokers, Voters, and Clientelism*, Cambridge University Press, 2013.

Szwarcberg, Mariela, "Who Monitors? Clientelism and Democratic Representation in Argentine Municipalities," 2011.

Vicente, Pedro, "Is Vote Buying Effective? Evidence from a Field Experiment in West Africa," *Economic Journal*, 2014, *124*, F356–F387.

Wasserman, Stanley and Katherine Faust, *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge University Press, 1994.

Table 1: Descriptive Statistics

	(1)	(2)	(3)
	Politician	Non-Politician	T-test
Network Meaures			
Between	0.018	0.001	-300.491
	(0.035)	(0.003)	[0.000]
Eigenvector	0.217	0.015	-255.480
	(0.259)	(0.047)	[0.000]
Municipal Percentiles			
Between	88.950	49.973	-90.305
	(18.650)	(25.956)	[0.000]
Eigenvector	86.765	49.975	-76.695
	(21.168)	(28.847)	[0.000]

Notes: N=3,907,448. The standard deviations are in (parentheses) (Columns 1-2). In Column 3, the test statistics are reported along with the p-values [bracket].

Table 2: Family Networks and the Decision to Run for Office

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between	0.150***	0.137***	0.122***	0.120***
	(0.008)	(0.012)	(0.012)	(0.012)
Observations	3,907,448	3,907,448	3,907,448	3,907,448
R-squared	0.023	0.023	0.037	0.038
Panel B: Eigenvector Centrality				
Eigenvector	0.128***	0.093***	0.091***	0.089***
	(0.006)	(0.009)	(0.010)	(0.010)
Observations	3,907,448	3,907,448	3,907,448	3,907,448
R-squared	0.016	0.018	0.034	0.035

Notes: Results from family-level regressions. The dependent variable is a (normalized) dummy equal to one if someone with the family name ran in the 2010 mayoral elections. The network measures are normalized as well. Regressions control for the number of individuals with the family name, the number of villages where someone from the family lives and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Municipal fixed effects are included in Column 4. 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 3: Candidate Networks and Vote Share (Municipal)

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between	0.047**	0.059*	0.080**	0.064
	(0.023)	(0.033)	(0.036)	(0.070)
Observations	1,856	1,856	1,856	1,856
R-squared	0.002	0.003	0.057	0.445
Panel B: Eigenvector Centrality				
Eigenvector	0.027	0.032	0.026	-0.040
	(0.024)	(0.040)	(0.040)	(0.079)
Observations	1,856	1,856	1,856	1,856
R-squared	0.001	0.002	0.054	0.445

Notes: Results from municipal*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. Regressions control for the number of individuals with the family name, the number of villages where someone from the family lives and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Municipal fixed effects are included in Column 4. 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 4: Candidate Networks and Vote Share (Village)

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between	0.062***	0.048***	0.047***	0.066***
	(0.006)	(0.006)	(0.006)	(0.011)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.813
Panel B: Eigenvector Centrality				
Eigenvector	0.064***	0.052***	0.049***	0.076***
	(0.006)	(0.007)	(0.007)	(0.012)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.813

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Village fixed effects are included in Column 4. 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 5: Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Outliers		w/o	w/	Non-	Degree	Hometown
	1%	5%	10%	ARMM	weights	Parametric		Dummy
Panel A: Betweenness Centrality								
Between	0.055***	0.062***	0.079***	0.042***	0.053***	0.040***	0.047***	0.046***
	(0.007)	(0.014)	(0.030)	(0.005)	(0.007)	(0.006)	(0.007)	(0.006)
Observations	19,159	17,152	14,643	42,252	50,181	50,181	50,181	50,181
R-squared	0.798	0.807	0.818	0.804	0.791	0.797	0.786	0.786
Panel B: Eigenvector Centrality								
Eigenvector	0.037***	0.037***	0.054***	0.035***	0.051***	0.035***	0.051***	0.047***
	(0.007)	(0.010)	(0.015)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Observations	23,404	21,742	19,234	42,252	50,181	50,181	50,181	50,181
R-squared	0.794	0.798	0.804	0.803	0.791	0.797	0.786	0.786

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects and control for the number of individuals with the family name, the number of female with the family name, education levels in the family and occupation in the family. The standard errors (in parentheses) account for potential correlation within municipality. In Column 6, the specification includes dummies for each distinct value of each control variable (except the network measure). In Column 7, the specification also controls for degree centrality. In column 8 we control for a "hometown dummy" that takes a value of one for precincts located in the village with the most number of relatives of the candidate. * denotes significance at the 10%, ** at the 5% and, *** at the 1% level.

Table 6: Candidate Networks and Vote Share (Village - New Families)

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between*New Family	0.074***	0.061***	0.059***	0.070***
	(0.013)	(0.013)	(0.012)	(0.023)
Between*Old Family	0.059***	0.045***	0.044***	0.065***
	(0.006)	(0.006)	(0.006)	(0.011)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.813
Panel B: Eigenvector Centrality				
Eigenvector*New Family	0.081***	0.069***	0.068***	0.089***
	(0.010)	(0.011)	(0.011)	(0.019)
Eigenvector*Old Family	0.059***	0.047***	0.044***	0.071***
	(0.006)	(0.007)	(0.007)	(0.013)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.813

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Village fixed effects are included in Column 4. 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 7: Mechanisms

	(1)	(2)	(3)	(4)
	Between		Eigenvector	
Panel A: Incumbent Vs. Challengers				
Centrality	0.069***	0.054***	0.068***	0.054***
	(0.006)	(0.007)	(0.006)	(0.007)
Centrality*Incumbent	-0.023**	-0.022**	-0.015	-0.015
	(0.010)	(0.009)	(0.010)	(0.010)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.786	0.785	0.786
Panel B: Only 'Serious' Candidates				
Centrality	0.066***	0.051***	0.069***	0.055***
	(0.007)	(0.007)	(0.006)	(0.008)
Observations	34,423	34,423	34,423	34,423
R-squared	0.610	0.613	0.611	0.613
Panel C: Municipalities with one candidate only				
Centrality	0.010	0.009	-0.006	-0.008
	(0.018)	(0.015)	(0.037)	(0.027)
Observations	1,187	1,187	1,187	1,187
R-squared	0.588	0.621	0.588	0.621

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures (between in Columns 1-2 and eigenvector in Columns 3-4) are also normalized. All regressions include candidate fixed-effects. Regressions in Columns 2 and 4 control for the number of individuals with the family name, the number of female with the family name, education levels in the family and 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 8: Party Mates Networks and Vote Share (Village)

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between	0.062***	0.045***	0.044***	0.068***
	(0.006)	(0.006)	(0.006)	(0.011)
Between	0.014***	0.018***	0.018***	0.030***
(party mates)	(0.003)	(0.003)	(0.003)	(0.007)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.787	0.814
Panel B: Eigenvector Centrality				
Eigenvector	0.063***	0.049***	0.047***	0.076***
	(0.006)	(0.006)	(0.007)	(0.012)
Eigenvector	0.009**	0.014***	0.016***	0.033***
(party mates)	(0.004)	(0.004)	(0.004)	(0.009)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.814

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). In Columns 2-4 the variables are included for both the candidate and her party mates.Village fixed effects are included in Column 4. 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 9: Candidate Networks, Policy Choices and Candidate Traits

	(1)	(2)	(3)	(4)	(5)	(6)
	Support	Policy		Traits		
	Policies	Alignment	Honest	Approachable	Experienced	Connected
Panel A: Betweenness Centrality						
Between	-0.029	0.017	-0.028	-0.004	0.031	0.083**
	(0.019)	(0.082)	(0.080)	(0.090)	(0.040)	(0.040)
Observations	629	658	658	658	658	658
R-squared	0.932	0.765	0.729	0.699	0.862	0.785
Panel B: Eigenvector Centrality						
Eigenvector	-0.036	0.026	-0.040	-0.032	0.042	0.060
	(0.024)	(0.075)	(0.073)	(0.083)	(0.041)	(0.046)
Observations	629	658	658	658	658	658
R-squared	0.932	0.765	0.729	0.699	0.862	0.783

Notes: Results from village*candidate regressions. The dependent variable is the normalized alignment between the candidate promises and voters preferences (Column 1), the normalized adjusted support for the candidate's proposed policies and programs (Columns 2), the normalized average 'honesty rating' given to the candidate (Column 3), the normalized average 'approachability rating' given to the candidate (Column 4), the normalized average 'experience rating' given to the candidate (Column 5) and the normalized average 'political connections rating' given to the candidate (Column 6). The network measures are also normalized. All regressions include candidate fixed-effects and candidate fixed-effects and control for the number of candidate's relatives in the village. 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 10: Candidate Networks and Vote Buying Incidence (Village)

	(1)	(2)	(3)
Panel A: Betweenness Centrality			
Between*Challenger	0.299***	0.292***	0.290***
	(0.105)	(0.110)	(0.106)
Between*Incumbent	-0.053	-0.054	-0.076
	(0.054)	(0.053)	(0.070)
Observations	142	142	142
R-squared	0.484	0.485	0.533
Panel B: Eigenvector Centrality			
Eigenvector*Challenger	0.234*	0.226*	0.241*
	(0.127)	(0.135)	(0.127)
Eigenvector*Incumbent	-0.007	-0.008	-0.037
	(0.085)	(0.084)	(0.091)
Observations	142	142	142
R-squared	0.457	0.457	0.510

Notes: Results from village-level regressions. The dependent variable is the normalized vote buying incidence. The network measures are also normalized. All regressions control for municipal fixed-effects. Regressions control for the number of the challenger's and the incumbent's relatives in the village (Columns 2-3). Regressions in Column 3 control for village density, the relevant village-level network measure, a dummy equal to one if the village is classified as rural, average education, age, household size and length of residence as well as the share of the population that is female, that 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.

Appendix for Online Publication

A.1 Additional Tables

Table A.1: Candidate Networks and Vote Share (Village)

	(1)	(2)	(3)	(4)
Between	0.033***	0.030***	0.030***	0.035***
	(0.007)	(0.007)	(0.007)	(0.013)
Eigenvector	0.038***	0.033***	0.032***	0.055***
	(0.007)	(0.008)	(0.008)	(0.014)
Observations	50,181	50,181	50,181	50,181
R-squared	0.785	0.785	0.786	0.813

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Village fixed effects are included in Column 4. 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: Descriptive Statistics - Exclude Candidates with Previous Experience

	(1)	(2)	(3)
	Politician	Non-Politician	T-test
Network Measures			
Between	0.009	0.001	-98.67
	(0.019)	(0.003)	[0.000]
Eigenvector	0.145	0.015	-100
	(0192)	(0.046)	[0.000]
Municipal Percentiles			
Between	86.11	49.973	-49.23
	(20.223)	(25.956)	[0.000]
Eigenvector	83.550	49.979	-41..15
	(23.001)	(28.847)	[0.000]

Notes: N=3,903,684. The standard deviations are in (parentheses) (Columns 1-2). In Column 3, the test statistics are reported along with the p-values [bracket].

Table A.3: Family Networks and the Decision to Run for Office [New Families]

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between*New Family	0.043***	0.036***	0.031***	0.029***
	(0.005)	(0.008)	(0.008)	(0.008)
Between*Old Family	0.159***	0.156***	0.136***	0.135***
	(0.026)	(0.027)	(0.028)	(0.028)
Observations	3,907,448	3,907,448	3,907,448	3,907,448
R-squared	0.156	0.156	0.160	0.160
Panel B: Eigenvector Centrality				
Eigenvector*New Family	0.037***	0.031***	0.030***	0.029***
	(0.003)	(0.006)	(0.007)	(0.007)
Eigenvector*Old Family	0.286***	0.276***	0.247***	0.245***
	(0.038)	(0.039)	(0.039)	(0.039)
Observations	3,907,448	3,907,448	3,907,448	3,907,448
R-squared	0.156	0.156	0.160	0.161

Notes: Results from family-level regressions. The dependent variable is a (normalized) dummy equal to one if someone with the family name ran in the 2010 mayoral elections. The network measures are normalized as well. Regressions control for the number of individuals with the family name, the number of villages where someone from the family lives and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Municipal fixed effects are included in Column 4. 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.4: Candidate Networks and Vote Share (Municipal - New Families)

	(1)	(2)	(3)	(4)
Panel A: Betweenness Centrality				
Between*New Family	0.174***	0.202***	0.270***	0.163
	(0.038)	(0.047)	(0.061)	(0.110)
Between*Old Family	-0.032	-0.006	0.017	0.042
	(0.027)	(0.036)	(0.037)	(0.070)
Observations	1,856	1,856	1,856	1,856
R-squared	0.112	0.115	0.161	0.505
Panel B: Eigenvector Centrality				
Eigenvector*New Family	0.135***	0.145***	0.161***	0.010
	(0.035)	(0.045)	(0.048)	(0.082)
Eigenvector*Old Family	-0.064**	-0.051	-0.053	-0.052
	(0.028)	(0.044)	(0.043)	(0.082)
Observations	1,856	1,856	1,856	1,856
R-squared	0.114	0.115	0.160	0.504

Notes: Results from municipal*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. Regressions control for the number of individuals with the family name, the number of villages where someone from the family lives and the number of female with the family name (Columns 2-4), education levels in the family (Columns 3-4), occupation in the family (Columns 3-4). Municipal fixed effects are included in Column 4. 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.5: Candidate Networks and Vote Share (Village) PPCRV Sample

	(1)	(2)	(3)
Panel A: Betweenness Centrality			
Between	0.060***	0.056***	0.118**
	(0.021)	(0.021)	(0.048)
Observations	728	728	728
R-squared	0.831	0.832	0.847
Panel B: Eigenvector Centrality			
Eigenvector	0.058**	0.052**	0.130**
	(0.023)	(0.023)	(0.057)
Observations	728	728	728
R-squared	0.831	0.832	0.847

Notes: Results from precinct*candidate regressions. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name (Column 2-3). Village fixed effects are included in Column 3. 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.6: Candidate Networks and Vote Share - Excluding outlying municipalities

	(1)	(2)	(3)	(4)
Subsample	Panel A: Betweenness Centrality			
Between	0.047***	0.045***	0.048***	0.049***
	(0.006)	(0.006)	(0.006)	(0.008)
Observations	46,673	41,394	36,507	21,634
R-squared	0.784	0.782	0.784	0.772
	Panel B: Eigenvector Centrality			
Eigenvector	0.050***	0.049***	0.050***	0.057***
	(0.007)	(0.007)	(0.008)	(0.010)
Observations	46,673	41,394	36,507	21,634
R-squared	0.784	0.782	0.784	0.772

Notes: Results from precinct*candidate regressions. The sample excludes municipalities in the top and bottom 1% (Column 1), 5% (Column 2), 10% (Column 3), 25% (Column 4) in the population distribution. The dependent variable is the normalized vote share (measured as a proportion of the registered population). The network measures are also normalized. All regressions include candidate fixed-effects. Regressions control for the number of individuals with the family name and 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.

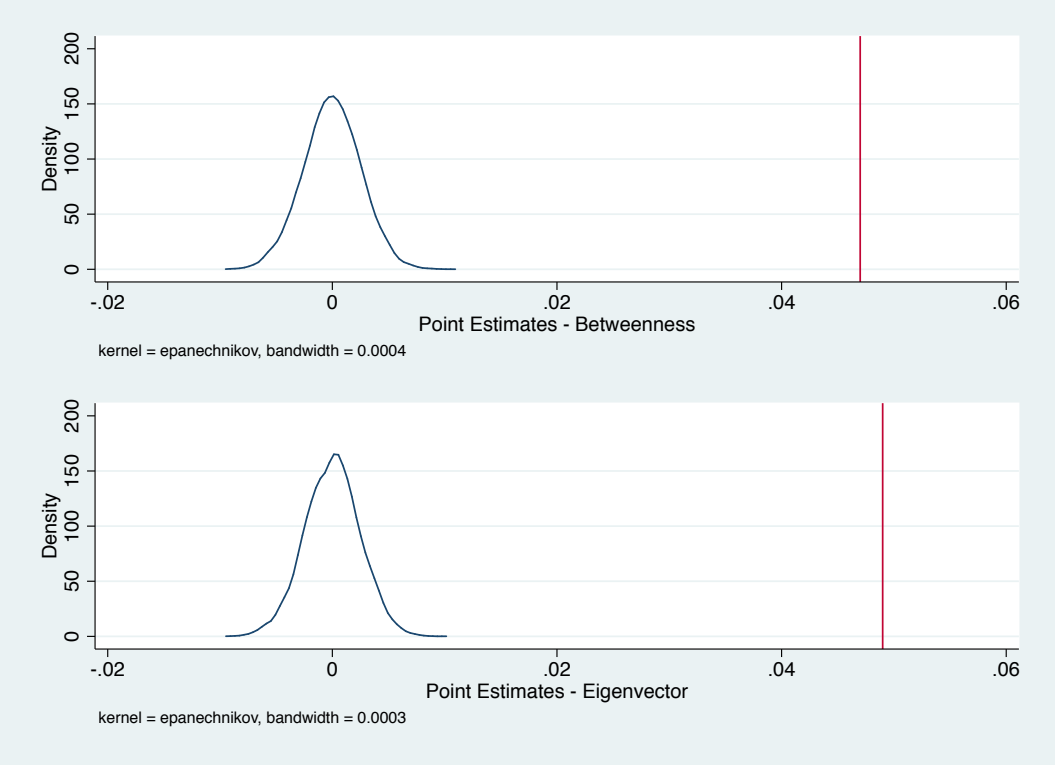


Figure A.1: Distribution of point estimates on betweenness [top panel] and eigenvector [bottom panel] centrality measures. Each point estimate is obtained from estimating regression (2) on a placebo dataset where the centrality values are randomly re-allocated across candidate and villages within the same municipality. In both panels, the red line represents the corresponding point estimates reported in Column 3 of Table 4.

A.2 Background on Centrality Measures

Degree Centrality

Degree centrality is the simplest measure, counting the number of ties that the politician's family has to other families. Following Wasserman and Faust (1994), we use two variants, a raw measure of the total number of connections, as well as an indexed measure that compares the total connections to the family with the highest total number of connections in the network. Since our ties represent intermarriages, they are undirected–that is, observing a tie from family A to family B implies that the same tie exists from family B to family A. As a result, we do not need to consider in-degree (inward) and out-degree (outward) ties.

$$Degree_i(f) = \sum F_{ij} \tag{4}$$

where F is the adjacency matrix of family network f, such that $F_{ij} = 1$ if there is a tie between nodes i and j, and 0 otherwise.

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 a family is proportional to the sum of centrality scores of the families it is connected to.

$$Eigenvector_i(f) \propto \sum F_{ij} * Eigenvector_j(f) \tag{5}$$

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).

Betweenness Centrality

Betweenness centrality is the extent to which the family serves as a link between different groups of families. It assesses centrality by looking at whether the family is an important hub in the paths traversing the network and is calculated using the number of shortest paths in the network that necessarily pass through the family (Freeman, 1977).

Following the notation in Jackson (2010), in the family network f, let $P_i(kj)$ indicate the number of shortest paths between family k and family j that necessarily pass through family i, while $P(kj)$ is the total number of shortest paths between k and j.

The ratio $P_i(kj)/P(kj)$ approximates the importance of family i in connecting k and j. If $P_i(kj) = P(kj)$, yielding a ratio of 1, then family i lies on all of the shortest paths connecting families k and j. Conversely, if $P_i(kj) = 0$, then family i is not important for connecting families k and j.

Betweenness centrality is calculated by averaging this ratio across all nodes (Freeman, 1977).

$$Betweenness_i(f) = \sum \frac{P_i(kj)}{P(kj)} \tag{6}$$

In our analysis, we normalize betweenness centrality for comparability:

$$Betweenness_i(f) = \sum \frac{P_i(kj)/P(kj)}{(n-1)(n-2)/2} \tag{7}$$