Electoral Rules and the Quality of Politicians:
Theory and Evidence from a Field Experiment in Afghanistan *
Andrew Beath† Fotini Christia‡ Georgy Egorov§ Ruben Enikolopov¶

8 March 2014

Abstract

We examine the effect of electoral rules on the quality of elected officials using a unique field experiment which introduced randomized variation in the method of council elections in 250 Afghan villages. In particular, we compare at-large elections with district elections. We propose a theoretical model where the difference in the quality of elected officials under the two electoral systems occurs because elected legislators have to bargain over policy, thus inducing citizens in district elections to vote strategically for candidates with more polarized policy positions that can secure resources for the district even at the expense of candidates’ competence. Empirical results prove consistent with the predictions of the model. Specifically, we find that elected officials in at-large elections are on average more competent than those in district elections and that this effect is stronger in more heterogenous villages.

We also find that elected officials in district elections have more biased preferences.

*We thank Hulya Eraslan, Horacio Larreguy, Mattias Polborn, and participants of the NBER Summer Institute Political Economy Public Finance meeting, Priorat Workshop on Bargaining and Politics, and Political Economy in the Chicago Area conference for helpful comments. Ruben Enikolopov acknowledges financial support from the Center for the Study of Diversity and Social Interactions at NES and financial support through Deutsche Bank Membership during his stay at IAS. Fotini Christia acknowledges support from ARO MURI award No. W911NF-12-1-0509.
†Office of the Chief Economist for East Asia and the Pacific, World Bank. abeath@worldbank.org
‡Department of Political Science, Massachusetts Institute of Technology. cfotini@mit.edu
§Northwestern University – Kellogg and NBER. g-egorov@kellogg.northwestern.edu
¶Barcelona IPEG, Universitat Pompeu Fabra, Barcelona GSE, New Economic School. ruben.enikolopov@upf.edu
1 Introduction

Selection of the right candidates both in in politics and in other settings is at least as important as providing them with right incentives. However, political economy literature has traditionally focused primarily on the issues of accountability.\(^1\) Recent literature starts to feel this gap by showing that the quality of politicians affects public policies (Besley, Pande, and Rao 2005; Martinez-Bravo 2013) and that political regime (Besley and Reynal-Querol 2011) and politicians’ wage (Ferraz and Finan, 2011, Gagliarducci and Nannicini, 2013) affect quality of politician. This paper contributes to this literature by examining how political institutions, and specifically electoral rules, affect the quality of elected officials.

We provide a theoretical model that generates a number of testable predictions, which prove consistent with empirical evidence from an innovative field experiment. In particular, we find that the quality of elected council members is higher in at-large elections, where all villagers vote for the whole council, than in district elections, where villagers can only vote for people in their electoral district. This effect is stronger in more heterogeneous villages. We also find evidence that district elections result in more polarized councils, a relationship that does not hold in at-large elections. We argue that these divergent results on candidate quality for district versus at-large elections are driven by strategic considerations of voters, who anticipate within-council bargaining and want to skew outcomes in their favor.

Our theoretical model considers a linear village with a uniform distribution of villagers; the location of a villager’s home also corresponds to his ideal point for location of a public good. Almost all villagers are uneducated, but there are a few (in the simplest version of the model just one) educated citizens, randomly distributed across the village. In line with the experimental setup, we consider a citizen-candidate model, where citizens elect a two-member council and may vote for any individual(s). In at-large elections, every citizen has two votes and can vote for any two villagers; in district elections, the village is split, geographically (and thus by policy preferences) into two districts, and each villager can cast a vote for a resident of the

---

\(^1\)The issue of political accountability has generated a rich line of theoretical and empirical work starting at least with the seminal Barro (1973). The importance of political selection, though emphasized as equally important as political accountability at least as early as in the Federalist papers (Madison, 1788), has been largely “neglected” (see Besley, 2005, p. 44).
same district. Once elected, the officials bargain over a joint policy decision. Thus, there is a tension between competence and political preferences of council members: In district elections, within-district median voters prefer more biased citizens to competent ones, because they expect them to achieve a better outcome in the bargaining game. In contrast, in at-large elections, the median voter is less bound by this strategic bias considerations and is more willing to elect competent citizens.

The empirical results come from a field experiment that randomized assignment of electoral rules for local council elections in 250 villages in Afghanistan. This experiment was conducted as part of a broader randomized impact evaluation of Afghanistan’s largest community driven development program, known as the National Solidarity Program (NSP). Each of the villages was randomly assigned one of two electoral rules – either district or at-large elections. Under district elections, the village is divided in several districts and candidates are elected to the village council from each district separately, with voters allowed to vote only for people who live within their assigned district. Under the alternative at-large procedure, voters have no restrictions for whom to vote and council members are elected based on the number of votes garnered across the whole village. Under both electoral rules, all villagers are considered candidates and formal political campaigning is prohibited.

We find that the quality of elected candidates, as measured by their educational attainment, is higher in villages with at-large elections and the difference is greater in more heterogeneous villages. To measure heterogeneity of villages we look at the divergence of villagers’ preferences measured before the intervention, as well as at village size and at the ethnic composition of villages. The results indicate that in homogenous villages there is no significant effect of electoral rules on the quality of elected officials. There is also some evidence that district elections lead to the election of candidates with more biased preferences, with location of the candidates’ houses serving as a proxy for their preferences with respect to the location of projects. Overall, the empirical results prove consistent with the theoretical predictions of the model.

An extensive literature analyzes the effects of electoral systems in general, and district magnitude in particular, on outcomes such as the number of viable candidates, representation of minorities, decisiveness, and support for redistribution. They find that the number of viable
candidates increases with district magnitude (Duverger, 1956; Cox, 1997; Norris, 2004), and that proportional representation, which is characterized by high district magnitude, is more favorable to minorities (Lijphart, 2004), although this effect depends on the geographic concentration of minorities (Moser, 2008), their social status (Moser and Scheiner, 2012), as well as the size of the minority (Trebbi, Aghion, and Alesina 2007). Electoral rules also affect geographic representation, with single-member districts and high-threshold proportional representation systems being better suited for district-bound, geographically-concentrated interests, with low threshold proportional representation systems favoring geographically-dispersed interests (Ferree, Powell, and Scheiner, 2013). Electoral systems also affect the incentive for personal vote, which decreases with district magnitude when there is no intra-party competition, but increases with district magnitude when such competition is present (Carey and Shugart, 1995). The type of electoral system also influences party formation, as well as government composition, with proportional representation systems favoring center-left governments that support higher levels of redistribution as compared to majoritarian systems (Iversen and Soskies 2006).

Looking at the link between electoral rules and the level of corruption and the types of public goods provided, proportional systems have been found to lead to higher levels of political rent extraction (Persson and Tabellini 1999, 2000). Larger voting districts are associated with less corruption, whereas larger shares of candidates elected from party lists are associated with more corruption (Persson, Tabellini, and Trebbi 2003). Meanwhile, majoritarian elections have been found to lead to smaller governments and welfare programs than those resulting from proportional elections (Persson and Tabellini 2004).²

Empirical literature on the determinants of the quality of politicians indicates that democracies are more likely to select highly educated leaders (Besley and Reynal-Querol 2011). Education appears to be an appropriate measure for the quality of politicians, as elected candidates with higher educational levels are less likely to use political power opportunistically (Besley, Pande, and Rao 2005) and improves public goods provision (Martinez-Bravo, 2013). Increasing intra-party competition improves politicians’ quality (Folke, Persson, and Rickne 2014) and

²Other works that look at the effect of electoral rules on the composition of government spending include Lizzeri and Persico (2001), Milesi-Ferretti, Perotti and Rostagno (2002).
inter-party competition also matters, as while a party that has an electoral advantage can slack on the quality of its candidates, a disadvantaged party has to provide candidates of higher quality in order to remain competitive (Banerjee and Pande 2007). There is also evidence that paying a higher wage has a positive effect on politicians’ quality (Ferraz and Finan, 2011, Gagliarducci and Nannicini, 2013), whereas higher budgets attract politicians of lower quality (Brollo et. al., 2013).

The theoretical literature examining the determinants of the quality of politicians suggests that suboptimal candidates can be elected when there is asymmetric information about their competence, as highly competent individuals have better outside options, leading to overrepresentation of less competent ones in the candidate pool (Caselli and Morelli 2004). Other papers emphasize strategic reasons behind choosing politicians other than the best available, such as a general equilibrium effect, whereby a party chooses mediocre individuals so as not to pay the very high market wage for the best and most competitive candidates (Mattozzi and Merlo 2007). Another line of argument, focuses on institutional constraints, like seniority rules, which give rise to an incumbency advantage that may prevent competent challengers from taking over (McKelvey and Reizman 1992). Yet others highlight how farsighted government members could oppose competence-improving changes out of fear that they would be ousted (Acemoglu et al. 2010) or betrayed in the future (Egorov and Sonin 2011). Also, competition over different political positions can lead to a competent but biased candidate’s failure to get elected (Aragones and Palfrey 2004).

Though the aforementioned works examine the effects of electoral rules and the quality of politicians separately, Myerson (1993) is the only work that considers them jointly, providing a theoretical model that examines the effect of district magnitude on the quality of politicians and arguing that small district magnitude together with strategic voting increases the barriers to entry in the electoral system, which in turn has a negative effect on the quality of politicians.4

Our paper provides a rare combination of an original theory along with results from a large-

---

3Banks and Sundaram (1998) study optimal contracts in a model where retention is the only reward mechanism. They find that retention helps select better candidates, but longer tenure undermines incentives to exert effort.

4See also Adams (1996), who considers the effect of a constitutional change in Illinois in 1980 on business friendliness of state legislators.
scale field experiment that confirms the model’s theoretical predictions. Specifically, our model makes a theoretical contribution by showing that electoral rules can affect the quality of elected officials even in a purely citizen-candidate environment, in the absence of entry barriers, party politics, and exogenously more numerous or advantaged groups. Our empirical contribution is also innovative, as to the best of our knowledge, ours is the first paper to exploit a field experiment that identifies the causal effect of electoral rules.

The rest of the paper is organized as follows: Section 3 describes the theoretical model, section ?? analyzes the model, and section 4 formulates the model’s empirical predictions. Section 2 presents the experimental design. We describe the data in section 5, offer the empirical results in section 6 and discuss them in section 7. We conclude in section 8.

2 Experimental Design

We examine the effect of electoral rules on the quality of elected officials using a field experiment, which randomized variation in the method of council elections in Afghan villages. This intervention was part of an impact evaluation of the National Solidarity Program (NSP) that randomized assignment of not only electoral rules, but also project selection procedures (Beath, Christia and Enikolopov 2013b), as well as the program itself (Beath, Christia, and Enikolopov 2012, 2013a). This section provides further details on NSP (subsection 2.1), describes the variation in electoral rules introduced across the 250 treatment villages under evaluation (subsection 2.3), details the sample and randomization procedures (subsection 2.4), and discusses the timing of the intervention and the data collection process (subsection 2.5).

2.1 Setting

Village council elections come as part of the National Solidarity Program (NSP), which was devised in 2002 by the Afghan Government to deliver services and infrastructure to the country’s rural population and build representative institutions for village governance. NSP has been implemented in over 32,000 villages in all of Afghanistan’s 34 provinces and has disbursed over $1.1 billion, making it the largest development program in Afghanistan. The program is
structured around two interventions: (i) the creation of an elected Community Development Council (CDC); and (ii) the disbursement of block grants to councils for implementation of village projects. The program is executed by the Afghan Ministry of Rural Rehabilitation and Development, implemented by contracted NGOs, and funded by bilateral and multilateral donors.

In order to facilitate the creation of representative institutions for village governance, NSP mandates the creation of gender-balanced councils through a secret-ballot, universal suffrage election.\textsuperscript{5} Once councils are formed, NSP disburses block grants valued at $200 per household, up to a village maximum of $60,000, to fund local development projects,\textsuperscript{6} with villages required to contribute at least 10 percent of project costs, which they largely do in the form of labor. Projects are selected by the council in consultation with the village community. Projects are ordinarily focused on either the construction or rehabilitation of infrastructure, such as drinking water facilities, irrigation canals, roads and bridges, or electrical generators; or the provision of human capital development, such as training and literacy courses. Overall, the main task of elected council members is to guide the choice of development projects and to then oversee project implementation.\textsuperscript{7}

NSP aspires to provide repeat block grants to participating villages, although villages receive no firm guarantees of when – or if – they will receive these. The process for conducting follow-up elections for the council is also uncertain. Per NSP rules, villages are supposed to hold re-elections for council positions every four years, although as follow-up elections are not facilitated, it is unclear whether these occur.\textsuperscript{8} Given this, and the general uncertainty that accompanies planned future development activity in Afghanistan, villagers perceive NSP as a one-shot event, which does not provide strong re-election incentives to council members.

The average population in our sample of villages is roughly one thousand people (see Table 1). There is a notable variation in the geographic size of villages, with quite a few villages

\textsuperscript{5}Note that this is the first time that the population is participating in a local election. Prior to that, villages in Afghanistan had only customary local governance structures (Beath, Christia, and Enikolopov 2013c)

\textsuperscript{6}The average block grant in the villages included in the sample was approximately $31,000.

\textsuperscript{7}There is also evidence that councils assume some additional responsibilities traditionally accorded to customary leaders, such as mediating conflicts, providing emergency assistance, and certifying documents (Beath et al. 2010).

\textsuperscript{8}No such elections had occurred by early 2012, when the data collection for this evaluation was completed.
that span several kilometers. The average distance between the house of a randomly selected survey respondent and the center of the village is about 400 meters, with a standard deviation of more than one kilometer. About 25 percent of villages are ethnically mixed, with the rest being exclusively Pushtun, Tajik, or Hazara (as well as one Turkmen village). The average education level in the sampled villages is very low, with more than seventy percent of adult male villagers having no formal education whatsoever and only four percent having finished high school. An average household consists of about ten people, of which about five are children under the age of fifteen. About two thirds of male heads of household are employed in agriculture and only forty five percent of respondents never or rarely have problems supplying food for their families.

2.2 Local Governance in Afghanistan

Afghanistan’s central government has historically lacked the strength and resources to exercise local control or provide public goods in many parts of the country. As a result, local communities have had their own structures of governance and accountability (Barfield 1984). The foundation of governance in rural Afghanistan is the local jirga or shura, a participatory council that has traditionally managed local public goods and adjudicated disputes (Nojumi, Mazurana and Stites 2004). Shura members tend to be the elders of families in the village (Rahmani 2006), although membership is ordinarily not fixed. Shuras generally convene when there is an issue to resolve and reach their decisions based on consensus (Boesen 2004). In addition to councils, villages ordinarily have a headman (termed a malik, arbab, or qariyadar) - usually a large landowner - who serves as liaison between the village and the central government (Kakar 2005). Traditional leadership in rural Afghanistan consists almost exclusively of males, as the principle of purdah - which stipulates that women should be generally hidden from public observation - precludes female involvement in communal gatherings and thus from local governance.

A key contrast between elected councils and customary governance institutions is the mode of selection and respective accountability structure. While elected councils involve a secret ballot, universal suffrage election, the position of headman is ordinarily inherited or otherwise derived on account of land holdings or other forms of economic authority. Although there is no formal assignment of local governance functions to elected councils, their authority in selecting,
implementing, and managing NSP-funded projects provides them with control over what is, for many Afghan villages, an unprecedented volume of resources. Thus, although the creation of an elected council does not directly usurp the major administrative tasks undertaken by the headman or other customary village institutions, the elected council exists as an institution vested with substantial authority and in parallel to customary governance structures.

2.3 Electoral Rules

Secret-ballot elections of the council were open to all adult residents and were conducted according to one of two sets of electoral rules that differ primarily in terms of district magnitude. Under both sets of rules, every village resident, whether male or female, aged eighteen years or older, who has lived in the village for at least one year, is eligible to vote or be elected to the council. NSP rules require that at least 60 percent of eligible voters must cast votes in the election for it to be valid. Villagers interested in getting elected to the council are prohibited from campaigning in any way for the position. The council has to contain an equal number of male and female members, with the total size being roughly proportional to the number of families residing in the village. All villages in the sample, regardless of the set of electoral rules, were segmented into geographically contiguous districts containing between 5 and 25 families depending on village size, and each district had its own polling station. A village map with districts and enclosed dwellings was displayed in a public area in the village. Further details of the two sets of electoral rules are provided below:9

District Election: Voters were restricted to casting a ballot for a single candidate, who had to reside in the same district. In each district, the one male and one female with the largest number of votes were elected to the council as representatives of their district. Thus, this method represents a single-ballot, simple plurality election with multiple districts (Cox, 1997), similar to the Anglo-American first-past-the-post system.

At-large Election: Under this method, voters could cast their ballot for anyone residing in the village regardless of district. Council members were the men and women receiving the most votes.

9A detailed guide on the procedures is available at: http://www.nsp-ie.org/sti/stile.doc
votes across the village. To ensure a sufficient number of elected members, voters had to cast ballots for a maximum of three different people, who were not ranked.\textsuperscript{10} The at-large election method represents a multi-member election under a plurality rule with a single district and multiple non-transferable votes. Thus, the two main differences from district elections are: 1) higher district magnitude (multiple elected members instead of one) and 2) number of votes cast (three instead of one).\textsuperscript{11}

In all villages, council elections were organized and administered by ‘social organizers’ employed by the contracted NGOs acting as implementing partners. Monitoring results from a randomly selected set of 65 villages that held district elections and 66 villages that held at-large elections, including data from the monitors’ 784 polling station reports and interviews administered to 1,675 male voters, indicate that election procedures were professionally executed by the implementing NGOs and that, in general, villagers exhibited a good understanding of the function of the different electoral rules.\textsuperscript{12} Thus, monitoring results confirm high levels of compliance with the assigned treatment status.

2.4 Sample and Randomization

The randomization of electoral rules occurred in 250 villages assigned to receive NSP that formed the treatment group for the randomized impact evaluation of NSP. The 250 villages are evenly split across ten districts in northern, northeastern, eastern, central, and western Afghanistan (see Figure 1). Despite the necessary exclusion of southern areas from the sample due to security concerns, the 10 districts are broadly representative of Afghanistan’s ethnolinguistic diversity, with five predominantly Tajik districts, four predominantly Pashtun districts, one predominantly Hazara district, and two districts with significant populations of Uzbek and Turkmen

\textsuperscript{10}This means the system allows plumping, but not cummulation (Cox, 1997). Permitting three votes in at-large elections was requested by participating NGOs who considered it a high probability that, if villagers were accorded only one vote in at-large elections, the number of candidates receiving votes would be fewer than the number of council seats, thereby necessitating multiple rounds of voting which would not be feasible.

\textsuperscript{11}According to the theoretical model the difference between electoral rules is driven by the difference in district magnitude, as predictions of the model are robust to changing the number of votes cast by each voter.

\textsuperscript{12}A detailed description of the monitoring results can be found at: http://www.nsp-ie.org/reports/CDCE-MR.pdf
Radom assignment of electoral rules was made concurrently with the assignment of the two alternative project selection procedures. Specifically, 25 treatment villages in each district were paired to minimize differences in background characteristics within each pair (leaving one village unpaired) and then matched in pairs of pairs to form quadruples. Unpaired villages across districts were also grouped into two quadruples (leaving two villages unmatched). Each village within the quadruple (and the two unmatched villages) was then randomly assigned one of four combinations of council election rules and project selection procedures.

This assignment procedure ensures that each village in the sample had an equal probability of being assigned to each of the two electoral rules and this assignment was orthogonal to the assignment of project selection procedures. To account for stratification at the randomization stage, we include quadruple fixed effects in the empirical analysis (Bruhn and McKenzie, 2009).

The randomization resulted in a well-balanced set of villages. Table 2 presents a comparison between the two groups of villages with regard to a number of pre-intervention characteristics. The differences between the two groups never exceed 13 percent of the standard deviation.

2.5 Phasing of Intervention and Data Collection

The baseline survey was administered in September 2007, prior to the assignment of allocation procedures. Council elections—which were monitored, providing additional data on the processes—occurred between October 2007 and May 2008. Project selection occurred between November 2007 and August 2008, whereas project implementation occurred between April 2008 and September 2011.

---

13 An assessment of the demographic and economic characteristics of the sample villages reveal few substantive differences with those of a random sample of villages surveyed by the 2007-08 National Risk and Vulnerability Assessment.

14 Development projects were selected either through a secret-ballot referenda or at a village meetings. For more on those results see Beath, Christia and Enikolopov 2013b.

15 These characteristics include village size (based on data collected by Afghanistan’s Central Statistics Office) and a set of geographic variables (distance to river, distance to major road, altitude, and average slope).

16 Pairs of pairs were formed by performing the same matching procedure treating each pair as a single village with background characteristics that equal the average of the respective characteristics for the two villages in a pair.
3 Theory

The setup of the theoretical model is designed to take into account important features of the setting of the experiment, such as the absence of entry costs for candidates, absence of parties and political campaigns, etc. However, the model was not backwards engineered to match the empirical findings and all the empirical predictions except one were first formulated as theoretical propositions and only then tested empirically (see section 4 for more details).

3.1 Setup

The society consists of a continuum of individuals distributed uniformly on a compact $S = [-B, B]$. The policy space that these agents care about coincides with set $S$. The literal interpretation of the model is a society which cares solely about the location of a public good, such as a school or a water well; however, the results of the model naturally extend to a much more general set of environments where an individual’s preferences are correlated with geographic location. We assume that if policy $p \in S$ is enacted, then an individual with bliss point $b \in S$ gets the baseline utility function $u(p, b)$, which we assume, for simplicity, to be quadratic:

$$u(p, b) = -k (p - b)^2,$$

where $k > 0$ measures the importance of the policy issue to the society. In addition to different bliss points, individuals in the society differ by their competence, or education, which may be high or low: $a \in \{0, h > 0\}$. In other words, each citizen $i$ is characterized by a pair $(a_i, b_i)$, where the first component is his competence ($a$ for ability), and the second is his location ($b$ for bliss point). To study the trade-off between policy position and competence, we assume that almost all individuals are incompetent ($a_i = 0$), except for a finite number $N$ randomly picked ones, who have high competence $a_i = N$. The results and effects are most transparent when $N = 1$, which we assume for the rest of the paper; in Appendix A, we allow $N$ to be any number, and also further show robustness of the effects in the model. We assume that the types of all individuals are known to all other individuals; given the context of the experiment, it is natural to assume that location of villagers’ dwellings and education are observable to fellow
villagers. We also make a technical assumption that for any \( b \in [-B, B] \) there is a citizen \( i \) with \((a_i, b_i) = (0, b)\); this assumption that there is an incompetent citizen for any policy position ensures existence of an equilibrium.

Policy \( p \) is chosen and implemented by a governing body (henceforth “council”), which is elected by the citizens and from the citizens. We assume that the council consists of two elected individuals (again, this assumption is relaxed in Appendix A as part of a set of robustness checks), and both must agree on a policy for it to be chosen. We also assume that the competence of council members increases the quality of policy implementation, say, provides the public good better or earlier, which is consistent with the data. If the two council members have types \((a_l, b_l)\) and \((a_r, b_r)\) and implement policy \( p \), then individual \( i \) will get utility

\[
w_i(a_l, a_r; p) = a_l + a_r + u(p, b_i) = a_l + a_r - k(p - b_i)^2.
\]

To simplify exposition, we assume that the council members pick a policy that maximizes their joint utility: \( p = \frac{b_l + b_r}{2} \). Notice that this policy will be the outcome of a bargaining game with alternating offers (Rubinstein, 1982) or a legislative bargaining game with random recognition (Baron and Ferejohn, 1989), in the limit where offers are made very frequently. (If the offers are not made frequently, \( p = \frac{b_l + b_r}{2} \) is still the expected outcome of the game, but there is some variance, which will result in disutility for the citizens.) We consider this game explicitly when we generalize the game to incorporate councils with more than two members. Here, slightly abusing notation, we assume that having a council with members \((a_l, b_l)\) and \((a_r, b_r)\) yields utility

\[
w_i(a_l, b_l, a_r, b_r) = a_l + a_r \left( \frac{b_l + b_r}{2} \right) = a_l + a_r - k \left( \frac{b_l + b_r}{2} - b_i \right)^2.
\]

We compare two electoral procedures: at-large elections and district elections. In district elections, the society is divided into two districts: left \( L \), containing individuals with \( b_i < 0 \), and right \( R \), containing individuals with \( b_i \geq 0 \). This division is made according to the location of an individual’s residence; the two districts also differ by the policy preferences of their inhabitants. Each individual casts a vote for one of the citizens living in his/her district, i.e., every individual living in the district is considered a candidate, again in line with the field experiment. Then
in each district, the individual who got the largest share of votes is elected, and in the case of a draw, a random person among those who received the most votes is chosen. In at-large elections, the entire society comprises a single district, and each individual casts two votes for two (different) citizens.\textsuperscript{17} The two candidates who received the most votes get elected.

The strategy of each voter in district elections is therefore $\mathbf{\lambda}(i)$, the identity of the individual in his district for whom he casts his vote (since only members of the same district may be elected, $b_{\mathbf{\lambda}(i)} \in [-B, 0)$ if $b_i \in [-B, 0)$ and $b_{\mathbf{\lambda}(i)} \in [0, B]$ if $b_i \in [0, B]$). The strategy of each voter in at-large elections is $\mathbf{\Lambda}(i) = (\lambda_1(i), \lambda_2(i))$, which corresponds to the (unordered) pair of individuals for whom he votes. All voting decisions are made simultaneously, which gives rise to a coordination problem. We make the following refinement.

\textbf{Definition 1} Voting strategies $\{\lambda_i\}$ in case of district elections or $\{\Lambda_i\}$ in case of at-large elections, constitute an equilibrium if for any electoral district (i.e., $L$ or $R$ in the first case, or the entire society $S$ in the second) there is no subset of voters $X$ in this district who would strictly improve the utility of all voters in $X$ by choosing different voting strategies.

In other words, we refine the (otherwise huge) set of Nash equilibria by allowing for deviations by coalitions of voters, but only within a district.

\subsection*{3.2 Analysis}

Our analysis of the game is greatly simplified by the fact that both in at-large and in district elections, the median voter theorem applies. In district elections, each of the two districts $L$ and $R$ will elect the council member most favored by the median voter in that district, holding the decision of the other district fixed (denote these median voters by $m_L$ and $m_R$, respectively). In at-large elections, the median voter of the entire society, $m_S$, will elect the pair of candidates that he likes best. These individuals (or this pair of individuals) will, in fact, be the Condorcet winners in their respective districts. These results hold because individual preferences exhibit the single-crossing property: if a citizen $i$ prefers policy $p_1$ to $p_2 < p_1$, then so does a citizen $j$.

\textsuperscript{17}This assumption ensures the existence of an equilibrium, and that the median voter effectively determines the composition of the council. The results would remain intact if citizens voted sequentially, i.e., they elected the second council member after observing who was elected in the first round.
with \( b_j > b_i \). The fact that one of the citizens is competent (denote his policy preference by \( q \), so his type is \((h, q)\)) makes the argument just marginally more involved. The proof of Proposition 1 (below) fills in the details.

**Proposition 1** In both district elections and at-large elections, equilibria exist, and the types of elected politicians are uniquely determined for almost all realizations of \( q \). Moreover:

1. In district elections, the district without the competent citizen elects the most biased individual (with \( b_i = \pm B \)), and the district with the competent citizen elects either this citizen with type \((h, q)\) or the most biased individual (with \( b_i = \pm B \)).

2. In at-large elections, the two elected citizens are the most competent individual \((h, q)\) and a citizen with the opposite political preferences \((0, -q)\).

Proposition 1 implies that the equilibrium concept we use (Definition 1) is sufficiently strong to pick a (generically) unique equilibrium. To build an intuition for types of citizens who get elected, consider at-large elections first. The median voter in the whole district, \( m_S \), has bliss point \( b_M = 0 \) and his ideal outcome is to elect two council members who negotiate and implement his ideal policy 0, at the same time making sure that one of the two is competent. This is feasible: he can achieve this ideal outcome by having the competent citizen \((h, q)\) and his political antipode \((0, -q)\) elected.

Consider district elections. The reason to elect the most competent citizen if he lives in the district is clear, but what is the rationale to elect the most biased individual? To answer this question, suppose that district \( L \) elects a citizen of type \((a_l, b_l)\); consider the best response of the median voter of district \( R \), \( m_R \). His ideal policy is \( \frac{B}{2} \), and if he elects a resident with type \((a_r, b_r)\), he would get utility

\[
w_{m_R}(a_l, b_l, a_r, b_r) = a_l + a_r - k \left( \frac{b_l + b_r}{2} - \frac{B}{2} \right)^2.
\]  

(1)

The right-hand side of (1) is strictly increasing in \( b_r \) for \( b_r \leq B \), because \( b_l \leq 0 \), and thus it reaches its maximum for \( b_r = B \). In other words, holding competence fixed, the median voter of district \( R \) prefers the most biased candidate, and this is true regardless of voting strategies.
of citizens in the left district.\footnote{This preference for the most biased candidate would not necessarily hold if the distribution of individuals were non-uniform, for example, in the case of non-bounded support. However, the tendency to elect a relatively biased candidate would remain. We maintain the assumption of a uniform distribution for expositional purposes.} The same effect causes the median voter in the left district, $m_L$, to favor a candidate with $b_l = -B$. Of course, it is also possible that the most competent candidate will be chosen over the most biased, and the next proposition tells us exactly when this happens.

**Proposition 2** In district elections, both districts elect the most biased and incompetent candidates if

$$|q| < \tilde{q} \equiv 2B - \sqrt{\frac{4h}{k} + B^2},$$

(2)

where $q$ is the ideal point of the competent citizen. If (2) does not hold, then one district elects the most biased of its residents, and the other one elects the competent citizen. The competent citizen is more likely to be elected if:

(i) the society is more homogenous, i.e. less polarized in their preferences ($B$ is lower);

(ii) competence is more pronounced ($h$ is higher);

(iii) policy matters less relative to competence ($k$ is lower).

The district without the competent citizen is bound to elect the most biased council member, $(0, -B)$ or $(0, B)$. The median voter of the other district faces a trade-off between electing the most biased citizen and the competent citizen. Thus, he is more likely to choose competence over policy if competence is more important ($h$ is high and $k$ is low) or if the competent citizen is also biased ($q$ is close to $\pm B$). Interestingly, polarization hurts the chances of the competent citizen, and the reason is that high polarization makes the median voter more sensitive to the political preferences of the council member he elects.

We can now compare the expected outcomes of at-large elections with those of district elections. Ex ante, the identity of the competent individual is not known, but in expectation the following proposition holds.
Proposition 3 In at-large elections, as compared to district elections:

1. the expected competence of an elected council member is higher (strictly higher if \(4h < 3B^2k\)) and this difference is increasing in \(B\) and \(k\);

2. the expected polarization (distance between preferences of a council member and the society’s median voter, normalized by dividing by \(B\)) is (strictly) lower, and this difference is increasing in \(B\) and \(k\);

3. there is no correlation between preferences and competence of council members in at-large elections, and in district elections, competence and distance from the median voter are negatively correlated.

These results follow from Propositions 1 and 2 and they are especially easy to see on Figure 1, where we depicted the outcome of elections (types of council members) for different realizations of \(q\). In at-large elections, the competent individual is always elected, and this is not true in district elections, provided that \(\hat{q} > 0\). Interestingly, if \(h\) is high or \(k\) is low or \(B\) is low, the competent individual will be elected in both cases, and the difference between the two types of elections disappears. The result on polarization is easy to see from the following consideration: the two council members elected in at-large elections are as far from the median as the competent one, while in district elections, one or both districts elect individuals who are further from the median than the competent individual. In addition, if \(\hat{q} > 0\), then in district elections, the most moderate types will never be elected. Finally, in at-large elections, there is no correlation between preferences and competence of a council member: as one can see from Figure ??, any council member with any political bliss point is equally likely to be competent or incompetent. In contrast, in district elections, the most biased council members are likely to be incompetent, and any council member with a more moderate ideal point is likely to be competent.\(^{19}\)

\(^{19}\)The ability of the median voter to choose both council members at once help him achieve the first best, but it does not drive the results in Proposition 3, as we show in Appendix A (Subsection 8.2 that at-large result in more competent council members even if one member is elected at a time, and Subsection 8.1 shows that at-large elections may lead to a worse outcome than district elections if offers are made infrequently). The results are entirely driven by the bargaining and compromise of the council members, which is anticipated by voters at the time of elections.
Apart from the empirical predictions about competence and polarization, our model has clear welfare implications.

**Proposition 4** In at-large elections, compared to district elections, the expected utility of every individual is higher, and thus social welfare is higher.

In light of Proposition 3, it is not surprising that social welfare is higher, in expectation, in at-large elections. It is more striking that the expected utility of every single individual is higher in at-large elections, provided that expectation is taken before the location of the competent citizen becomes known. The intuition, however, is simple: the expected policy is \( \mathbb{E}p = 0 \) under both procedures, and by moving from at-large to district elections, the society makes the policy outcome more uncertain and runs the risk of electing two incompetent citizens; both effects hurt every citizen equally.

## 4 Empirical Predictions

The theoretical results give rise to several empirical predictions, which we test using data from a field experiment.\(^{20}\) Based on the first statement of Proposition 3 we can formulate the following.

\(^{20}\)Note that although the theoretical model was formulated after the completion of the field experiment, theoretical predictions were not constructed by backwards engineering the empirical results. All the empirical predictions except the first one were first formulated as the theoretical predictions and only then tested empirically.
empirical prediction:

1. The quality of elected candidates is higher in at-large elections as compared with district elections.

Similar to Besley et al. (2005), we use educational attainment as a measure of candidate quality, since there is evidence that the leaders’ level of education has a positive effect on governance outcomes. Specifically, using cross-country evidence, Besley et al. (2010) show that higher education of a country’s leaders is associated with higher economic growth, whereas Martinez-Bravo (2013) provide evidence of the causal effect of education of village leaders on public goods provision in Indonesia. In Section 6 we show that in the context of our experiment, educational attainment is associated with better council performance.

The first statement of Proposition 3 also asserts that the effect of the electoral system on the quality of politicians is stronger in communities with more diverse preferences (higher $B$) where the choice of policy is relatively more important than the politicians’ competence (higher $k$), so that for communities with sufficiently homogenous preferences (where $3B^2k < 4h$) there should be no differences between electoral systems. Thus, we can formulate the following two empirical predictions:

2. The difference in the quality of elected candidates between at-large and district elections is higher in more heterogenous villages.

3. In homogenous villages the quality of candidates does not depend on the electoral method.

In the empirical analysis, we use three alternative measures of heterogeneity – fractionalization of preferences over projects, ethnic heterogeneity, and geographic size of village. The choice of these measures is driven by the nature of the tasks performed by council members. As noted above, the main tasks of the elected council members is to guide the choice over development projects and then oversee the implementation of these projects. At the project selection stage, a candidate’s preferences over projects can affect both the type and the location of development projects, which can be treated as the policy dimension in the model.

Fractionalization of preferences over projects, directly measures heterogeneity of interests in terms of project type. More divergent villager preferences with respect to project type,
correspond to higher $B$ in the model. In addition, quality of implementation is likely to be more important if preferences are aligned and people get their preferred project; at the same time, if preferences are diverse, then policy choice becomes more important. Thus, fractionalized preferences correspond to the case where policy is important relative to competence, i.e., high $k$.

Ethnic heterogeneity is often used in the literature to capture differences in tastes (e.g. Alesina, Baqir, Easterly 1999) that might affect preferences over both project type \(^{21}\) and location (if villages are ethnically segregated),\(^{22}\) so ethnically heterogenous villages are also likely to have higher $B$ and $k$.

Finally, preferences over project location are likely to be driven by the location of voter and candidate residences. The larger the village, the more divergent the preferences of villagers with respect to project location (which corresponds to higher $B$), and the more important the policy dimension relative to the quality of candidates (which corresponds to higher $k$).

According to the model, the quality of elected candidates improves outcomes for any policy choice. In the context of the experimental set up, the quality of elected candidates is likely to affect project implementation regardless of their type and location. In particular, council members of high quality are likely to implement projects faster. Thus, we can formulate the following empirical prediction:

4. The speed of project implementation is higher in at-large elections as compared with district elections.

According to the second statement of Proposition 3, expected polarization is lower in at-large elections, leading us to the following prediction.

5. Elected council member preferences are less biased under at-large elections as compared to district elections.

In the model, polarization corresponds to the distance between council member preferences and the median voter. In the empirical analysis, we focus on candidate preferences with respect to

\(^{21}\)Which is confirmed in our context, as ethnic heterogeneity is correlated with fractionalization of preferences over types of projects (the correlation is significant at the 5 percent level)

\(^{22}\)Unfortunately, the number of observations within villages is not large enough to test for the existence of segregation directly.
project location, taking literally the assumption that the geographic location of elected candidates and voters reflects their policy preferences. Thus, we use the distance between an elected council members’s house (which reflects his preferences) and the center of the village (which reflects the preferences of the median voter) as a proxy for the bias in his preferences. Finally, the third statement of Proposition 3 corresponds to the following empirical prediction:

6. In district elections there is a negative correlation between candidate quality and his bias in terms of preferences, with no such relationship in at-large elections.

We test this prediction using the same measures as above - educational attainment as a proxy of candidate quality and the distance between an elected candidate’s house and the center of the village as a proxy for the bias in his preferences.

5 Data

The data used in the empirical analysis come from several different sources. Information on the characteristics of elected council members was supplied by the implementing NGOs. Specifically, the data contains information on electoral results and personal characteristics of all candidates, including their gender, age, educational attainment, and occupation, as well as their district of residence. Data was provided for 2,044 male candidates from 241 villages.\textsuperscript{23} We construct a dummy variable that equals one if the person has full secondary education (i.e. finished high school) and zero otherwise and a similar measure for getting lower secondary education (i.e. finishing middle school). Among male candidates only 9% have finished high school, and 17% have finished middle school (see Table 2).

Information on the geographic distribution of elected council members is based on GPS coordinates of their residences collected during the monitoring of program implementation. This information is available only for a subset of villages, since some monitors failed to follow the instructions and did not collect the required information. In addition, some coordinates con-

\textsuperscript{23}Of the 9 villages for which the data was not received, 7 villages did not comply with the NSP treatment assignment, an error driven primarily by confusion between villages with similarly-sounding names. For the remaining 2 villages, the relevant NGO did not provide the necessary information. In both cases, attrition was not correlated with the assigned electoral rules.
tained obvious mistakes and were excluded from the analysis.\textsuperscript{24} The resulting database contains coordinates of the residences of 1,104 male council members in 140 communities in nine out of ten evaluation districts. Although sample attrition is substantial, there are no reasons to believe that it was systematic in a way that can be driving the corresponding results.

We restrict our analysis to male candidates for two reasons. First, in Afghan villages, as discussed above, women are traditionally excluded from community-level decision making. The female council members elected through the gender quota are therefore not expected to have a significant effect on project selection and implementation.\textsuperscript{25} Second, the level of education of female candidates was so low, that there was almost no variation in educational attainment. In particular, more than 90\% of female candidates did not have any formal education and only 0.8\% had finished high school.

Information on the size and ethnic composition of villages, as well as information on the preferences over the types of development projects comes from the baseline survey, conducted before the start of the program. The survey was administered to ten randomly selected male heads of household in each village. The resulting dataset contains information on demographic and socio-economic characteristics for 2,387 male heads of household.

To construct a measure of heterogeneity of preferences with regard to the types of development projects, we use a question from the baseline survey in which respondents were asked to indicate – from a list of fifteen potential projects that correspond to the types of projects sponsored by NSP – the project that they believed should be selected if the village was provided with a $60,000 grant.\textsuperscript{26} In each village \(v\) for each type of project \(j\) we calculate the share \(s_{vj}\) of respondents that indicated this project and calculate fractionalization of preferences in the village \(f_v = 1 - \sum s_{vj}^2\). Next, we construct a dummy variable, which equals one if fractionalization is above the median in the sample and zero otherwise. We also calculate the average distance

\textsuperscript{24}These mistakes were identified by superimposing location of council members’ houses and satellite images of the respective villages and identifying the instances in which coordinates did not belong to the village. We documented all such instances and ensured that the data cleaning exercise was conducted in absence of information about the treatment status of the villages, so that it does not bias the results.

\textsuperscript{25}This is confirmed by the results in Beath, Christia and Enikolopov (2013b) which indicate that the preferences of even the most important female villagers do not have a significant effect on the choice of projects.

\textsuperscript{26}Results in Beath, Christia, and Enikolopov (2013b) indicate that the preferences revealed in the baseline survey are highly correlated with the subsequent choice of development projects.
between the households of the respondents based on the GPS coordinates of their houses and construct a dummy variable for large villages which equals one if the average distance is above the median in the sample and zero otherwise. As a measure of ethic heterogeneity of villages, we use a dummy variable that equals zero if all villagers belong to one ethnicity and one otherwise.\textsuperscript{27}

Information on project implementation comes from the administrative records of NSP’s central office. The data contains information for 1,119 projects that were implemented in the villages as part of NSP and includes the dates for when implementation of each project started and when it ended. The start dates range from April 2, 2008 to February 22, 2011. The completion dates range from July 23, 2008 to September 24, 2011.\textsuperscript{28} Based on this information, we construct five dummy variables for whether a project started before the beginning of the fourth quarter of 2008 and each of the quarters in 2009. Similarly we construct five dummy variables for whether a project started before the beginning of each of the quarters in 2009 and the first quarter of 2010. Figures A1 and A2 in the Online Appendix show the distribution of dates for project commencement and completion, as well as the dates for which we construct the indicators used in the empirical analysis.

6 Empirical Results

To provide evidence that educational attainment is a reasonable proxy for council members’ quality, we check that it is associated with better council performance. In particular, we examine how council member education affects the speed of project implementation using the following OLS model:\textsuperscript{29}

\textsuperscript{27}The measure is based on a question that asks heads of households to indicate their ethnicity with seven options: Pashtun, Tajik, Hazara, Uzbek, Turkmen, Baluch, and other. All results hold, if we use a similar measure based on the question from the focus group of village leaders that asked them to indicate what ethnicities reside in the village or if we use a measure of ethnic fractionalization instead, although the latter measure of fractionalization is not very reliable given the small number of observations per village.

\textsuperscript{28}Six projects failed and were not completed.

\textsuperscript{29}In all specifications with binary dependent variables we use linear models, since our specifications are close to a saturated model (Angrist and Pischke 2008), while linear specification makes interpretation of the coefficients easier and avoids the incidental parameter problem caused by the inclusion of quadruple fixed-effects. The results are robust to using a logit model instead.
\[ \text{Progress}_{ijv} = \alpha + \beta \cdot \text{EducationCouncil}_v + \gamma \cdot \text{EducationVillagers}_v + \phi_q + \varepsilon_{jv} \] (3)

where \( \text{Progress}_{ijv} \) is an indicator for whether project \( j \) in village \( v \) has started or finished by a specific date, \( \text{EducationCouncil}_v \) is an indicator for whether at least one council member has finished high school (which is true for 32% of the villages), \( \text{EducationVillagers}_v \) is an indicator for whether at least one of the villagers in the baseline survey has finished high school (which is true for 26% of the villages), \( \phi_q \) is the quadruple fixed effect, and \( \varepsilon_{jv} \) is the error term.

Results of the analysis indicate that both villager education and council member education are important predictors of the project’s implementation progress (see Table 3). In villages with more educated council members projects were more likely to start earlier and to be completed faster. Although these results do not prove causal effect of council member education on project implementation, they provide evidence that their educational attainment is associated with better outcomes, and thus can be used as a proxy for candidate quality.

To test the first empirical predictions of the model we use the following OLS model:

\[ \text{Education}_{vi} = \alpha + \tau \cdot \text{AL}_v + \phi_q + \varepsilon_{vi} \] (4)

where \( \text{Education}_{vi} \) is a dummy variable for whether candidate \( i \) in village \( v \) has finished high school, \( \text{AL}_i \) is a dummy variable, which equals one if the village \( v \) has been assigned at-large elections and zero if the village has been assigned district elections, \( \phi_q \) is the quadruple fixed effect, and \( \varepsilon_{vi} \) is the error term. Standard errors are clustered at the village level. The first empirical prediction posits that \( \tau > 0 \).

Results indicate that male council members elected through at-large elections are better educated as compared with male council members elected through district elections (see column 1 in Table 3). The share of council members who have finished high school among the council members elected under at-large elections is higher by 4 percentage points as compared to the

\(^{30}\)The results are robust to using dummy variable for finishing middle school (see Table A1 in the Online Appendix)
average of 7 percent among councils members elected under district elections (the difference is statistically significant at the 1% level). Although modest in absolute magnitude, this effect constitutes a 57 percent increase in the share of council members who have finished high school. Given the overall low level of education of council members and the very low level of education and literacy in the Afghan rural context, it can have a noticeable effect. Thus, the results confirm the first empirical prediction of the model.

The second empirical prediction posits that the effect of at-large elections on the quality of elected council members is higher in more heterogeneous villages, while the third empirical prediction states that the effect of at-large elections is insignificant in homogenous villages. To test the second and the third empirical predictions of the model we use the following OLS model:

$$\text{Education}_{v_i} = \alpha + \gamma \cdot AL_v + \delta \cdot AL_v \cdot \text{Heterogen}_v + \mu \cdot \text{Heterogen}_v + \phi_q + \varepsilon_{v_i}$$

(5)

where \(\text{Heterogen}_v\) is a measure of heterogeneity of village \(v\) and all other variables are the same as in equation (4). The second empirical prediction posits that \(\delta > 0\) and the third one that \(\gamma = 0\) (as long as the measure of heterogeneity is normalized to zero in homogenous villages).

Results of the empirical analysis indicate that the effect of at-large elections is indeed significantly stronger in more heterogeneous villages for all measures of heterogeneity (columns 2-4 in Table 3), which is consistent with the second empirical prediction. Moreover, consistent with the third empirical prediction, in homogenous villages, there is no significant effect of electoral rules on the quality of elected candidates. The results also provide some evidence that increasing heterogeneity is associated with lower quality of elected candidates in villages with district elections, which is consistent with Proposition 2. The latter result, however, should be treated with caution, as it is not based on randomized variation in electoral rules, so it can be driven by endogeneity bias.

According to the fourth empirical prediction, the speed of project implementation should be higher in at-large elections as compared with district elections. To test this hypothesis we estimate the following OLS model:
\[ \text{Progress}_{jv} = \alpha + \lambda \cdot AL_v + \phi_q + \varepsilon_{jv} \]  

(6)

where all the variables are the same as in (3) and (4). According to the fourth empirical prediction \( \lambda > 0 \).

The results indicate that in villages with at-large elections, project implementation was more likely to start earlier and to be completed faster. The magnitude of the effect is noticeable, with projects being 5 percent more likely to have started before July 2009 and 5 percent more likely to have finished before July 2009. The results are robust to controlling for the type of implemented projects, so that the effect is not driven by the choice of different types of projects in villages that held at-large elections. This is consistent with the finding that there is no significant effect of electoral system on the type of selected projects, which is consistent with the prediction of the model that on average, there is no difference in policies implemented under at-large and district elections. Thus, the results are consistent with the fourth empirical prediction.\(^{31}\)

The fifth empirical prediction posits that district elections will lead to the election of candidates with more biased preferences. We test this prediction by estimating a model similar to (4) with the logarithm of the distance between the house of an elected council member and the village center (calculated by averaging the coordinates of the houses of baseline survey participants) as the outcome variable.\(^{32}\) To account for differences in the size of the village, we can include as an additional control a measure of the distance between the villagers’ homes in the baseline survey. Results indicate, consistent with the empirical prediction, that the distance between the homes of elected officials and the village centers is smaller in at-large elections (see Table 6). Not surprisingly, the distance turns out to be large in bigger villages, i.e. in villages with greater distance between houses of randomly selected villagers who constituted the sample

\(^{31}\)If the electoral system affects project implementation only through its effect on the education of council members, one can estimate the causal effect of council members education on the speed of project implementation using an IV specification in which electoral system is used as an instrument for the education of council members. The results of such an estimation are generally higher in magnitude than the result of the respective OLS estimation (see Table A2 in the Online Appendix for the results). These results, however, should be taken with extreme caution, as the exclusion restriction is unlikely to hold, since the electoral system can affect project implementation through other channels. For instance, greater difference in preferences of elected candidates in district elections, as predicted by the theoretical model, may cause a gridlock, which would delay the implementation of the projects.

\(^{32}\)We winsorize the distance at the 95 percent quintile to limit the influence of potential outliers.
Finally, according to the sixth empirical prediction we should observe a negative correlation between a candidate’s quality and the extent of his bias. We test this prediction by estimating the following model:

\[ Education_{vi} = \alpha + \varphi \cdot AL_{vi} + \rho \cdot AL_{vi} \cdot Distance_{vi} + \kappa \cdot Distance_{vi} + \phi q + \varepsilon_{vi} \quad (7) \]

where \(Distance_{vi}\) is the distance between the house of candidate \(i\) and the center of the village. The empirical prediction posits that \(\kappa < 0\) and \(\varphi + \rho = 0\).

Results indicate that in district elections there is indeed a negative, although not statistically significant, correlation between a candidate’s education and the distance between his house and the center of the village (see Table 7). There is, however, a statistically significant difference in this correlation between villages with at-large and district elections, with the correlation being positive, rather than negative in at-large elections, although it is statistically significant only if we don’t control for village size.\(^{33}\) Thus, we find limited support for the sixth empirical prediction.

7 Discussion

According to our theoretical model, electoral rules affect the quality of elected officials by changing voters’ incentives to support candidates with more biased preferences at the expense of their quality. The results from the field experiment prove to be fully consistent with the predictions of the theoretical model. In particular, we find that the quality of elected candidates is higher in at-large elections and that this difference is higher in more heterogenous villages. There is also evidence that in more heterogenous villages, district elections lead to the election of candidates with more biased preferences at the expense of their quality.

There are several possible alternative explanations for the positive effect of at-large elections on the quality of candidates. First of all, this effect can be driven by the restriction on the

\(^{33}\) We also test and find that for the random sample of villagers in the baseline survey there is no significant relationship between their education and the distance from their houses to the center of a village.
residence of candidates in the district elections. If there are two high-quality candidates that live in the same district, only one of them can be elected under district elections, but both of them can be elected in at-large elections. This restriction can have a negative effect on the quality of elected candidates in the situation in which candidates of good quality are scarce. To see if the difference in the quality of the elected council members is driven by this restriction, we look at the distribution of elected council members across districts. Villages with district elections had exactly one male and one female candidate elected to the council from each district. Although in at-large elections there were no formal restrictions on the distribution of candidates across districts, the distribution turned out to be not much different from that in villages with district elections. Specifically, in villages with at-large elections in 93 percent of districts there was either a male or a female candidate in the council that lived in the district and 73 percent of districts had both male and female residents as council members. There were only 37 out of 125 at-large villages where not all districts had a resident council member. Of these, in 25 villages there was only one district that did not have a resident in the council. Thus, the negative effect of at-large elections on the probability of a district having a resident council member was very small.

To further address this point, we exclude from the sample candidates in at-large elections from districts that had more than one candidate elected to the council. Thus, we look only at the quality of candidates for whom the restriction on the number of candidates from the same district was not binding. Although this restriction is endogenous and these results should be interpreted with caution, it provides some evidence regarding the robustness of our results to this alternative explanation. The results in Table A3 in the Online Appendix indicate that the effects obtained in the benchmark specification are robust to such a sample restriction. Finally, the restriction on candidates' residence in district elections should be more important in smaller villages. However, empirical results indicate that the effect of electoral rules on the quality of candidates is stronger in larger, rather than smaller, villages. Overall, the empirical results suggest that the effect of electoral rules on the quality of elected candidates is not driven by the restriction on candidates in district elections.

Another potential explanation is that an increase in district magnitude in at-large elections
will make it harder for the incumbent to coordinate voting, which will reduce the incumbency advantage of the members of the pre-existing elites and increase the quality of candidates. To examine the potential role of incumbency, we look at the share of council members who were identified as pre-existing elite members. Specifically, as the results in Table A4 in the Online Appendix indicate, the share of council members that were members of the preexisting elite, even if we use the most inclusive definition of pre-existing elites, is 39 percent. While sizable, this result suggests that pre-existing elites do not dominate the elected council. Importantly, electoral rules have almost no effect on the share of pre-existing elites among council members. The only marginally significant result indicates that the share of council members who were named as the main decision makers is somewhat higher in villages with at-large elections. We obtain similar results if we use an alternative measure of incumbency advantage and look at the proportion of pre-existing elite members who subsequently were elected to the council. The share of elite members elected to the council varies between 19 and 44 percent depending on the measure, but again there is no significant difference between villages that used alternative electoral rules. Overall, results indicate that electoral rules have no significant effect on incumbency advantage.

Another potential explanation is that formal education serves as a proxy for candidate quality, if voters do not have exact information on his quality. In district elections the size of the districts is smaller and voters are likely to have better information about the candidates, so they do not need to rely on formal education as a proxy for candidate quality, whereas larger district size in at-large elections leads to less information about actual candidate quality and more reliance on formal education as a proxy for their quality. This interpretation, however, cannot explain why the results are stronger in more heterogeneous villages and why the quality of elected candidates is on average lower in more heterogeneous villages. In more heterogeneous villages, voters are likely to have less information about the candidates, so according to this explanation voters should rely more on formal education as a proxy for quality. This would lead to higher levels of education of elected candidates in more heterogeneous villages, which contradicts our empirical findings. In addition, this explanation cannot account for the results on the location of the homes of elected candidates.

In general, the difference in the quality of elected representatives can also reflect barriers
to entry, which are higher in electoral systems with small district magnitudes (Myerson, 1993). However, this explanation is not relevant in our context, as all citizens are considered candidates and there are no entry barriers.

The observed difference in the quality of politicians may also be driven by the fact that the at-large elections format gives a disproportionate advantage to higher quality candidates for reasons other than voting decisions by citizens. Imagine, for example, that candidates get support through rallies (public speeches) and bribing (vote-buying) during the electoral campaign. Arguably, in bigger districts, speaking is more important due to economies of scale. As long as high quality candidates have a comparable advantage at public speaking, this would give them an advantage in electoral systems with larger districts. This explanation relies on a strong assumption that higher quality candidates are relatively better at speaking and not at bribing, which may or may not be true. Moreover, this explanation, while attractive theoretically, cannot explain the results of the field experiment, because political campaigning was forbidden, and based on our monitoring data vote-buying did not take place.

Higher quality of candidates elected under at-large electoral rules may not only lead to better implementation of development projects, but also improve villagers’ attitudes toward elected local leaders, making them more likely to support elections as the preferable selection method of local leaders. The results presented in Appendix Table A5 provide some evidence that at-large elections have a positive effect on villagers’ attitudes toward local leaders and elections as a method of selection of the village head. These results are broadly consistent with the model’s prediction that at-large elections lead to higher social welfare.

8 Conclusion

In this paper, we examine the effect of electoral rules on the levels of competence of elected representatives. We consider two alternative electoral rules – district and at-large elections. We provide a theoretical model in the tradition of citizen-candidate models with free entry of candidates (Besley and Coate, 1998). In district elections, citizens elect one legislator in each district separately. In at-large elections, citizens elect all the legislators jointly. The legislators
then make a joint policy decision. In each district, the local median voter anticipates the legislative bargaining process with the representatives of other districts, which makes him more concerned about policy relative to competence and leads to the election of more biased and less competent politicians. The model predicts that the quality of politicians will be higher in at-large elections and that this difference will be stronger in more heterogeneous communities.

We exploit the results of a field experiment conducted in 250 Afghan villages to test the empirical predictions of the model. Each of the villages was randomly assigned to elect a village council through one of two electoral rules – district or at-large elections. The results indicate that the quality of elected candidates, as measured by their educational attainment, is higher in villages with at-large elections. The difference is greater in more heterogeneous villages, as measured by the fractionalization of preferences over the type of development projects, geographic size and ethnic composition of these villages, so that in homogenous villages there is no significant effect of the electoral system on the quality of elected officials. Overall, the empirical results are consistent with the theoretical predictions of the model, and are not consistent with a number of alternative explanations for the effect of electoral rules on the quality of elected representatives.

The setting of our experiment allows us to isolate a specific mechanism through which the electoral system affects the quality and bias of elected candidates – strategic voting with a fixed set of candidates and no political campaigning. In most other contexts, however, electoral systems will also affect the outcomes through their effect on the entry of candidates, political campaigning, party involvement, and primary elections among others. Taking into account the role of these mechanisms, would be a fruitful avenue for future work.
References


Besley, Timothy, and Marta Reynal-Querol (2011) “Do democracies select more educated leaders?” American Political Science Review, 105(3).


Ferraz, Claudio and Frederico Finan (2011) “Motivating Politicians: The Impacts of Monetary
Incentives on Quality and Performance.” Working paper.


Online Appendix

Appendix A: Robustness

The model above is simple and makes clear predictions. In Appendix A, we show that these predictions are not due to excess simplification of the environment and that our results are robust.

Several competent individuals

The results of the paper are driven by scarcity of competent individuals; if for any policy position it were possible to find a competent citizen with such preferences, there would be no trade-off between policy and competence. Yet the assumption that there is only one competent individual may seem somewhat extreme. The truth is, it simplifies exposition considerably, but is not critical.

To show this claim formally, assume that the society includes $N$ competent citizens and, as before, needs to elect two council members. Formally, assume that citizens with ideal points $q_1, \ldots, q_N$ are competent, where $q_1, \ldots, q_N$ are independent random variables distributed uniformly on $[-B, B]$ (as usual, we will denote the order statistics by $q(1) \leq \cdots \leq q(N)$). As before, assume everyone knows who is competent and who is not. The case $N = 1$ was considered in Section 3.

We start by showing that for any $N$ and any realization of $q_1, \ldots, q_N$, there exists an equilibrium in pure strategies, both in district and in at-large elections. Median voter theorem applies again, and for at-large elections, a pair of citizens that maximizes the utility of the median voter, $w_{mx}(a_l, b_l, a_r, b_r)$, may be elected in an equilibrium. Notice that the median voter only needs to consider $N (N - 1) / 2$ pairs of competent citizens plus a combination of one competent citizen with type, say, $(h, q_1)$ and his political antipode $(0, -q_1)$; since he only needs to choose among a finite number of pairs, the maximum is attained at some pair.

The argument is only slightly more involved in the case of district elections. Suppose that in some pure strategy equilibrium the left district $L$ elects a citizen $(a_l, b_l)$. The best response by
the right district’s median voter is either to elect the most extreme of the competent individuals \((h, q_N(N))\), provided that there is a competent individual in the district \((q_N(N) \geq 0)\), or to elect the most extreme individual \((0, B)\); this only depends on \(b_l\). Thus, the political preferences of the best-response individual is \(BR_R(b_l) \subset \{q_N(N), B\}\). Moreover, this best-response function is monotone: if \(b_l' < b_l\) and \(B \in BR_R(b_l)\), then \(B \in BR_R(b_l')\), and if \(q_N(N) \in BR_R(b_l)\), then \(q_N(N) \in BR_R(b_l)\). Similarly, if the right district elects a citizen \((a_r, b_r)\), the political preferences of the best-response individual in the left district \(L\) is \(BR_L(b_r) \in \{-B, q(1)\}\). It also satisfies monotonicity: if \(b_r' > b_r\) and \(-B \in BR_L(b_r)\), then \(-B \in BR_L(b_r')\), and if \(q(1) \in BR_L(b_r)\), then \(q(1) \in BR_L(b_r)\). This monotonicity of best responses already implies existence. Obviously, if \(B \in BR_R(-B)\) and \(-B \in BR_R(B)\), then there is an equilibrium where \((a_l, b_l) = (0, -B)\) and \((a_r, b_r) = (0, B)\) are elected. If the first inclusion fails, then \(BR_R(b_l) = q_N(N)\) for any \(b_l\), and thus there is an equilibrium where \(R\) elects individual with type \((h, q_N(N))\) and \(L\) elects \((a_l, b_l)\), where \(b_l \in BR_L(q_N(N))\). Similarly, if the second inclusion fails, then there is an equilibrium where \(L\) elects \((h, q(1))\) and \(R\) elects \((a_r, b_r)\) with \(b_r \in BR_R(q(1))\). In any case, there is an equilibrium in pure strategies. The argument above applies, with obvious modifications, to \(N = 0\) as well.

Notice, however, that it in the case of district elections, the equilibrium need not be unique (even in terms of elected types). For example, take \(N = 2\), \(B = 1\), \(k = 1\), \(h = \frac{1}{4}\), and suppose \(q_1 = -\frac{1}{2}\), \(q_2 = \frac{1}{2}\). Then there is an equilibrium where \((h, q_1)\) and \((h, q_2)\) are elected: indeed, the median voter in district \(R\) gets \(\frac{1}{4} + \frac{1}{4} - \left(\frac{1+(-\frac{1}{2})}{2} - \frac{1}{2}\right)^2 = \frac{1}{4}\) by electing the competent citizen, but only \(\frac{1}{4} - \left(\frac{1-\frac{1}{2}}{2} - \frac{1}{2}\right)^2 = \frac{3}{16}\) by electing the extreme one, and thus does not want to deviate (and the calculation for district \(L\) is symmetric). At the same time, there is an equilibrium where \((0, -B)\) and \((0, B)\) are elected: in this case, the median voter in district \(R\) gets \(-\left(\frac{1+(-1)}{2} - \frac{1}{2}\right)^2 = -\frac{1}{4}\) by electing the most extreme one, but only \(\frac{1}{4} - \left(\frac{3+(-1)}{2} - \frac{1}{2}\right)^2 = -\frac{5}{16}\) by electing the competent one. This multiplicity of equilibria is due to strategic complementarity: the median voter in either district is more willing to elect an extreme council member if the other district elects an extreme one.

We thus have the following result.

**Proposition 5** Suppose that there are \(N\) competent individuals, where \(N\) is a non-negative
integer. Then for any realization of their political preferences there exists an equilibrium.

To proceed further, we need the following technical lemma.

**Lemma 1** Suppose that \(N \geq 2\) random variables \(q_1, \ldots, q_N\) are independent and uniformly distributed on \([-1, 1]\). Fix any real number \(z \in (0, 1)\). Let

\[
\begin{align*}
P(N, z) &= \Pr(\exists j \in \{1, \ldots, N\} : 1 - q_j \leq z), \\
Q(N, z) &= \Pr(\exists i, j \in \{1, \ldots, N\}, i \neq j : |q_i + q_j| \leq z).
\end{align*}
\]

Then \(P(N, z)\) and \(Q(N, z)\) are strictly increasing in \(z\) and in \(N\), and \(P(N, z) \leq Q(N, z)\) for all \(N\) and \(z\).

**Proof.** The proof of this result is in Appendix B, along with other proofs.

In what follows, assume that \(4h < kB^2\); this assumption means that the political dimension is sufficiently important. It says that any citizen prefers his ideal point implemented by an incompetent council member to a point at distance \(B/2\) implemented by a competent one. The assumption guarantees that in within-district elections, a competent citizen with \(b_i\) close to 0 will not be elected, so there is a real competence-vs.-bias trade-off in district elections.

Consider at-large elections. The median voter can always guarantee himself utility \(h\), by electing any competent citizen \((h, q_1)\) and an incompetent citizen \((0, -q_1)\). However, he could do better if there were two competent citizens \(q_i\) and \(q_j\) with \(k \left(\frac{q_i + q_j}{2}\right)^2 < h\); in this case, he would get \(2h - k \left(\frac{q_i + q_j}{2}\right)^2\). Therefore, in at-large elections, if there are two competent citizens with \(|q_i + q_j| \leq 2\sqrt{\frac{h}{k}}\), then the council will consist of two competent citizens, and otherwise will contain one competent and one incompetent one. By Lemma 1 (which we may apply with an appropriate normalization), the probability that both members are competent equals \(Q(N, z)\), where we denoted \(z = \frac{2}{B} \sqrt{\frac{h}{k}} < 1\).

Now take district elections. Two competent citizens will be elected only if both districts elect competent citizens. Suppose district \(L\) elected a citizen with political position \(b_i\); then the median voter in district \(R\) will elect a competent citizen only if there is one with political position \(q_j\) such that \(h - k \left(\frac{b_i + q_j}{2} - \frac{B}{2}\right)^2 \geq -k \left(\frac{b_i + B}{2} - \frac{B}{2}\right)^2\), i.e., if \(q_j \geq B - b_i - \sqrt{4\frac{h}{k} + b_i^2}\).
This is equivalent to $1 - \frac{q_t}{B} \leq \frac{b_t}{B} + \sqrt{4 \frac{h}{kB} + \left( \frac{b_t}{B} \right)^2}$; therefore, by Lemma 1, for any given $b_l$, the probability that district $R$ elects a competent citizen is $P(N, z(b_l))$, where $z(b_l) = \frac{b_l}{B} + \sqrt{4 \frac{h}{kB} + \left( \frac{b_l}{B} \right)^2}$. Notice that $z(b_l) \leq z$ for all $b_l \in [-B, 0]$, and the inequality is strict for $b_l \neq 0$.

From the reasoning above, in district elections, district $R$ would only elect council members with $b_r \geq B - b_l - \sqrt{4 \frac{h}{k} + b_l^2} > 0$, and, similarly, district $L$ would only elect those with $b_l \leq -B - b_r + \sqrt{4 \frac{h}{k} + b_r^2} < 0$. Therefore, holding any council member from the left district who may be elected fixed, we have $z(b_l) < z$, and thus $P(N, z(b_l)) < P(N, z) \leq Q(N, z)$, which means that the probability that district $R$ elects a competent politician is strictly less likely than the probability that two competent politicians are elected in at-large elections. Consequently, the probability that both council member are competent is strictly smaller in district elections than in at-large ones. This establishes the following result.

Proposition 6 For any number of competent citizens $N \geq 1$, the expected quality of council members under at-large elections is higher than under district elections. Moreover, the number of competent council member under at-large elections first-order stochastically dominates that under district elections.

It is trivial to extend the result to the case where $N$ is random (say, a Poisson variable); in this case, the villagers would observe the identities of competent citizens, and thus $N$, prior to voting, and then the reasoning above for this $N$ applies.

8.1 Legislative bargaining game

In this subsection, we modify the game from Section 3 by assuming that the two elected council members do not automatically choose the policy midway between their ideal points, but rather participate in a Baron and Ferejohn (1989)-style legislative bargaining game. Namely, the two council members, $l$ with type $(a_l, b_l)$ and $r$ with type $(a_r, b_r)$ (where $b_l < b_r$) play the following game.

There are an infinite number of periods, starting with period 0. In each period, each of the council members becomes agenda setter with probability $\frac{1}{2}$. The agenda-setter proposes policy
\( p \), and the other member either accepts or rejects it. If \( p \) is accepted in period \( t \), then each citizen \( i \) (including the two council members) get \( u(p, b_i) = -k(p - b_i)^2 \) in each subsequent period. In every period before a policy is accepted, all citizens suffer a penalty \(-P\), where \( P > 4kB^2 \) (since the payoff from policy \( u(p, b_i) \) is non-positive, we need to assume that the payoff without any policy is even worse, even if the distance to that policy is \( 2B \)). All citizens maximize their discounted expected payoff, and \( \beta \in (0, 1) \) is a common discount factor.

We first solve for the outcome of the bargaining game. It is characterized by an acceptance set \( A \subset X \), which is a connected compact, and each of the council members, when he becomes an agenda-setter, picks the policy from set \( A \) which maximizes his \( u(p, b_i) \) over \( p \in A \). The immediate acceptance result applies; along the equilibrium path, the first policy proposed will be accepted. We can easily prove the following result.

**Lemma 2** Suppose that two council members, \( l \) and \( r \), have ideal points \((a_l, b_l)\) and \((a_r, b_r)\) with \( b_l < b_r \). Then in equilibrium:

(i) If \( \beta \geq \frac{P - k(b_r - b_l)^2}{p - k(b_r - b_l)^2} \), then \( l \) and \( r \) propose \( \frac{b_l + b_r}{2} \) and \( \frac{b_l + b_r}{2} + \left( \frac{P}{2\beta} + \frac{(b_r - b_l)^2}{2(1 - \beta)} \right)^2 \beta (2 - \beta) - \frac{b_r - b_l}{2(1 - \beta)} \), respectively, and other council member is indifferent between accepting and rejecting these proposals.

(ii) If \( \beta < \frac{P - k(b_r - b_l)^2}{p - k(b_r - b_l)^2} \), then \( l \) and \( r \) propose their ideal points, \( b_l \) and \( b_r \), and the other council member strictly prefers to accept it.

Lemma 2 says the following. If the discount factor \( \beta \) is sufficiently high, then the acceptance set \( A \) is sufficiently narrow; it lies strictly between the ideal positions of the two council members, and each agenda-setter proposes the policy at the extreme of the acceptance set. If the discount factor is sufficiently low, then the acceptance set is wide, as the politicians are too impatient and are willing to accept policies that are far from their ideal point. This allows each politician to insist on their ideal policy in equilibrium. It is easy to see that if punishment \( P \) is very high, then the acceptance set is likely to be large, and politicians will propose their ideal policy.

As in many bargaining models, the extreme case where \( \beta \) is close to 1, i.e., politicians are either patient or are able to make proposals frequently, is the most interesting one. However, the
opposite case where $\beta$ is close to 0 is also noteworthy. The following characterizes comparative statics in these extreme cases.

**Proposition 7** There exist $0 < \beta_1 < \beta_2 < 1$ such that:

(i) If $\beta > \beta_2$, then the expected competence of council members elected in at-large elections is higher than that in district elections. Moreover, as $\beta \to 1$, the types of elected council members converges (in distribution) to the case where they chose the midpoint automatically, as in Section 3;

(ii) If $\beta < \beta_1$, then the expected competence in district elections is higher than the expected competence in at-large elections. Moreover, when bargaining, each council member proposes his own ideal point.

Proposition 7 gives two important takeaways. First, if offers are made frequently and $\beta$ is close to 1, the outcomes of elections are similar to the outcomes of the game studies in Section 3, and this implies robustness of those results. Second, if offers are made rarely, the results are overturned, and district elections lead to more competent council members. This goes in contrast to the previous results; to see the intuition, it is helpful to observe that if $\beta$ is low enough, then each council member will propose his ideal point. This creates very different incentives to voters in district elections: instead of electing a very biased council member in hope that his influence would moderate the council member from the other district, the median voter in a district would prefer to elect someone with ideal point close to him. Indeed, this median voter has no hope of influencing the offer made by the council member elected by the other district, and instead he wants to get higher utility from offers made by his own delegate. As a result, a district which lacks a competent individual elects his median voter to the council, whereas a district with the competent person elects him if he is close enough to the median voter, i.e., if $k (q - \frac{B}{2})^2 \leq h$ in district $R$ and if $k (q + \frac{B}{2})^2 \leq h$ in district $L$, and otherwise it elects the median voter in that district. The incentives are also changed in at-large elections. Now, the median voter prefers to elect one council member with $b_i = 0$, and also the competent person, provided that $kq^2 \leq h$. Thus, to get elected, the competent person needs to be within $\sqrt{\frac{h}{k}}$ distance from 0 in at-large elections, and within such distance from either $-\frac{B}{2}$ or $\frac{B}{2}$ in district elections, and the second is...
clearly more likely.

We therefore see that district elections dominate at-large elections if offers are sufficiently infrequent, and the reason is that the size of each district is smaller, and therefore even relatively extreme citizens in the district are not so extreme from the perception of the district’s median voter. Thus, the effect that at-large elections produce more competent council members (which we see in the data) is due to legislative bargaining considerations, rather than the ability of all voters to coordinate in at-large elections.

8.2 Multiple districts

We expand the previous subsection to the case of multiple districts. Suppose that in district elections, the village is divided into $M$ equally-sized contiguous districts, so for $j \in \{1, M\}$, district $D_j = [-B + \frac{2B}{M}(j - 1), -B + \frac{2B}{M}j]$, and each district needs to elect one council member. In at-large elections, the entire village elects $M$ council members (to keep the model similar to the previous case, it is natural to assume that each citizen has $M$ votes, but this is not consequential to the analysis as everything will be decided by the median voter in any case). As in Subsection 8.1, each council member is chosen randomly to make a proposal, and a proposal is accepted if sufficiently many council members support it. To be concrete, let us focus on simple majority rules (which generalizes Subsection 8.1): a proposal is accepted if more than $\frac{M}{2}$ council members support it. Furthermore, we consider the case where $\beta$ is sufficiently high; this makes the results comparable to those in the main part of the paper (Section 3).

It is easy to show that the median voter theorem applies in both at-large and district elections, and thus decisions will be made by median voters in respective districts. To understand their incentives, consider the outcome of bargaining between four council members with political preferences $b_1, \ldots, b_M$. It is not hard to show that if $\beta$ is close to 1, then the acceptance set converges to a point (see Austen-Smith and Banks [[[REFERENCE]]]). Moreover, this point coincides with the preferences of the median council member $b_{\frac{M+1}{2}}$ if $M$ is odd and it lies halfway between the two median council members (i.e., $\frac{1}{2} \left( b_{\frac{M}{2}} + b_{\frac{M}{2}+1} \right)$) if $M$ is even. To see why,

Given this characterization, it is easy to understand the election incentives. In at-large elections, the median voter $m_0$ would always elect the competent politician, whereas the rest
will be at or very close to his ideal point. In other words, the result that the most competent politician is elected will hold, provided that $\beta$ is close to 1. In district elections, the incentives are different. Suppose $M$ is even; then the median voters in two median districts $D_{M/2}$ and $D_{(M/2)+1}$ have incentives to elect more extreme council members because this will sway bargaining in the direction that they prefer. At the same time, the other districts do not have this strategic incentive; their choice will not alter the acceptance set or the policy proposed by their delegate. Consequently, they will elect the competent citizen if he happens to reside there. Thus, a competent person is elected, unless he lives too close to the center, more precisely, he is elected if $|q| \geq \frac{2B}{M} - 2\sqrt{\frac{b}{k}}$. If $M$ is odd, then if the competent person lives in the central district $D_{(M+1)/2}$, then he is elected if and only if $q \leq \sqrt{\frac{h}{k}}$, while if he lives in any other district, he is elected for sure. Thus, a competent person is elected if $q \notin \left( \frac{B}{M} - \sqrt{\frac{h}{k}}, \frac{B}{M} \right)
$.  

\begin{align*}
\frac{h}{k} &< \left( \frac{x - 2B}{2} - \frac{B}{M} \right)^2 \\
\frac{h}{k} &< \left( \frac{2}{3} - \frac{2B}{M} \right)^2
\end{align*}

**Election one legislator at a time**

In the main model in Section 3, at-large elections led to more competent council members partly because the voters were able to perfectly balance the competent individual they wanted to elect with someone who has exactly the opposite policy preferences. In Subsection 8.1, we showed that this result disappears if both council members are elected at the same time, but instead of working out a joint decision, each of them is able to choose his ideal policy with equal probability. This suggested that the results are driven by joint policy decisions rather than coordination.

In this Subsection, we emphasize this further by showing that if the two council members are elected sequentially, then our result of Section 3 go through. (For example, the U.S. Senate is elected this way: each state elects two senators, but only one at a time.) More precisely, we take one council member as given, and his type is $(a_0, b_0)$. Without loss of generality, assume that $b_0 < 0$, and consider two possibilities: in at-large elections, the whole society votes for the other member, and in district elections, only district $R$ votes.
his Our goal here is to show that the result of the paper is not driven by the simultaneous elections of both legislators in at-large elections.

More precisely, let us assume that only one legislator is elected in the game, and the other one is inherited from the previous period. (This resembles the way the U.S. Senate is elected.) In at-large elections, the old legislator has type \((a_0, b_0)\), and the society elects another legislator, \((a_n, b_n)\) from the same pool of candidates as in the main model (i.e., there is only one competent individual). In district elections, the old legislator with type \((a_0, b_0)\) belongs to one of the districts, and the new legislator \((a_n, b_n)\) is elected in the other district from the restricted pool of candidates. Without loss of generality, assume \(b_0 < 0\), so the district elections take place in the right district.

We can again prove that the single-crossing conditions hold, so elections are determined by the median voter in the corresponding elections. Let us again fix the competence of the competent individual at \(b\). Consider at-large elections first. The median voter is effectively choosing between mirroring the old legislator (thus electing someone with type \((0, -b_0)\) and getting utility \(a_0\)) and electing the competent legislator, thus getting utility \(a_0 + h - k \left( \frac{b_0 + b}{2} \right)^2\).

He will choose the competent legislator if and only if \((b_0 + b)^2 \leq 4h/k\), i.e., if \(b\) is in \(2\sqrt{h/k}\)-neighborhood of \(-b_0\).

In district elections, the median voter is choosing between the biased candidate (which will give him utility \(a_0 - k \left( \frac{b_0 + B}{2} - \frac{B}{2} \right)^2 = a_0 - k \left( \frac{b + b}{2} \right)^2\) and the competent one (which will give him utility \(a_0 + h - k \left( \frac{b_0 + b}{2} - \frac{B}{2} \right)^2\); this is only possible, of course, if \(b \geq 0\) and thus the competent individual is available). The competent candidate is elected if and only if \(4h/k + (b_0)^2 \geq (b_0 - B + b)^2\), i.e., if \(b\) is in the \(\sqrt{4h/k + (b_0)^2}\)-neighborhood of \(B - b_0\). Since \(b_0 < 0\), this is true for \(b \in \left[ B + |b_0| - \sqrt{4h/k + (b_0)^2}, B \right]\); the length of this interval is less than \(2\sqrt{h/k}\). It is now clear that in expectation (taken over the value of \(b_0\)), at-large elections are still more likely to elect more competent candidate; one can also prove that the result for polarization holds as well.

The intuition for this result is the following. In at-large elections, the induced ideal point for the median voter for the new legislator is \(-b_0\), while in district elections, this point is \(B - b_0\). Thus, in the former case, the ideal point is strictly in the interval of \([0, B]\), and in the larger
case it is beyond this interval. This immediately leads to polarization, but given the quadratic disutility function, the voters are also more sensitive to policy in the latter case, and thus they are more willing to elect an incompetent individual. As a result, even when only one politician were to be elected, at-large elections would produce superior results (e.g., in the sense of Proposition 4). It is worth noting that this would be true even if in at-large elections, citizens had to elect someone from the right district (thus potentially restricting their ability to elect the most competent candidate).

Other decision-making in legislature

So far, we have assumed that the two legislators make a joint policy decision, and in doing so, they bargain efficiently. This seems to be a reasonable approximation to the environment we are interested in. One could, however, consider different models of decision-making in legislatures.

Suppose, for example, that the legislative body makes decisions on a number of questions, and only share $\alpha$ requires a joint decision, while for $1 - \alpha$, a random legislator is appointed to make a unilateral decision. The previous case corresponds to $\alpha = 1$, while $\alpha < 1$ may correspond to situations where some policy decisions are local, and the local legislator has the sole responsibility of making the decision.

It turns out that our results stay for $\alpha$ sufficiently high, but as $\alpha$ becomes smaller, district elections become preferred. To see why, consider the extreme, $\alpha = 0$, and notice that in this case the median voter in district elections does not have a strategic reason for voting for biased candidates. His ideal candidate has the same ideal point as he does ($-B/2$ or $B/2$), and moreover, the problems of the two districts are independent. Now, the reason why district elections would lead to more competent candidates is clear: the median voter in the district is not too averse to any of the candidates in this district; for example, if $h > k (B/2)^2$, the most competent candidate is guaranteed to be elected. In at-large elections, the median voter (at 0) would be quite a bit averse to competent but biased candidates; in this case, we can only guarantee that the competent citizen will be elected if $h > kB^2$, which is a stronger condition.

This comparison tells us that what is really driving the results are strategic considerations of voters when they anticipate joint decision making in the legislature. It is worth noting that
efficiency of bargaining also contributes to this result. For example, suppose that the decision in the legislature was made through a Baron-Ferejohn (or Rubinstein) legislative bargaining game, with a randomly chosen first proposer and discount factor $\beta$ between proposals. As $\beta$ approaches 1, the equilibrium decisions converge to the average of the legislators’ ideal points, which leaves us with the model that we analyzed in Section ??B. But if $\beta$ is close to 0, the first proposer becomes able to enact his ideal policy, and the payoffs are similar to the case $\alpha = 0$ studied earlier. This means that for $\beta$ sufficiently low, district elections are likely to be preferred to at-large elections.

Together, these considerations deliver an important take-away. At-large elections are preferred if legislators make a joint decision. If they have multiple policy questions which they split between themselves, or even if they don’t split, the number of decisions to make is so high that it takes a long while to return back to the question that was left undecided, then district elections should have an edge. Studying such trade-offs in more detail seems to be a fruitful area for future research.

Appendix B: Proofs

Proof of Proposition 1. Part 1. Let us show that the following increasing differences property holds. In district elections, for any distribution of types $(a_l, b_l)$ elected by district $L$, we have that for two citizens $i, j$ with $b_i > b_j$, and any candidates $(a_r, b_r), (a_r', b_r')$ such that $b_r > b_r'$,

$$
\mathbb{E}w_i (a_l, b_l, a_r, b_r) - \mathbb{E}w_i (a_l, b_l, a_r', b_r') > \mathbb{E}w_j (a_l, b_l, a_r, b_r) - \mathbb{E}w_j (a_l, b_l, a_r', b_r'),
$$

where the expectation is taken over the distribution of $(a_l, b_l)$. Indeed, we have

$$
\mathbb{E}w_i (a_l, b_l, a_r, b_r) - \mathbb{E}w_i (a_l, b_l, a_r', b_r') = \mathbb{E}a_l + a_r - \mathbb{E}k\left(\frac{b_l + b_r}{2} - b_l\right)^2 - \mathbb{E}a_l - a_r' + \mathbb{E}k\left(\frac{b_l + b_r'}{2} - b_l\right)^2 = (a_r - a_r') + k\left(\frac{b_r - b_r'}{2}\right)\left(2b_i - \mathbb{E}b_l - \frac{b_r + b_r'}{2}\right),
$$

46
which is again increasing in \( b_i \). Obviously, a similar increasing differences condition holds for elections in district \( L \), holding the distribution is district \( R \) fixed.

Suppose that \( \sigma \) is an equilibrium in district elections. Take district \( L \) and consider the set of types \( Z \) that maximize the payoff of median voter \( m_L \), holding the strategies of voters in district \( R \) fixed (this set is nonempty, since the space of types is compact: it is a segment \( \{ a, b : a = 0, b \leq 0 \} \), plus perhaps a point \( (h, q) \), if \( q \leq 0 \). Let us show that district \( L \) must elect a council member from set \( Z \) with probability 1. Suppose not, i.e., there is a probability distribution over the elected types \( (a_l, b_l) \), and there is a positive probability that some type \( (a, b) \notin Z \) is elected. Take \( (a', b') \in Z \) and let us show that there is a coalition that is able and willing to deviate and elect \( (a', b') \). Indeed, we have that the median voter \( m_L \) prefers \( (a', b') \) over the distribution of types in \( \sigma \). Then if \( b' > E b_t \), then all individuals with \( b_t \geq -\frac{B}{2} \) prefer \( (a', b') \) because of increasing differences, and some of those with \( b_t < -\frac{B}{2} \) prefer \( (a', b') \) by continuity, and thus there is a majority which can elect \( (a', b') \) and profit from it. A similar argument applies if \( b' < E b_t \), whereas if \( b' = E b_t \), then all citizens of district \( L \) strictly prefer \( (a', b') \), and thus there is a profitable deviation. This shows that only types that maximize the utility of the median voter may get elected, and a similar argument applies to district \( R \).

Consider the expected utility of the median voter in district \( L \) if type \( (a_l, b_l) \) is elected. It is given by

\[
\mathbb{E} w_{m_L} (a_l, b_l, a_r, b_r) = a_l + \mathbb{E} a_r - \mathbb{E} k \left( \frac{b_l + b_r}{2} + \frac{B}{2} \right)^2
\]

\[
= a_l + \mathbb{E} a_r - k \left( \frac{b_l + \mathbb{E} b_r}{2} + \frac{B}{2} \right)^2 - \frac{k}{4} \text{Var} (b_r),
\]

and is monotonically decreasing in \( b_l \). Thus, the only possible types that can maximize the utility of \( m_L \) are \((0, -B)\) or \((h, q)\), provided that \( q \leq 0 \). Similar considerations apply to district \( R \), which proves that the district without the competent citizen elects the most biased individual, and the district with it electses either of the two. Moreover, the median voter in a district with the competent citizen (say, district \( L \)) is only indifferent between him and the biased voter if

\[
w_{m_L} (0, -B, 0, B) = w_{m_L} (h, q, 0, B),
\]

47
as in this case district $R$ elects the type $(0, B)$ as we just showed; this is equivalent to

$$-k \left( \frac{B}{2} \right)^2 = h - k \left( \frac{q + B}{2} + \frac{B}{2} \right)^2;$$

and this can hold for exactly one value of $q$, $q = -\hat{q}$. Similarly, the median voter in district $R$ may be indifferent only if $q = \hat{q}$. This proves that for almost all values of $q$ the types elected in equilibrium are uniquely determined.

It remains to prove that there exists an equilibrium. For $|q| \neq \hat{q}$, consider voting strategies where in every district, every voter votes for the candidate specified above. Then there is no profitable deviation by any coalition; any such coalition must gather support of at least half of voters in the the district and thus must make the median voter at least as well off; however, for these $q$, there is no such alternative. If $q = \hat{q}$, then there is an equilibrium where voters to the left $m_R$ in district $R$ vote for $(h, q)$ and the rest vote for $(0, B)$; each gets half of votes and wins with probability $\frac{1}{2}$; the strategy is similar if $q = -\hat{q}$. It is easy to show that in these cases, too, there is no profitable deviation by any coalition, and this finishes the proof of existence.

**Part 2.** Let us establish the following increasing differences property. In at-large elections, for two citizens $i, j$ with $b_i > b_j$, and any candidates $(a_l, b_l), (a_r, b_r), (a'_l, b'_l), (a'_r, b'_r)$ such that $\frac{b_l + b_r}{2} > \frac{b'_l + b'_r}{2}$,

$$w_i (a_l, b_l, a_r, b_r) - w_i (a'_l, b'_l, a'_r, b'_r) > w_j (a_l, b_l, a_r, b_r) - w_j (a'_l, b'_l, a'_r, b'_r).$$

To see this, consider

$$w_i (a_l, b_l, a_r, b_r) - w_i (a'_l, b'_l, a'_r, b'_r)$$

$$= a_l + a_r - k \left( \frac{b_l + b_r}{2} - b_i \right)^2 - a'_l - a'_r + k \left( \frac{b'_l + b'_r}{2} - b_i \right)^2$$

$$= (a_l + a_r - a'_l - a'_r) + k \left( \frac{b_l + b_r}{2} - \frac{b'_l + b'_r}{2} \right) \left( 2b_i - \frac{b_l + b_r}{2} - \frac{b'_l + b'_r}{2} \right),$$

which is increasing in $b_i$.

Let us show that there is an equilibrium where individuals with types $(h, q)$ and $(0, -q)$ are
elected. Fix the voting strategies where each citizen casts one vote for \((h, q)\) and another vote for \((0, -q)\), and show that there is no collective deviation that increases utility of all its members. Indeed, suppose that a subset of citizens \(X\) can deviate and get types \((a_l, b_l), (a_r, b_r)\) elected. If \(b_l + b_r = 0\) and not all citizens are indifferent, it must be that \(a_l = a_r = 0\), but in this case, all citizens are worse off, so \(X\) must be empty and cannot make any deviation. Thus, \(b_l + b_r \neq 0\), and without loss of generality suppose \(b_l + b_r < 0\). Then for median voter \(m_0\), \(w_{m_0} (h, q, 0, -q) > w_{m_0} (a_l, b_l, a_r, b_r)\), and by increasing differences, \(w_i (h, q, 0, -q) > w_i (a_l, b_l, a_r, b_r)\) for any \(i\) with \(b_i > 0\); continuity implies that the same inequality holds in the neighborhood of 0, if \(b_i > \frac{b_l + b_r}{4}\) (which is negative). Thus, the share of voters who strictly prefer \((a_l, b_l), (a_r, b_r)\) to \((h, q), (0, -q)\) is less than \(\frac{1}{2}\), and \(X\) is a subset of this set. Thus, after deviation, \((h, q)\) and \((0, -q)\) will share the votes of \(S \setminus X\), thereby each getting more than \(\frac{1}{4}\) of all votes. At the same time, any candidate supported by voters in \(X\) will get less than \(\frac{1}{4}\), even if all citizens in \(X\) give him one of their votes. This implies that coalition \(X\) is unable to alter the results of the elections, a contradiction that proves existence of an equilibria with the required properties.

Now, suppose that there is an equilibrium \(\sigma\) which induces some distribution over pairs of individuals \((a_l, b_l), (a_r, b_r)\) who get elected. Suppose first that \(E (b_l + b_r) = 0\). If the individual with \((h, q)\) is elected with probability 1, then individual with type \((0, -q)\) is also elected with probability 1, and thus \(\sigma\) is an equilibrium stipulated by the Proposition. If \((h, q)\) is not part of the pair with a positive probability, then \(E (a_l + a_r) < h\). In this case, the entire society \(S\) has a deviation, where each citizen casts votes for \((h, q)\) and \((0, -q)\); this will not change the expected policy, will not increase policy variance, but will increase the expected competence of the council. Now suppose that \(E (b_l + b_r) \neq 0\); without loss of generality, \(E (b_l + b_r) < 0\). Consider coalition \(X\) of citizens with \(b_i > \frac{E(b_l + b_r)}{4}\); each of them prefers policy 0 to policy \(\frac{E(b_l + b_r)}{2}\), and therefore each of them strictly prefers to have \((h, q)\) and \((0, -q)\) elected. They can also achieve this by voting for these individuals; in this way, they will get more than \(\frac{1}{4}\) votes each, whereas all other individuals will be left with less than \(\frac{1}{4}\) votes each. This is a profitable deviation, showing that only equilibria where \((h, q)\) and \((0, -q)\) are elected may exist. This completes the proof. □

Proof of Proposition 2. As shown in the proof of Proposition 1, district \(L\) elects the
competent citizen if \( w_{mL}(0, -B, 0, B) < w_{mL}(h, q, 0, B) \), i.e., if \( q < -\hat{q} \), and similarly, district \( R \) does so if \( q > \hat{q} \). Thus, two most biased individuals are elected in the complementary case, i.e., if \( |q| < \hat{q} \). This set is nonempty if \( \hat{q} > 0 \), which holds if and only if \( \frac{3}{4} B^2 > \frac{h}{2} \). When this is true, the probability that the competent citizen is elected is

\[
R = 1 - \frac{\hat{q} - (-\hat{q})}{2B} = 1 - \frac{\hat{q}}{B} = \sqrt{\frac{4h}{kB^2} + 1} - 1. \tag{8}
\]

Thus, \( R \) is increasing in \( h \) and decreasing in \( k \) and \( B \). This completes the proof. ■

**Proof of Proposition 3.** Part 1. In at-large elections, one council member is competent and the other is not, thus expected competence is \( C_a = \frac{h}{2} \). In district elections, the expected competence is \( C_d = R \frac{h}{2} \) (where \( R \) is given by (8)). Thus, \( C_a \geq C_d \) because \( P \leq 1 \), and the inequality is strict whenever \( R < 1 \), which may be simplified to \( 4h < 3B^2k \). The difference is \( C_a - C_d = (1 - P) \frac{h}{2} = \left( 2 - \sqrt{\frac{4h}{kB^2} + 1} \right) \frac{h}{2}, \) which is increasing in \( B \) and \( k \).

Part 2. In at-large elections, for a given \( q \), both council members lie at distance \( q \) from 0, and thus expected polarization equals \( P_a = \frac{1}{2} \int_0^B \frac{1}{B} q dq = \frac{1}{4} \). In district elections, it equals \( P_d = \frac{1}{2} \left( \int_0^{\hat{q}} \frac{1}{B} Bdq + \int_{\hat{q}}^B \frac{1}{B} \left( \frac{q + B}{2} \right) dq \right) = \frac{1}{4} \left( 3 - \frac{\hat{q}}{B} \right) \left( 1 + \frac{\hat{q}}{B} \right), \) provided that \( \hat{q} > 0 \), and equals \( P_d = \frac{3}{4} \) otherwise. Thus, \( P_a - P_d = \frac{1}{4} \left( 1 + 2 \frac{\hat{q}}{B} - \left( \frac{\hat{q}}{B} \right)^2 \right) > 0 \). In addition, \( P_a - P_d \) is increasing in \( \frac{\hat{q}}{B} = 2 - \sqrt{\frac{4h}{kB^2} + 1}, \) and thus is increasing in \( k \) and \( B \).

Part 3. In at-large elections, for a council member \((a, b)\), the \( \Pr \left( \frac{|b|}{B} < x \mid a = h \right) = \Pr \left( \frac{|b|}{B} < x \mid a = 0 \right) = x \) (for \( x \in [0, 1] \)). Thus, in elected council members, competence and bias are independent and thus uncorrelated. In district elections, if \( a = h \), the conditional distribution is uniform on \( \left[ \frac{\hat{q}}{B}, 1 \right] \), so \( \Pr \left( \frac{|b|}{B} < x \mid a = h \right) = \frac{x - \frac{\hat{q}}{B}}{\frac{\hat{q}}{B}} \) for \( x \in \left[ \frac{\hat{q}}{B}, 1 \right] \). At the same time, if \( a = 0 \), the conditional distribution is an atom at 1: \( \Pr \left( \frac{|b|}{B} = 1 \mid a = 0 \right) = 1 \). Hence,

\[
\mathbb{E} \left( \frac{|b|}{B} \mid a = h \right) = \frac{1}{2} \left( 1 + \frac{\hat{q}}{B} \right) < 1 = \mathbb{E} \left( \frac{|b|}{B} \mid a = h \right),
\]

because \( \hat{q} < B \). Consequently, in district elections, \( a \) and \( b \) are negatively correlated. This completes the proof. ■
Proof of Proposition 4. Consider the utility of a voter $i$ with ideal point $b_i$ if the location of the competent person is $q$. In case of at-large elections, it is equal to

$$ U_a (q, b_i) = w_i (h, 0; 0) = h - kb_i^2. $$

In case of district elections, it equals

$$ U_d (q, b_i) = \begin{cases} h - k \left( \frac{q + B}{2} - b_i \right)^2 & \text{if } q < -\hat{q} \\ -kb_i^2 & \text{if } |q| < \hat{q} \\ h - k \left( \frac{q - B}{2} - b_i \right)^2 & \text{if } q > \hat{q} \end{cases}. $$

Taking expectation over $q$, $E U_a (q, b_i) = h - kb_i^2$, and

$$ E U_d (q, b_i) = h \left( 1 - \frac{\hat{q}}{B} \right) - k \left( b^2 + \frac{1}{12} \left( 1 - \frac{\hat{q}}{B} \right)^3 \right). $$

Thus,

$$ E U_a (q, b_i) - U_d (q, b_i) = h \frac{\hat{q}}{B} + \frac{1}{12} k \left( 1 - \frac{\hat{q}}{B} \right)^3 > 0. $$

This completes the proof. ■

Proof of Lemma 1 (from Appendix A). The fact that $P(N, z)$ and $Q(N, z)$ are strictly increasing in both variables is trivial. Denote the c.d.f. of each of $q_j$ by $F(x)$; then $F(x) = \frac{x+1}{2}$ for $x \in [-1, 1]$. Let us first show that

$$ P(N, z) = 1 - \left( \frac{2 - z}{2} \right)^N. $$

Indeed,

$$ P(N, z) = \Pr (1 - q(N) \leq z) = \Pr (q(N) \geq 1 - z) = 1 - \Pr (q(N) \leq 1 - z) = 1 - F^N (1 - z) = 1 - \left( \frac{1 - z + 1}{2} \right)^N = 1 - \left( \frac{2 - z}{2} \right)^N. $$
We prove that \(Q(N, z) \geq P(N, z)\) (with equality only if \(N = 2\)) by induction by \(N\), separately for even and odd \(N\). We start with even \(N\).

Suppose \(N = 2\). Then

\[
Q(2, z) = \Pr(|q_1 + q_2| \leq z) = \Pr(-z \leq q_1 + q_2 \leq z) = 2 \Pr(0 \leq q_1 + q_2 \leq z),
\]

where the last equality follows from symmetry of distribution of \(q_1 + q_2\). The p.d.f. of the distribution of \(q_1 + q_2\) is \(\frac{2-|x|}{4}\) for \(|x| \leq 2\), and thus

\[
Q(2, z) = 2 \Pr(0 \leq q_1 + q_2 \leq z) = 2 \int_0^z \frac{2-x}{4} \, dx = \frac{z(4-z)}{4}.
\]

Now take \(N \geq 4\)

\[
Q(N, z) = \Pr(\exists i, j \in \{1, \ldots, N\}, i \neq j : |q_i + q_j| \leq z) \\
> \Pr(|q_1 + q_2| \leq z \vee \cdots \vee |q_{N-1} + q_N| \leq z) \\
= 1 - \Pr(|q_1 + q_2| \geq z \wedge \cdots \wedge |q_{N-1} + q_N| \geq z) \\
= 1 - \Pr(|q_1 + q_2| \geq z) \times \cdots \times \Pr(|q_{N-1} + q_N| \geq z) \\
= 1 - (1 - Q(2, z))^N = 1 - \left(\frac{2-z}{2}\right)^N = P(N, z),
\]

which proves the result for even \(N\).
Consider the case of odd $N$. Suppose $N = 3$. Then we have
\[
Q(N, z) = \Pr (|q_1 + q_2| \leq z \vee |q_1 + q_3| \leq z \vee |q_2 + q_3| \leq z) > \Pr (|q_1 + q_2| \leq z \vee |q_1 + q_3| \leq z) = 1 - \Pr (|q_1 + q_2| \geq z, |q_1 + q_3| \geq z) = 1 - \Pr (q_1 + q_2 \leq -z \vee q_1 + q_2 \geq z, |q_1 + q_3| \geq z) = 1 - \Pr (|q_1 + q_3| \geq z) \Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z) = 1 - (1 - Q(2, z)) \Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z) = 1 - \left(\frac{2-z}{2}\right)^2 \Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z).
\]

It therefore suffices to prove that $\Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z) \leq \frac{2-z}{2}$. For that, it suffices to prove that $\Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1) \leq 1 - \frac{2-z}{2}$. If $z + q_1 \leq 1$, then $\Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1) = \frac{z-q_1}{2} + \frac{1-(z-q_1)}{2} = 1 - z \leq \frac{2-z}{2}$. If $z + q_1 > 1$, then $\Pr (q_2 \leq -z - q_1 \vee q_2 \geq z - q_1) = \Pr (q_2 \geq z - q_1) = \frac{1-(z-q_1)}{2} < \frac{2-z}{2}$. Therefore,
\[
Q(3, z) > 1 - \left(\frac{2-z}{2}\right)^2 \frac{2-z}{2} = P(3, z).
\]

Now suppose $N \geq 5$. We have
\[
Q(N, z) = \Pr (\exists i, j \in \{1, \ldots, N\}, i \neq j : |q_i + q_j| \leq z) > \Pr (|q_1 + q_2| \leq z \vee \cdots \vee |q_{N-2} + q_{N-1}| \leq z \vee |q_{N-2} + q_N| \leq z \vee |q_{N-1} + q_N| \leq z) = 1 - \Pr (|q_1 + q_2| \geq z \wedge \cdots \wedge |q_{N-2} + q_{N-1}| \geq z \wedge |q_{N-2} + q_N| \geq z \wedge |q_{N-1} + q_N| \geq z) = 1 - \Pr (|q_1 + q_2| \geq z) \times \cdots \times \Pr (|q_{N-4} + q_{N-3}| \geq z)
\]
\[
\times \Pr (|q_{N-2} + q_{N-1}| \geq z \wedge |q_{N-2} + q_N| \geq z \wedge |q_{N-1} + q_N| \geq z) = 1 - (1 - Q(2, z))^\frac{N-3}{2} \times (1 - Q(3, z)) = 1 - \left(\left(\frac{2-z}{2}\right)^2\right)^\frac{N-3}{2} \left(\frac{2-z}{2}\right)^3 = P(N, z).
\]

This completes the proof.
9 Appendix C: Effect on Villagers’ Satisfaction

Higher quality of politicians may improve attitudes of the voters toward local leaders and the elections process in general and we can test this prediction by looking at the attitudes of the villagers elicited in the surveys conducted after the start of the program. Because of data limitations, however, we cannot separate satisfaction with newly elected council from the satisfaction with traditional leaders, and satisfaction with the performance of NSP-related projects from satisfaction with other activities of the village leaders.

In particular, information on villagers’ attitudes comes from two follow-up surveys. The midline survey was conducted in May-October 2009, at which point all complying villages had elected councils and selected projects, but only 18 percent of projects had been completed. By the time the endline survey was conducted in May-October 2011 virtually all projects were complete. The survey was designed to be administered to ten randomly-selected households surveyed at baseline, with separate questionnaires for male household heads and a senior woman in the household. The data provide information on 2,367 male respondents and 2,141 female respondents in the midline survey and 2,130 male and 1,858 female respondents in the endline survey.\footnote{Because of the deterioration in security conditions, we were not able to conduct surveys of male heads of household in 11 villages and of female heads of household in 33 villages. In both cases, there were no significant differences in attrition between villages with different procedures of project selection. Enumerators administering the male household questionnaire were instructed to locate and interview the same households and, whenever possible, the same villagers who participated in the baseline survey. Enumerators were able to successfully locate such respondents in 65 percent of households in which male respondents were interviewed during the baseline survey. The predominant reason for enumerators not being able to interview baseline respondents was that the person was away from home on the day that the survey team visited the village, as it was the time of harvest. Differences between villages with different procedures of project selection in individual-level attrition are not statistically significant. We also check that the effect on attrition of such characteristics of respondents as age, income, assets, size of household, education, and ethnicity are similar in villages with different project selection rules.}

To measure villagers’ satisfaction with local economic and governance outcomes, we use four perception-based binary indicators from male and female household surveys at midline and endline: (i) whether respondent thinks that village headman should be elected; (ii) whether the respondent attributes positive economic changes to actions of the village leadership;\footnote{This indicator is available only for the midline survey.} (iii) whether the respondent is satisfied with the work of the village leaders; and (iv) whether the respondent perceives that the household is better off than it was last year.
To test this prediction we estimate the following model:

$$Y_{vit} = \alpha + \gamma_1 \cdot AL_v \cdot \tau_{1t} + \gamma_2 \cdot AL_v \cdot \tau_{2t} + \phi_1 \cdot \tau_{1t} + \phi_2 \cdot \tau_{2t} + \varepsilon_{vit}$$ 

(9)

Results from the analysis of the effect of electoral rules on villagers’ attitudes toward local leaders and their perception of their economic situation are presented in Table A5 in the Online Appendix. They indicate that male villagers are more likely to support election of the village headman in the midline survey, but this effect decreases in magnitude and loses statistical significance in the survey. Female villagers in the midline survey are more likely to attribute positive economic changes to village leaders and to perceive an improvement in the economic situation of their household in villages with at-large elections, although there is no significant difference in their levels of satisfaction with the work of village leaders or in the instances of disagreement with the decisions of village leaders. There is also no evidence that electoral rules affect attitudes of male or female respondents in the endline survey.

Overall, the results are is weakly consistent with the model’s predictions, as there is some evidence that attitudes of the villagers are more positive in villages with at-large elections in the midline survey. We do not observe any effect of electoral rules on villager’s satisfaction in the endline survey, but this is not surprising, given the timing of the endline survey. The survey was conducted four years after the selection of the councils and on average a year after all development projects were finished, i.e. at this time all the activities for which the council was responsible were over and the quality of elected candidates should not have an obvious effect on the satisfaction of the villagers.
<table>
<thead>
<tr>
<th></th>
<th>Mean (Full Sample)</th>
<th>Standard Deviation (Full Sample)</th>
<th>Observation (Full Sample)</th>
<th>Mean in District Villages</th>
<th>Mean in At-large Villages</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many households are in this village in total?</td>
<td>118.43</td>
<td>115.71</td>
<td>2264</td>
<td>121.70</td>
<td>115.08</td>
<td>0.06</td>
</tr>
<tr>
<td>How many people live in this household in total?</td>
<td>9.79</td>
<td>5.00</td>
<td>2374</td>
<td>9.57</td>
<td>10.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Number of children under 15 in household</td>
<td>4.58</td>
<td>2.76</td>
<td>2374</td>
<td>4.50</td>
<td>4.66</td>
<td>0.06</td>
</tr>
<tr>
<td>Distance to the center of the village</td>
<td>402</td>
<td>1011</td>
<td>2078</td>
<td>351</td>
<td>455</td>
<td>0.10</td>
</tr>
<tr>
<td>Respondent employed in agriculture</td>
<td>0.66</td>
<td>0.47</td>
<td>2360</td>
<td>0.67</td>
<td>0.64</td>
<td>0.07</td>
</tr>
<tr>
<td>Age of the respondent</td>
<td>43.81</td>
<td>13.30</td>
<td>2336</td>
<td>43.97</td>
<td>43.65</td>
<td>0.02</td>
</tr>
<tr>
<td>Do not have formal education</td>
<td>0.71</td>
<td>0.45</td>
<td>2387</td>
<td>0.73</td>
<td>0.69</td>
<td>0.09</td>
</tr>
<tr>
<td>Finished Middle school</td>
<td>0.07</td>
<td>0.26</td>
<td>2387</td>
<td>0.07</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>Finished High school</td>
<td>0.04</td>
<td>0.20</td>
<td>2387</td>
<td>0.04</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Mother tongue is Dari</td>
<td>0.70</td>
<td>0.46</td>
<td>2387</td>
<td>0.72</td>
<td>0.69</td>
<td>0.05</td>
</tr>
<tr>
<td>Village is Ethnically Mixed</td>
<td>0.24</td>
<td>0.43</td>
<td>250</td>
<td>0.21</td>
<td>0.26</td>
<td>0.13</td>
</tr>
<tr>
<td>Never of rarely have problems supplying food</td>
<td>0.45</td>
<td>0.50</td>
<td>2387</td>
<td>0.43</td>
<td>0.47</td>
<td>0.09</td>
</tr>
<tr>
<td>Main source of drinking water is unprotected spring</td>
<td>0.27</td>
<td>0.44</td>
<td>2387</td>
<td>0.28</td>
<td>0.26</td>
<td>0.03</td>
</tr>
<tr>
<td>Have access to electricity</td>
<td>0.15</td>
<td>0.35</td>
<td>2387</td>
<td>0.14</td>
<td>0.15</td>
<td>0.04</td>
</tr>
<tr>
<td>Have a mobile phone</td>
<td>0.18</td>
<td>0.38</td>
<td>2387</td>
<td>0.19</td>
<td>0.17</td>
<td>0.04</td>
</tr>
<tr>
<td>Have a radio</td>
<td>0.75</td>
<td>0.43</td>
<td>2387</td>
<td>0.74</td>
<td>0.76</td>
<td>0.05</td>
</tr>
<tr>
<td>Expenditure on food in the last 30 days</td>
<td>3561</td>
<td>1982</td>
<td>2340</td>
<td>3524</td>
<td>3600</td>
<td>0.04</td>
</tr>
<tr>
<td>Received a loan</td>
<td>0.47</td>
<td>0.50</td>
<td>2387</td>
<td>0.48</td>
<td>0.46</td>
<td>0.05</td>
</tr>
<tr>
<td>Most preferred project is drinking water</td>
<td>0.29</td>
<td>0.46</td>
<td>2387</td>
<td>0.30</td>
<td>0.28</td>
<td>0.05</td>
</tr>
<tr>
<td>Most preferred project is irrigation</td>
<td>0.13</td>
<td>0.33</td>
<td>2387</td>
<td>0.11</td>
<td>0.15</td>
<td>0.11</td>
</tr>
<tr>
<td>Most preferred project is electricity</td>
<td>0.06</td>
<td>0.24</td>
<td>2387</td>
<td>0.06</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>Most preferred project is road or bridge</td>
<td>0.15</td>
<td>0.36</td>
<td>2387</td>
<td>0.16</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>Attended shura meetings</td>
<td>0.32</td>
<td>0.47</td>
<td>2387</td>
<td>0.33</td>
<td>0.31</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Table 2. Summary Statistics for outcome variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard error</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Council member has finished high school</td>
<td>8.7</td>
<td>28.2</td>
<td>2,016</td>
</tr>
<tr>
<td>Council member has finished middle school</td>
<td>16.8</td>
<td>37.4</td>
<td>2,016</td>
</tr>
<tr>
<td>Distance between Residences of Council Members and Village Center (meters)</td>
<td>406</td>
<td>867</td>
<td>1,018</td>
</tr>
<tr>
<td>Project Implementation Started Before:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October 2008</td>
<td>0.13</td>
<td>0.34</td>
<td>1317</td>
</tr>
<tr>
<td>January 2009</td>
<td>0.41</td>
<td>0.49</td>
<td>1317</td>
</tr>
<tr>
<td>April 2009</td>
<td>0.76</td>
<td>0.43</td>
<td>1317</td>
</tr>
<tr>
<td>July 2009</td>
<td>0.82</td>
<td>0.39</td>
<td>1317</td>
</tr>
<tr>
<td>October 2009</td>
<td>0.87</td>
<td>0.34</td>
<td>1317</td>
</tr>
<tr>
<td>Project Implementation Finished Before:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 2009</td>
<td>0.41</td>
<td>0.49</td>
<td>1317</td>
</tr>
<tr>
<td>April 2009</td>
<td>0.82</td>
<td>0.39</td>
<td>1317</td>
</tr>
<tr>
<td>July 2009</td>
<td>0.87</td>
<td>0.34</td>
<td>1317</td>
</tr>
<tr>
<td>October 2009</td>
<td>0.9</td>
<td>0.31</td>
<td>1317</td>
</tr>
<tr>
<td>January 2010</td>
<td>0.91</td>
<td>0.29</td>
<td>1317</td>
</tr>
</tbody>
</table>
Table 3. Educational Attainment of Council Members and Project Implementation Rate

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Implementation of a Project has Started Before:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>At Least One Council Member has Finished High school</td>
<td>(1) 0.019</td>
<td>(2) 0.035</td>
<td>(3) 0.241***</td>
<td>(4) 0.251***</td>
<td>(5) 0.259***</td>
</tr>
<tr>
<td>At Least One Respondent from Baseline Survey has Finished High school</td>
<td>(1) 0.099</td>
<td>(2) 0.165***</td>
<td>(3) 0.221***</td>
<td>(4) 0.255***</td>
<td>(5) 0.195***</td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.33</td>
<td>0.55</td>
<td>0.47</td>
<td>0.43</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>Implementation of a Project has Finished Before:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>At Least One Council Member has Finished High school</td>
<td>(1) 0.035</td>
<td>(2) 0.251***</td>
<td>(3) 0.240***</td>
<td>(4) 0.216***</td>
<td>(5) 0.200***</td>
</tr>
<tr>
<td>At Least One Respondent from Baseline Survey has Finished High school</td>
<td>(1) 0.165***</td>
<td>(2) 0.255***</td>
<td>(3) 0.199***</td>
<td>(4) 0.267***</td>
<td>(5) 0.264***</td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
<td>1,285</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.55</td>
<td>0.43</td>
<td>0.33</td>
<td>0.42</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
Table 4. Effect of Electoral Rules on Educational Attainment of Council Members

<table>
<thead>
<tr>
<th></th>
<th>Council Member Has Finished High school</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>At-large elections</td>
<td>3.96***</td>
</tr>
<tr>
<td></td>
<td>[1.32]</td>
</tr>
<tr>
<td>Fractionalized preferences over projects</td>
<td>7.97***</td>
</tr>
<tr>
<td>At-large elections</td>
<td></td>
</tr>
<tr>
<td>Fractionalized preferences over projects</td>
<td>-3.1</td>
</tr>
<tr>
<td></td>
<td>[1.99]</td>
</tr>
<tr>
<td>Ethnically Mixed Village* At-large elections</td>
<td>7.96**</td>
</tr>
<tr>
<td></td>
<td>[3.21]</td>
</tr>
<tr>
<td>Ethnically Mixed Village</td>
<td>-3.27</td>
</tr>
<tr>
<td></td>
<td>[2.21]</td>
</tr>
<tr>
<td>Large Village* At-large elections</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Village</td>
<td>-3.36*</td>
</tr>
<tr>
<td></td>
<td>[1.96]</td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,016</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
### Table 5. Effect of Electoral Rules on Project Implementation

#### Panel A

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>At-Large</td>
<td></td>
<td>0.001</td>
<td>0.016</td>
<td>0.073**</td>
<td>0.052**</td>
<td>0.053***</td>
</tr>
<tr>
<td></td>
<td>Quadruple fixed effects</td>
<td>[0.05]</td>
<td>[0.05]</td>
<td>[0.03]</td>
<td>[0.03]</td>
<td>[0.02]</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
<tr>
<td>R-squared</td>
<td></td>
<td>0.36</td>
<td>0.58</td>
<td>0.53</td>
<td>0.43</td>
<td>0.17</td>
</tr>
</tbody>
</table>

#### Panel B

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>At-Large</td>
<td></td>
<td>0.016</td>
<td>0.052**</td>
<td>0.052***</td>
<td>0.014**</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Quadruple fixed effects</td>
<td>[0.05]</td>
<td>[0.03]</td>
<td>[0.02]</td>
<td>[0.01]</td>
<td>[0.00]</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
<tr>
<td>R-squared</td>
<td></td>
<td>0.58</td>
<td>0.43</td>
<td>0.18</td>
<td>0.15</td>
<td>0.06</td>
</tr>
</tbody>
</table>

---

*Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.*
<table>
<thead>
<tr>
<th>At-large elections</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln (Distance between Residences of Council Members and Village Center)</td>
<td>-0.32***</td>
<td>-0.28***</td>
</tr>
<tr>
<td></td>
<td>[0.11]</td>
<td>[0.08]</td>
</tr>
<tr>
<td>Ln (Median Distance between Residences of Villagers)</td>
<td>0.75***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.06]</td>
<td></td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,003</td>
<td>1,003</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
<td>0.36</td>
</tr>
</tbody>
</table>

*Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
### Table 7: Location of Council Members’ Residences and their Educational Attainment

<table>
<thead>
<tr>
<th></th>
<th>Council Member Has Finished High school</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Ln (Distance between Residences of Council Members and Village Center)</td>
<td>-0.65</td>
</tr>
<tr>
<td></td>
<td>[1.22]</td>
</tr>
<tr>
<td>At-Large * Ln (Distance between Residences of Council Members and Village Center)</td>
<td>4.45**</td>
</tr>
<tr>
<td></td>
<td>[2.24]</td>
</tr>
<tr>
<td>At-Large</td>
<td>-16.16</td>
</tr>
<tr>
<td></td>
<td>[11.20]</td>
</tr>
<tr>
<td>Ln (Median Distance between Residences of Villagers)</td>
<td>4.10**</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>p-value for the effect of distance in at-large villages</td>
<td>0.05</td>
</tr>
<tr>
<td>Observations</td>
<td>857</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
</tr>
</tbody>
</table>

*Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.*
Table A1. Effect of Electoral Rules on Educational Attainment of Council Members

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At-large elections</td>
<td>3.43*</td>
<td>-2.48</td>
<td>0.6</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>[1.93]</td>
<td>[2.76]</td>
<td>[1.98]</td>
<td>[2.55]</td>
</tr>
<tr>
<td>Fractionalized preferences over projects * At-large elections</td>
<td>12.17***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[4.27]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fractionalized preferences over projects</td>
<td>-3.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[3.05]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village* At-large elections</td>
<td>11.33**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[5.49]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village</td>
<td></td>
<td>-4.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[4.26]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Village* At-large elections</td>
<td></td>
<td></td>
<td>7.24*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[4.11]</td>
<td></td>
</tr>
<tr>
<td>Large Village</td>
<td></td>
<td></td>
<td>-4.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[2.80]</td>
<td></td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2,016</td>
<td>2,016</td>
<td>2,016</td>
<td>2,016</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
### Table A2. Educational Attainment of Council Members and Project Implementation (IV results)

#### Panel A

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td></td>
</tr>
<tr>
<td>At Least One Council Member has Finished High school</td>
<td>0.004</td>
<td>0.083</td>
<td>0.374**</td>
<td>0.266*</td>
<td>0.272**</td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>[0.22]</td>
<td>[0.25]</td>
<td>[0.18]</td>
<td>[0.14]</td>
<td>[0.11]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>

#### Panel B

<table>
<thead>
<tr>
<th>Implementation of a Project has Finished Before:</th>
<th>January 2009</th>
<th>April 2009</th>
<th>July 2009</th>
<th>October 2009</th>
<th>January 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td></td>
</tr>
<tr>
<td>At Least One Council Member has Finished High school</td>
<td>0.083</td>
<td>0.266*</td>
<td>0.265**</td>
<td>0.072**</td>
<td>0.015</td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>[0.25]</td>
<td>[0.14]</td>
<td>[0.11]</td>
<td>[0.03]</td>
<td>[0.01]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>

**Note:** Results of the IV estimation in which the variable “At Least One Council Member Has Finished High school” is instrumented with the dummy variable for at-large elections. Standard errors clustered at the village level in parentheses.

*significant at 10%; ** significant at 5%; *** significant at 1%.
Table A3. Educational Attainment of Male Council Members  
(excluding districts with more than one member elected)

<table>
<thead>
<tr>
<th>Council Member has Finished High school</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At-large elections</td>
<td>4.04***</td>
<td>0.36</td>
<td>2.46*</td>
<td>-1.60</td>
</tr>
<tr>
<td></td>
<td>[1.43]</td>
<td>[1.94]</td>
<td>[1.40]</td>
<td>[1.56]</td>
</tr>
<tr>
<td>Fractionalized preferences over projects</td>
<td>7.75**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At-large elections</td>
<td>[3.51]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fractionalized preferences over projects</td>
<td>-2.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.02]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village* At-large elections</td>
<td>6.11*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[3.21]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village</td>
<td>-2.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.15]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Village* At-large elections</td>
<td>11.98***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[3.15]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Village</td>
<td>-3.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.02]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,716</td>
<td>1,716</td>
<td>1,716</td>
<td>1,716</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.21</td>
</tr>
</tbody>
</table>

*Note: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
Table A4. Electoral Rules and Incumbency Advantage

<table>
<thead>
<tr>
<th>Definition of Elite</th>
<th>Percent of Male Council Members who were Members of Pre-Existing Elite</th>
<th>Percent of Members of Pre-Existing Elite Elected to Council &amp; Office</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean in District Elections</td>
<td>Difference</td>
</tr>
<tr>
<td>Member Of The Baseline Focus Group (Including Those Who Could Not Attend)</td>
<td>31.9</td>
<td>2.43</td>
</tr>
<tr>
<td>Observations</td>
<td>1055</td>
<td>2058</td>
</tr>
<tr>
<td>Decisions-Maker According To Male Focus Group</td>
<td>13.2</td>
<td>-0.54</td>
</tr>
<tr>
<td>Observations</td>
<td>1055</td>
<td>2058</td>
</tr>
<tr>
<td>Decisions-Maker According To Male Head Of Household Survey</td>
<td>20.7</td>
<td>3.24*</td>
</tr>
<tr>
<td>Observations</td>
<td>1055</td>
<td>2058</td>
</tr>
<tr>
<td>Decisions-Maker According To Female Individual Survey</td>
<td>14.9</td>
<td>-0.66</td>
</tr>
<tr>
<td>Observations</td>
<td>1055</td>
<td>2058</td>
</tr>
<tr>
<td>Either Of The Four Above</td>
<td>38.9</td>
<td>3.20</td>
</tr>
<tr>
<td>Observations</td>
<td>1055</td>
<td>2058</td>
</tr>
</tbody>
</table>

Note: The difference between district and at-large elections from regression (1). Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
<table>
<thead>
<tr>
<th>Table A5: Effect of Electoral Rules on Voter Attitudes</th>
<th>Respondent Prefers to Elect Village Head</th>
<th>Respondent is Satisfied with Work of Village Leaders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male Respondents</td>
<td>Female Respondents</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>At-large *Midline Survey</td>
<td>0.05**</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>[0.02]</td>
<td>[0.03]</td>
</tr>
<tr>
<td>At-large *Endline Survey</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>[0.02]</td>
<td>[0.03]</td>
</tr>
<tr>
<td>Quadruple* Survey fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>4,440</td>
<td>3,578</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.23</td>
<td>0.12</td>
</tr>
</tbody>
</table>

|                                                      | Male Respondents | Female Respondents | Male Respondents | Female Respondents |
|                                                      | (1)             | (2)             | (3)             | (4)             |
| At-large *Midline Survey                             | -0.28           | 1.27**          | -0.78           | 4.45*           |
|                                                    | [0.832]         | [0.557]         | [2.473]         | [2.517]         |
| At-large *Endline Survey                             | -1.18           |                  | -1.59           |                  |
|                                                    |                  |                  | [1.705]         | [1.949]         |
| Quadruple* Survey fixed effects                      | Yes             | Yes             | Yes             | Yes             |
| Observations                                        | 2,355           | 2,135           | 4,493           | 3,997           |
| R-squared                                           | 0.08            | 0.04            | 0.19            | 0.15            |

Note: Outcomes measured in percent. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.
Figure A1. Distribution of Dates when Project Implementation Started.

Note: Red lines indicate the dates for which we construct indicators for whether each project has started or not, which is used in the empirical analysis.

Figure A2. Distribution of dates when Project Implementation Finished.

Note: Red lines indicate the dates for which we construct indicators for whether each project has finished or not, which is used in the empirical analysis.