Electoral Rules and the Quality of Politicians: Theory and Evidence from a Field Experiment in Afghanistan

Andrew Beath† Fotini Christia‡ Georgy Egorov§ Ruben Enikolopov¶

October 30, 2013

Abstract

Electoral rules matter for how individuals’ information and policy preferences get aggregated, but little is known about how they affect the quality of elected officials. To address this question we exploit a unique field experiment which introduced randomized variation in the method of elections of village councils in 250 Afghan villages. In particular, we compare at-large elections with district elections. We propose a theoretical model where the difference in electoral systems occurs because elected legislators have to bargain over policy, and in district elections, citizens have a strategic motive to elect candidates with more polarized policy positions at the expense of competence. Empirical results prove consistent with the predictions of the model. Specifically, we find that competent candidates are more likely to be elected in at-large elections and that this difference is larger in more heterogenous villages.

*Second draft. Comments are very welcome. We thank Mattias Polborn, Horacio Larreguy, Hulya Eraslan and participant of the NBER Summer Institute Political Economy Public Finance meeting and Priorat Workshop on Bargaining and Politics 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. cfofotini@mit.edu
§Kellogg School of Management and NBER. g-egorov@kellogg.northwestern.edu
¶Barcelona Institute for Political Economy and Governance, Universitat Pompeu Fabra; New Economic School. renikolopov@nes.ru
1 Introduction

Our paper contributes to the relatively recent and growing literature that looks at the determinants of political selection. In particular, we look at the effect of electoral rules on the quality of elected representatives.\textsuperscript{1} We provide a theoretical model that generates a number of testable predictions along with empirical evidence based on a field experiment that proves consistent with predictions of the model.

The model compares two sets of electoral rules – district and at-large elections – with respect to the quality of elected officials. Once elected, the officials bargain over a joint policy decision. The model predicts that the quality of elected officials will be higher in at-large elections and that this difference will be bigger in more heterogeneous communities.

Our model falls into the tradition of citizen-candidate models, which consider political competition as a three-stage game of entry, voting, and policy making (Besley and Coate, 1997, Osborne and Slivinski, 1996). In our paper we assume that running for office is costless, so we do not consider the entry stage of the game (cf. Besley and Coate, 1998). This assumption, which is consistent with the setup of the field experiment from which the empirical evidence is derived, makes the model more tractable without being a major driving force for our results.

The empirics on the effect of electoral rules on the quality of elected representatives come from a field experiment that varies such 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 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 they are allowed to vote and council members are elected based on the number of votes garnered across the whole village. In addition, voters in

\textsuperscript{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).
at-large elections can vote for up to three different candidates, while voters in district elections can only cast a single vote. Under both electoral rules all villagers are considered candidates and there is no formal political campaigning. Empirical results indicate 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 village size and ethnic composition. The results also indicate that in small or ethnically homogenous villages there is no significant effect of electoral rules on the quality of the elected official. There is also some evidence that the satisfaction of citizens with their local leaders is higher in villages with at-large elections. Overall, the empirical results are consistent with the theoretical predictions of the model.

One reason for a society to elect incompetent politicians is asymmetry of information, i.e. the problem of not knowing who is competent. In Caselli and Morelli (2004), able individuals have better outside options, and thus incompetent politicians are overrepresented in the pool of candidates. Other papers emphasize strategic reasons behind choosing politicians other than the best available. Mattozzi and Merlo (2007) describe a general equilibrium effect: if a party hires the best candidates, it makes them attractive for lobbying firms and thus has to pay a very high wage, so in equilibrium, mediocre individuals are chosen. McKelvey and Reizman (1992) show that seniority rules, common in legislatures, give rise to an incumbency advantage that may prevent competent challengers from taking over. Acemoglu et al. (2010), in a model where the current government has some influence on the composition of future government, show that farsighted government members would oppose competence-improving changes out of fear that they would be outed in the future. Egorov and Sonin (2011) show that fear of betrayal may also lead to selection of a less-than-optimal agent. In Aragones and Palfrey (2004), two partisan candidates of different quality compete by choosing political positions, and a competent but biased candidate may fail to get elected.

Our interest is in the interplay between political preferences of voters and competence of elected politicians. Besley et al. (2005) examine how the quality of politicians affects corruption.

---

2 Banks 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.
and find that politicians with a higher level of education are less likely to use political power opportunistically. This result indicates that the quality of politicians has tangible effects on policy outcomes and that at least in developing countries education is a good proxy for their competence. In a related paper, Banerjee and Pande (2007) show that a party that has an electoral advantage can slack on the quality of its candidates, whereas a disadvantaged party has to provide candidates of higher quality in order to remain competitive. Our paper suggests that even without party politics, in a purely citizen-candidate environment, and in the absence of exogenously more numerous or advantaged groups, heterogeneity of voters’ preferences nevertheless may hurt the competence of elected politicians. Moreover, we show that this negative effect depends on the nature of electoral rules.

Existing literature that compares different electoral rules focuses mainly on their effect on the level of corruption and the types of public goods provided. Persson and Tabellini (1999, 2000) argue that proportional systems lead to higher levels of rent extraction by politicians. Persson, Tabellini, and Trebbi (2003) provide evidence that larger voting districts are associated with less corruption, whereas larger shares of candidates elected from party lists are associated with more corruption. Persson and Tabellini (2004) show that majoritarian elections lead to smaller governments and smaller welfare programs than proportional elections. Other 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). However, this literature does not consider the effect of electoral rules on the quality of politicians. An exception is Myerson (1993) who provides a theoretical model that examines the effect of district magnitude on quality of politicians and argues 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.3 Our paper suggest that electoral rules can affect the quality of elected officials even in the absence of entry barriers. To the best of our knowledge, ours is the first paper that exploits a field experiment to provide empirical evidence on the causal effect of electoral rules on the quality of elected officials.

3See also Adams (1996), who considers the effect of a constitutional change in Illinois in 1980 on business friendliness of state legislators.
The rest of the paper is organized as follows: Section 2 describes the theoretical model. Section 3 provides an analysis for the model. Section 4 formulates the model’s empirical predictions. Section 5 presents the experimental design. We describe the data in section 6. Section 7 offers the empirical results. We discuss the results in section 8 and conclude in section 9.

2 Model

The society consists of a continuum of individuals who have different policy preferences in a one-dimensional policy space. Their ideal points are uniformly distributed on \([-B, B]\), and the utility of an individual \(i\) with ideal point \(b_i \in [-B, B]\) if policy \(p\) is chosen is given by

\[
u_i(p) = -k (b_i - p)^2
\]

for some positive constant \(k\), naturally interpreted as the importance of the policy issue. Most individuals have low competence \(a_i = 0\), but there are a few individuals with high human capital that have competence \(a_i = h > 0\). The main question of interest in the analysis is the probability of electing high-quality candidates under different electoral rules. To keep the model as simple as possible, assume there is only one such individual.\(^4\) His ideal point \(b\) is uniformly distributed on the same interval \([-B, B]\). To ensure existence of an equilibrium, let us assume that for any \(x \in [-B, B]\) there is a low-competence individual with that ideal point (i.e., \(i\) with \(a_i = 0\), \(b_i = x\)).

The society elects a governing body, say a council, from among its members. We compare the following two procedures: district elections and at-large elections. In district elections, the society is divided into two districts: left, containing individuals with \(b_i < 0\), and right, containing individuals with \(b_i \geq 0\). Each individual casts a vote for one of the members of his/her district (i.e. every individual is considered a candidate). Then in each district, the individual who got the largest share of votes gets elected. In the case of a draw a random person from those receiving the most votes gets elected. In at-large elections, there is only one district containing

\(^4\)This simplifies the analysis considerably. In Appendix A, when discussing the robustness of our results, we argue that the main insights hold with two or more competent individuals as well.
the entire society, and each individual casts two votes for two different citizens. In this case, the
two candidates who received the most votes get elected.

The two elected legislators, of types \((a_l, b_l)\) and \((a_r, b_r)\) then make a joint decision on policy
\(p\). We assume that they find a point that maximizes their joint utility (thus \(p = \frac{b_l + b_r}{2}\)).\(^5\) In
addition, all individuals in the society benefit from having competent politicians; so the total
utility of an individual \(i\) from this legislative body is
\[
\begin{align*}
    w_i(a_l, b_l, a_r, b_r) &= a_l + a_r + u_i \left( \frac{b_l + b_r}{2} \right) \\
    &= a_l + a_r - k \left( \frac{b_l - b_r}{2} \right)^2.
\end{align*}
\]

We study the strategic behavior of voters under the following equilibrium concept. The
strategy of each voter in district elections is \(\lambda (i)\), the identity of the individual in his district
he casts his vote for. The strategy of each voter in at-large elections is \(\Lambda (i) = (\lambda_1 (i), \lambda_2 (i))\),
which reflects the unordered pair of individuals he votes for. In both cases, we assume that
voters have perfect knowledge of all their fellow society members, in particular, they know their
policy preferences as well as the identity of the most competent member.\(^6\) All voting decisions
are made simultaneously, which gives rise to a coordination problem. We make the following
refinement.

**Definition 1** Voting strategies \(\{\lambda_i\}\) or \(\{\Lambda_i\}\) constitute an equilibrium if for any electoral district
(\(l\) or \(r\) in the first case, or the entire society in the second) there is no subset of voters from this
district who would strictly improve the utility of all members of this subset by choosing different
voting strategies.

In other words, we allow for deviations by coalitions of voters, but only within a district.

### 3 Analysis of the Model

We are comparing the expected outcome of district elections with that of at-large elections.

\(^5\)This could be modeled explicitly by specifying a Baron and Ferejohn (1989)-type extensive-form legislative
bargaining game. We can also consider a situation where only one politician (each with equal probability) decides
the policy; as these two cases would be the extremes of Baron-Ferejohn legislative bargaining game, with very
patient and very impatient legislators, respectively. We consider the simplest case to save on notation and space.

\(^6\)We could assume that they know the competent person’s policy preferences only if he is in the same voting
district; this would not change the results qualitatively.
Our first result is that in the model, the median voter theorem applies, and each district will elect the legislator (or the pair of legislators) most favored by the median voter in that district. In fact, this individual (or pair of individuals) will be the Condorcet winner, who exists in our model.

Consider first the district elections, and consider the voting decisions of people in the right district, holding those in the left district fixed. In other words, suppose that the left district elects a citizen of type \((a_l, b_l)\) distributed with some probability distribution \(L\) the uncertainty reflects the possibility of draws \(M\). The expected utility of a citizen with ideal point \(x\) if \((a_r, b_r)\) is elected is

\[
U_x (a_r, b_r) = \mathbb{E}_{a_l, b_l} \left( a_l + a_r - k \left( \frac{b_l + b_r}{2} - x \right)^2 \right)
\]

\[
= \mathbb{E} a_l + a_r - k \frac{b_l^2}{4} - \frac{1}{2} k b_l \mathbb{E} b_l - \frac{1}{4} k \mathbb{E} b_l^2 - k x^2 + k x \mathbb{E} b_l + k x b_r
\]

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

One can easily see that for two alternative candidates \((a_i, b_i)\) and \((a_j, b_j)\) with \(b_i > b_j\), \(U_x (a_i, b_i) - U_x (a_j, b_j)\) is increasing in \(x\). This implies that the problem has a (generically unique) Condorcet winner, who coincides with the alternative candidate that the median voter (in the right district) prefers, i.e., the voter with \(b = B/2\). The same logic applies to the left district, and we further establish that these candidates will be the ones elected in equilibrium.

In the case of at-large elections, citizens vote for two alternative candidates simultaneously. The utility of voter \(x\) if \((a_{i1}, b_{i1})\) and \((a_{i2}, b_{i2})\) are elected equals

\[
V_x (a_{i1}, b_{i1}; a_{i2}, b_{i2}) = \left( a_{i1} + a_{i2} - k \left( \frac{b_{i1} + b_{i2}}{2} - x \right)^2 \right);
\]

notice that \(V_x (a_{i1}, b_{i1}; a_{i2}, b_{i2}) - V_x (a_{j1}, b_{j1}; a_{j2}, b_{j2})\) is increasing in \(x\) whenever \(b_{i1} + b_{i2} > b_{j1} + b_{j2}\). Again, this problem has a Condorcet winner, who is the pair most preferred by the median voter of this district 0, and we will show that only such pairs may be elected.

**Proposition 1** In both district elections and at-large elections, equilibria exist and the types of
elected politicians are (generically) uniquely determined. 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 \( a_i = h \)) or the most biased individual (with \( b_i = \pm B \)).

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

The proposition states that the equilibrium concept stated in Definition 1 is sufficiently strong to pick a (generically) unique equilibrium. To build an intuition for the predictions of the model, consider the at-large elections first. For the median voter with bliss point 0, an ideal outcome is where one of the elected candidates is competent, and both would negotiate his ideal policy 0. He can achieve this ideal outcome by choosing the smartest politician and the candidate opposite in terms of politics.

Consider now the citizens voting in district elections. The reasons to elect the most competent of them are clear. But what is the rationale to elect the -biased individual? To answer this question, suppose that the competent individual is in the left district, and we consider the problem of the voters in the right district, in particular, that of the median voter there (with preference \( \frac{B}{2} \)). If an individual of type \((0, b_r)\) is elected, his expected utility will be

\[
U_{B/2} (0, b_r) = \mathbb{E}a_l - k \left( \frac{\mathbb{E}b_l + b_r}{2} - \frac{B}{2} \right)^2 - \frac{k}{4} \text{Var} (b_l).
\]

But \( \mathbb{E}b_l \leq 0 \) and \( b_r \leq B \), which implies that his utility is strictly increasing in \( b_r \), thus reaching its maximum for \( b_r = B \). In other words, regardless of voting strategies of citizens in the left district, the median voter of the right district prefers the most biased candidate.\(^7\) The same effect causes the median voter in the left district 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 result tells us when exactly this happens.

\(^7\)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 a non-bounded support. However, the tendency to elect relatively biased candidate would remain. We maintain the assumption on uniform distribution for expositional purposes.
Proposition 2 In district elections, both districts elect the most biased candidates if

\[(3B - |b|)(B - |b|) > \frac{h}{k},\]  \hspace{1cm} (3)

where \(b\) is the ideal point of the competent citizen. If (3) does not hold, then one district elects the most biased of its members, and the other elects the competent one. The competent citizen is more likely to be elected if:

1. the society is more homogenous, i.e. less polarized in their preferences (\(B\) is lower);
2. the competent person holds more biased views (\(|b|\) is higher);
3. his competence is more pronounced (\(h\) is higher);
4. policy matters less relative to competence (\(k\) is lower).

This result comes from the comparison of the district’s median voter’s utility from electing either the most biased citizen or the most competent one. Given this trade-off, the comparative statics results are natural, but noteworthy nonetheless: competent individuals are likely to get elected when competence is more important (\(h\) is high and \(k\) is low), but more interestingly, the election of such an individual depends on the degree of polarization in society. Higher polarization makes the median voter more sensitive to the ideal point of the legislator he elects, and this hurts the prospects of a more moderate competent individual. In addition, since biased, rather than moderate, candidates are preferred, the competent one has higher chances if he is himself biased.

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 elected legislators is (weakly) higher;
2. the expected polarization (measured by either the distance of elected legislators from each other or from the society’s median voter) is (strictly) lower.
There is no difference in expected competence if voters care little about policy (k is low) or there is little polarization (B is low); otherwise the gap in competence is increasing in both k and B. The difference in polarization is strictly increasing in k and B.

The result for competence easily follows from the fact that the competent individual is always elected in at-large elections, which is not always the case in district elections. 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 if one recalls that the two legislators elected in at-large elections are as far from the median as the competent one, while in district elections, both districts elect individuals weakly further away from the median. As a result, district elections lead to more polarization. As B or k become larger, the gap for competence increases solely due to adverse effects in district elections. At the same time, higher B or k increases polarization in both cases; despite that, we get an unambiguous comparative statics result in this case, too.

Apart from the empirical predictions, our model has clear welfare implications. The next results demonstrate that at-large elections dominate district elections in a number of dimensions.

**Proposition 4 (Voter Satisfaction.) In at-large elections, compared to district elections:**

1. Social welfare is higher (as measured by the integral of citizens’ expected utilities);

2. Total competence of the two elected politicians is higher;

3. Policy is closer in expectation to the median voter’s ideal point, and the voters’ aggregate utility from policy is higher.

In addition, the ideal points of a random citizen and a random elected legislator are closer in expectation in at-large elections than in district elections.

**4 Empirical Predictions**

The theoretical results give rise to several empirical predictions, which we test using data from a field experiment.
1. The quality of elected candidates is higher in at-large elections as compared with district elections.

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

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

4. People’s satisfaction is higher in at-large elections as compared with district elections.

The first three predictions follow from Proposition 3, where increasing village heterogeneity corresponds to higher polarization $B$ or to an increasing relative importance of the policy dimension $k$ (or both). The fourth prediction follows from Proposition 2 and the last prediction follows from Proposition 4.

5 Experimental Design

We test the model’s predictions using randomized variation in electoral rules during the elections of village councils in Afghan villages. This study is 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 2013) as well as the program itself (Beath, Christia, and Enikolopov 2012, forthcoming). This section provides further details on NSP (5.1), describes the variation in electoral rules introduced across the 250 treatment villages under evaluation (5.2), details the sample and randomization procedures (5.3), as well as the phasing of the intervention and the data collection (5.4).

5.1 Setting

The elections of the village councils are part of the National Solidarity Program (NSP), which was devised in 2002 by the Government of Afghanistan to deliver services and infrastructure to the country’s rural population and build representative institutions for village governance. NSP has been implemented in 31,000 villages in all of Afghanistan’s 34 provinces and has disbursed over $1.1 billion to participating villages, 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 CDCs 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 a gender-balanced CDC through a secret-ballot, universal suffrage election. Once CDCs are formed, NSP disburses block grants valued at $200 per household to fund local development projects, 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 CDC in consultation with the village community. Selected projects are ordinarily focused on the construction or rehabilitation of infrastructure (e.g., drinking water facilities, irrigation canals, roads and bridges, or electrical generators) and human capital development (e.g., training and literacy courses). Overall, the primary task of the newly elected village council is to design, select and implement NSP-funded projects.

NSP intends 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 CDC is also uncertain. Per NSP rules, villages are supposed to hold re-elections for CDC positions every four years, although as follow-up elections are not facilitated, it is unclear whether these occur. Given this and the general uncertainty which accompanies planned future development activity in Afghanistan, villagers perceive NSP as a one-shot event, which does not provide strong reelection incentives to CDC members.

5.2 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

---

8The average block grant in the villages included in the sample was approximately $31,000.
9However, there is evidence that CDCs assume some responsibilities traditionally accorded to customary leaders, such as mediating conflicts, providing emergency assistance, and certifying documents (Beath et al. 2010).
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 in order for it to be valid. Campaigning is prohibited- that is, villagers interested in being elected to the council should not campaign in any way for the position. The council should 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 were segmented into geographically contiguous districts containing between 5 and 25 families regardless of the set of electoral rules. 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:\footnote{A detailed guide on the procedures is available at: http://www.nsp-ie.org/sti/stile.doc}

\textit{District Election:} Voters are restricted to casting a ballot for a single candidate, who must also live in their district. In each district, the one male and one female with the largest number of votes are elected to the council as representatives of their district. This method represents a single-ballot, simple plurality election with multiple districts (Cox, 1997), similar to the Anglo-American first-past-the-post system.

\textit{At-large Election:} Under this method, no geographical constraints are placed on vote choice. Council members are the male and female candidates receiving the most votes across the village. To ensure a sufficient number of elected members, voters cast ballots for a maximum of three different people, who are not ranked.\footnote{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.} The at-large election method is 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) the higher district magnitude (multiple elected members instead of one) and 2) the number of votes cast (three instead of one).\footnote{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.}

In all villages, council elections were organized and administered by ‘social organizers’ employed by the contracted NGOs acting as implementing partners. After the elections, the newly
elected council selected four of its members of each gender to the executive council, which consists of the council head, deputy head, secretary and treasurer.

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.\footnote{A detailed description of the monitoring results can be found at: http://www.nsp-ie.org/reports/CDCE-MR.pdf}

\subsection*{5.3 Sample and Randomization}

The randomization of allocation procedures occurred in 250 villages assigned to receive NSP that formed the treatment group for the randomized impact evaluation of NSP (Beath et al. 2010, Beath, Christia, and Enikolopov 2012, forthcoming). 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 minorities.

Data from the 2007–08 National Risk and Vulnerability Assessment (NRVA) allows for a comparison of the 250 treatment villages with a randomly-selected stratified sample of the population of rural Afghanistan. Although there are no significant differences in the age of respondents or in their income (see Table A1 in the Online Appendix), evaluation villages are more likely to be engaged in production activities related to agriculture, have slightly worse access to medical services and better access to electricity, although the magnitude of these differences is quite small. These small differences are likely to be driven by the fact that villages located close to provincial centers were more likely to receive NSP in the first phase of the program, which concluded before the start of the study.

In addition to electoral rules, the NSP impact evaluation also randomized the project selec-
tion procedures (for more on those results see Beath, Christia and Enikolopov 2013). 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 allocation procedures and council election procedures. This assignment procedure ensures that each village in the sample had an equal probability of being assigned to either procedure.

The randomization resulted in a well-balanced set of villages. Table 1 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.

5.4 Phasing of Intervention and Data Collection

The baseline survey was administered in September 2007, prior to the assignment of allocation procedures. CDC elections occurred between October 2007 and May 2008 and the disbursement of monetary grants between November 2007 and August 2008. Council elections were monitored, providing additional data on the processes. Two follow-up survey were administered: a midline follow-up survey between May and October 2009 and an endline follow-up survey between May and October 2011.

6 Data

The data employed in the empirical analysis come from several different sources. Information on the characteristics of elected council members was supplied for by 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

---

14 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).

15 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.
district of residence. Overall, data was provided for 5,129 candidates from 241 villages of whom 4,068 became council members. In particular, we construct a dummy variable that equals one if the person has full secondary education (i.e. finished high school) and zero otherwise. We restrict our analysis to male candidates, as the level of education of female candidates was so low, that there was almost no variation in educational attainment. Even among male candidates only 9% have finished high school (see Table 2). We also show that our results are robust to using a dummy variable for getting lower secondary education (i.e. finishing middle school) as a measure of candidate quality.

Information on the size and ethnic composition of villages comes from the baseline survey, conducted before the start of the program. Specifically, the baseline data used in the analysis comes from surveys that were administered to ten randomly selected male heads of household in each village and from a focus group of male village leaders in each village, which contain information on a wide range of characteristics, including demographic and socio-economic information for both ordinary villagers and village leaders prior to council elections. The survey covered 2,387 male heads of household and 1,999 male village leaders. The information on the geographic size of the village was inferred from the GPS coordinates of the houses of the ten randomly selected heads of household. We calculate the average distance between the households of the respondents 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 ethnic heterogeneity of villages we use a dummy variable that equals zero if all villagers belong to one ethnicity and one otherwise. The measure is based on a question that asks heads of households to indicate their ethnicity.

16 Of the 9 villages for which the data was not received, 7 villages did not comply with the assignment of NSP treatment, which was driven primarily by a confusion between villages with similarly sounding names. For the remaining 2 villages, the NGO did not provide the necessary information. In both cases attrition was not correlated with the assigned electoral rules.

17 There is also evidence that education of the leaders affects governance outcomes. In particular, Besley et al (forthcoming) show that higher education of the country leaders leads to higher growth.

18 More than 90% of female candidates did not have any formal education and only 0.8% had finished high school.

19 The results are similar if we use measure of ethnic fractionalization instead, although the measure of fractionalization is not very reliable given the small number of observations per village.

20 The question distinguished seven ethnicities: Pashtun, Tajik, Hazara, Uzbek, Turkmen, Baluch, and other.
Information on villagers’ satisfaction with local economic and governance outcomes 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. 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 the respondent disagrees with a recent decision or action of the village leadership; (ii); whether the respondent attributes positive economic changes to actions of the village leadership; (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.

7 Empirical Results

In the empirical analysis we use educational attainment as a proxy for candidate quality. To test the first empirical predictions of the model we use the following OLS model:

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.

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.

This indicator is available only for the midline survey.
where $Y_{vi}$ is the measure of educational attainment for candidate $i$ in village $v$, $AL_i$ 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 $\tau > 0$.

The 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 compared to the average of 7 percent among councils members elected under district elections (the difference is statistically significant at 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:

$$ Y_{vi} = \alpha + \tau \cdot AL_v + \phi_q + \varepsilon_{vi} $$

(4)

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

---

23 We include quadruple fixed effects to account for the allocation of treatment to villages (villages were first assigned to either the treatment and control groups, with treatment villages then assigned to either referenda or consultation meetings for project selection and either at-large elections or neighborhood elections for council elections) through quadruple-wise matching (Bruhn and McKenzie, 2009).
\( \gamma = 0 \) (as long as the measure of heterogeneity is normalized to zero in homogenous villages).

We use two alternative measures of heterogeneity – geographic size of the village and ethnic heterogeneity. The choice of these two measures is driven by the nature of the tasks performed by council members. The main tasks of elected council members is to guide the choice of development projects to be implemented and then to oversee the implementation of the projects. Whereas, at the implementation stage the quality of the candidates is likely to be the most relevant characteristic of the candidates, at the project selection stage, candidate’s own preferences over projects play a more important role. In particular, council members can affect both the type and the location of development projects, which can be treated as the policy dimension in the model. Ethnic heterogeneity is often used in the literature to capture differences in tastes (e.g. Alesina, Baqir, Easterly 1999) that might affect preferences over types of projects. Preferences over the location of the projects are likely to be driven by the location of the residences of the voters and candidates. The larger the village the more divergent the preferences of the villagers with respect to the location of the projects (which corresponds to higher \( B \)) and the more important the policy dimension relative to the quality of candidates (which corresponds to higher \( k \)). Thus, both ethnic heterogeneity and the size of the village are likely to capture heterogeneity of villages referred to in the empirical predictions of the model.

The results indicate that the effect of at-large elections is indeed significantly stronger in more heterogeneous villages for both measures of heterogeneity (columns 2 and 3 in Table 3). Moreover, in homogenous villages there is no significant effect of electoral rules on the quality of elected candidates. Thus, the results confirm the second and third empirical predictions of the model. The results also indicate 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 a randomized variation in electoral rules, so it can be driven by endogeneity bias.

The fourth empirical prediction posits that at-large elections will lead to higher villager satisfaction with their local leaders. To test this prediction we estimate the following model:

\[
Y_{vit} = \alpha + \gamma_1 \cdot AL_v \cdot \tau_{1t} + \gamma_2 \cdot AL_v \cdot \tau_{2t} + \phi_q \cdot \tau_{1t} + \phi_q \cdot \tau_{2t} + \varepsilon_{vit}
\]  

(6)
where \( Y_{vit} \) is the measure of satisfaction of respondent \( i \) in village \( v \) during the midline (1) or endline (2) survey \( t \in 1, 2 \), \( \tau_{1t} \) and \( \tau_{2t} \) are dummy variables for the midline and endline surveys respectively.

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 4. They indicate that there is no significant difference in the attitudes of male villagers in either midline or endline survey. Female villagers in midline survey, however, 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 female respondents in the endline survey. Overall, the results provide some limited support for the fifth empirical prediction, as the attitudes of female respondents are more positive in villages with at-large elections in the midline survey, while the attitudes of female respondents in the endline survey and the attitudes of male respondents in both midline and endline surveys appear unaffected by electoral rules.

8 Discussion

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.

The effect of electoral rules on villagers’ satisfaction is weakly consistent with the model’s predictions. In particular, women are more satisfied with their leaders in villages with at-large elections during the midline survey. There is however no such effect for male villagers. We also do not observe any effect of electoral rules on villager’s satisfaction in the endline survey, but the latter result 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.

According to our theoretical model, electoral rules affect the quality of elected candidates by changing voters’ incentives to support 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 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 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 are looking only at the quality of candidates for which 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 5 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, empirical results suggest that the effect of electoral rules on the quality of elected candidates is not driven by the restriction on candidates’ residence in district elections.

Another potential explanation is that formal education serves as a proxy for the quality of the candidates if the voters do not know the actual quality of the candidates. In district elections the size of the districts is smaller and voters are likely to have better information about the candidates and do not need to rely on formal education as a proxy for quality of candidates, whereas larger district size in at-large elections leads to less information about actual quality of the candidates and more reliance on formal education as a proxy for the quality. This explanation, 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, which would lead to higher levels of education of elected candidates in more heterogeneous villages, which contradicts our empirical findings.

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 and find modest levels of such an effect. Specifically, as the results in Table 6 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. The set of electoral rules has 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 by the male head of household survey participants
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, the results indicate that electoral rules have no significant effect on the incumbency advantage.

In general, difference in the quality of elected representatives can also reflect barriers to entry, which are higher in electoral systems with small district magnitude (Myerson, 1993). However, this explanation is not relevant in our context in which 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 to the best of our knowledge vote-buying did not take place.

9 Conclusion

In this paper we examine the effect of electoral rules on the 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 small district, the local median voter anticipates the legislative bargaining process, which makes him more concerned about policy relative to competence, which 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 geographic size and ethnic composition of these villages, so that in small and ethnically homogenous villages there is no significant effect of the electoral system on the quality of elected officials. There is also some evidence that the satisfaction of citizens with local leaders is higher in villages with at-large elections, but only in the short-run. 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.
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, but rather due to the fact that joint decisions made by legislators indeed give voters incentives to elect biased individuals, which hurts legislators' competence if competence is scarce.

ELECTING ONE LEGISLATOR AT A TIME

In the main model, at-large elections were superior 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. 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 \(a_0 - k \left(\frac{b_0 + B}{2} - \frac{B}{2}\right)^2 = a_0 - k \left(\frac{b_0}{2}\right)^2\)) 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).

More than one competent individual

The results of the paper are of course driven by scarcity of competent individuals: if for any policy position it would be possible to find a competent citizen with such preferences, there would not be a 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 make this case, assume again that in at-large elections, politicians are elected simultaneously (in sequential elections as in the previous subsection, it is clear that the results would hold for any finite number of competent individuals). Suppose that there are two competent individuals in the society. Then at-large elections already have an advantage in that there, both competent citizens may be elected even if they reside in the same district. Consider, however,
the more interesting case, where the two individuals are from different districts.

The analysis in the latter case is going to be more involved because multiple equilibria are possible (in the case of district elections). To see why, suppose that the two competent individuals have bliss points $-\varepsilon$ and $\varepsilon$ for $\varepsilon$ small. Then one can find parameter values with two pure strategy equilibria: where these two competent individuals are elected, and where the two individuals with biased preferences are elected (intuitively, the reason is the median voters’ success or failure to coordinate). This raises the issue of equilibrium selection. But in any case, one can verify that regardless of the way equilibrium is selected in cases where there are multiple, at-large elections dominate according to either of the criteria as in Proposition 4.

**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 3. 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. Consider district elections and suppose that $\sigma$ is an equilibrium. Let $(a^*, b^*)$ be any candidate with a positive chance of election in the right district. Let us prove that he maximizes the payoff of the right district’s median voter, holding the strategies of voters in the left district fixed. Suppose this is not the case, and candidate $(a, b)$ would yield a higher payoff to that median voter. By definition of maximum, this median voter would strictly prefer to elect $(a, b)$ with probability 1 over the equilibrium results. This, however, implies that a majority of voters (all in $[0, B/2 + \varepsilon]$ if $\mathbb{E}b' < b$, all in $[B/2 - \varepsilon, B]$ if $\mathbb{E}b' > b$, and everyone if $\mathbb{E}b' = b$) prefers $(a, b)$ as well, and therefore they have a joint profitable deviation to a strategy where they all vote for $(a, b)$. This is impossible, thus all candidates with a positive chance of election are preferred by the median voter. The same, of course, holds for the left district,
holding the strategies in the right district fixed.

In the text, we have proved that in the district without the competent individual, the district’s median voter’s preferred candidate is the biased one for any voting strategies in the other district. Now, this means that only this individual may be elected in an equilibrium. Similarly, in the district where the competent individual resides, either he or the biased individual will be elected, depending on which one the median voter prefers subject to the other district electing the biased politician. It remains to prove that these outcomes may be supported in an equilibrium. But this is trivial. If only one politician maximizes the median voter’s payoff, it suffices to assume that all individuals in each district vote for the same politician. This is an equilibrium because a successful deviation must include at least half of voters, and the single-crossing condition then requires that it be supported by the median voter, which is impossible. If, in the district with the competent individual, the median voter is indifferent between him and the biased type, then we can have either of the two elected, or we can have a pure strategy equilibrium where both may be elected, with voters more biased than the median voter voting for the biased type and voters less extreme voting for the competent individual. This proves the result for district elections.

For at-large elections, single-crossing considerations similar to those above imply that the pair of candidates elected must be one most preferred by the median voter at 0. In this case, clearly, the most preferred pair consists of the most competent individual and his political antipode. Again, if everyone votes for these candidates, it is an equilibrium. This completes the proof. ■

Proof of Proposition 2. Without loss of generality, assume that the competent individual belongs to the right district. For the median voter of that district with \( b_i = B/2 \), the expected utilities from electing the biased and the competent citizen (taking account that the left district elects the biased candidate \((0, -B)\)) are:

\[
U_{B/2}(0, B) = -k\left(\frac{-B + B}{2} - \frac{B}{2}\right)^2;
\]

\[
U_{B/2}(h, b) = h - k\left(\frac{-B + b}{2} - \frac{B}{2}\right)^2.
\]
Trivial algebra shows that $U_{B/2} (0, B) > U_{B/2} (h, b)$ if and only if $(3B - b) (B - b) > 4k^2$. If the competent individual is in the left district, so $b < 0$, we get an analogous condition. Combining the two yields (3).

The condition (3) is more likely to hold if the right-hand side is smaller ($h$ is small or $k$ is large) or the left-hand side is larger ($B$ is large or $|b|$ is small, given that $|b| \leq B$ and thus both factors are nonnegative). The competent citizen is likely to be elected in the opposite cases. This completes the proof. ■

Proof of Proposition 3. By Proposition 1, in at-large elections, the total competence of legislators is $0 + h = h$. In district elections, it is either $0$ or $h$, and it is zero with a positive probability whenever the condition (3) holds for some $b$. Hence, the total expected competence is larger in at-large elections, and this holds strictly if and only if $3B^2 > 4h/k$. The latter is more likely to hold if $B$ is high or $k$ is high; moreover, once this condition holds, the probability that (3) holds equals $2 - \sqrt{1 + \frac{4h}{B^2k}}$ and thus is increasing in $B$ and $k$. The result for expected competence follows.

To prove the result on polarization, fix the position of the most competent individual $b$. Then the distance from the median voter equals $2|b|$ in the case of at-large elections, and either $|b| + B$ or $2B$ in the case of district elections. Hence, polarization is always strictly higher (in expectation) in district elections. The difference is higher if $(3B - |b|)(B - |b|) > 4k^2$ holds for more values of $b$, i.e., if $B$ is higher or $k$ is higher. This completes the proof. ■

Proof of Proposition 4. Part 1. If the two elected legislators have types $(a_l, b_l)$ and $(a_r, b_r)$, then the social welfare is

$$W = \int_{-B}^{B} \left( a_l + a_r - k \left( x - \frac{b_l + b_r}{2} \right)^2 \right) \frac{1}{2B} \, dx$$

$$= a_l + a_r - k \frac{B^2}{3} - k \left( \frac{b_l + b_r}{2} \right)^2 .$$

In at-large elections, we always have $a_l + a_r = h$ and $b_l + b_r = 0$. In district elections, for all parameter values, $b_l + b_r \neq 0$ with a positive probability, and $a_l + a_r \leq h$. Hence, expected social welfare is higher in at-large elections.
Part 2. This follows from Proposition 3.

Part 3. Following the notation of the proof of Part 1, in at-large elections, \( b_L + b_r = 0 \), the ideal point of the median voter, while in district elections, \( b_L + b_r \neq 0 \) is possible. The voters’ aggregate utility from policy alone is

\[
W_p = \int_{-B}^{B} \left( -k \left( x - \frac{b_L + b_r}{2} \right)^2 \right) \frac{1}{2B} dx = -k \frac{B^2}{3} - k \left( \frac{b_L + b_r}{2} \right)^2 ,
\]

and thus the expectation of \( W_p \) is higher in at-large elections than in district elections. This completes the proof. ■

References


Beath, Andrew, Fotini Christia and Ruben Enikolopov (2012) “Winning Hearts and Minds


Table 1. Balance of Pre-treatment Covariates

<table>
<thead>
<tr>
<th></th>
<th>District</th>
<th>At-large</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many households are in this village in total?</td>
<td>108</td>
<td>108</td>
<td>0.00</td>
</tr>
<tr>
<td>How many people live in this household in total?</td>
<td>9.57</td>
<td>10.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Age</td>
<td>44.03</td>
<td>43.65</td>
<td>0.03</td>
</tr>
<tr>
<td>Do not have formal education</td>
<td>0.73</td>
<td>0.69</td>
<td>0.09</td>
</tr>
<tr>
<td>Mother tongue is Dari</td>
<td>0.72</td>
<td>0.69</td>
<td>0.05</td>
</tr>
<tr>
<td>Never or rarely have problems supplying food</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>0.28</td>
<td>0.26</td>
<td>0.03</td>
</tr>
<tr>
<td>Have access to electricity</td>
<td>0.14</td>
<td>0.15</td>
<td>0.04</td>
</tr>
<tr>
<td>Male health-worker available</td>
<td>0.14</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>Female health-worker available</td>
<td>0.12</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>Have a mobile phone</td>
<td>0.19</td>
<td>0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>Have a radio</td>
<td>0.74</td>
<td>0.76</td>
<td>0.05</td>
</tr>
<tr>
<td>Have sheep</td>
<td>0.56</td>
<td>0.55</td>
<td>0.02</td>
</tr>
<tr>
<td>Total expenditure for food in the last 30 days (AFA)?</td>
<td>3524</td>
<td>3600</td>
<td>0.04</td>
</tr>
<tr>
<td>Received a loan</td>
<td>0.48</td>
<td>0.46</td>
<td>0.05</td>
</tr>
<tr>
<td>People should pay taxes</td>
<td>0.41</td>
<td>0.40</td>
<td>0.02</td>
</tr>
<tr>
<td>Most preferred project is drinking water</td>
<td>0.30</td>
<td>0.28</td>
<td>0.05</td>
</tr>
<tr>
<td>Most preferred project is school</td>
<td>0.18</td>
<td>0.16</td>
<td>0.08</td>
</tr>
<tr>
<td>Most preferred project is road or bridge</td>
<td>0.12</td>
<td>0.16</td>
<td>0.12</td>
</tr>
<tr>
<td>Attended shura meetings</td>
<td>0.33</td>
<td>0.31</td>
<td>0.05</td>
</tr>
<tr>
<td>Women own private land</td>
<td>0.30</td>
<td>0.31</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Standard error</td>
<td>Min</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------</td>
<td>----------------</td>
<td>-----</td>
</tr>
<tr>
<td>Male council member has finished high school</td>
<td>8.7</td>
<td>28.2</td>
<td>0</td>
</tr>
<tr>
<td>Male council member has finished secondary school</td>
<td>16.8</td>
<td>37.4</td>
<td>0</td>
</tr>
<tr>
<td>Ethnically mixed village</td>
<td>0.2</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>Average Distance between villagers</td>
<td>194.0</td>
<td>213.1</td>
<td>16.6</td>
</tr>
<tr>
<td>Average Distance between council members</td>
<td>472.5</td>
<td>525.4</td>
<td>0</td>
</tr>
<tr>
<td>Respondents Disagreed With Decision or Action of Village Leaders (male respondent)</td>
<td>7.9</td>
<td>26.9</td>
<td>0</td>
</tr>
<tr>
<td>Respondents Disagreed With Decision or Action of Village Leaders (female respondent)</td>
<td>9.8</td>
<td>29.8</td>
<td>0</td>
</tr>
<tr>
<td>Respondent Attributes Positive Change in Economic Situation to Village Leaders (male respondent)</td>
<td>3.0</td>
<td>17.0</td>
<td>0</td>
</tr>
<tr>
<td>Respondent Attributes Positive Change in Economic Situation to Village Leaders (female respondent)</td>
<td>2.0</td>
<td>13.9</td>
<td>0</td>
</tr>
<tr>
<td>Respondent is Satisfied with Work of Village Leaders (male respondent)</td>
<td>82.7</td>
<td>37.8</td>
<td>0</td>
</tr>
<tr>
<td>Respondent is Satisfied with Work of Village Leaders (female respondent)</td>
<td>72.2</td>
<td>44.8</td>
<td>0</td>
</tr>
<tr>
<td>Household’s Economic Situation Has Improved in Past 12 Months (male respondent)</td>
<td>29.1</td>
<td>45.4</td>
<td>0</td>
</tr>
<tr>
<td>Household’s Economic Situation Has Improved in Past 12 Months (female respondent)</td>
<td>28.1</td>
<td>45.0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3. 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>Large Village* At-large elections</td>
<td>9.96***</td>
</tr>
<tr>
<td></td>
<td>[3.01]</td>
</tr>
<tr>
<td>Large Village</td>
<td>-13.32***</td>
</tr>
<tr>
<td></td>
<td>[4.38]</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>-11.23**</td>
</tr>
<tr>
<td></td>
<td>[4.78]</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: Male council members only. 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 Voter Satisfaction

<table>
<thead>
<tr>
<th></th>
<th>Respondent Disagreed With Decision or Action of Village Leaders</th>
<th>Respondent Attributes Positive Change in Economic Situation to Village Leaders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male Respondents</td>
<td>Female Respondents</td>
</tr>
<tr>
<td>At-large *Midline Survey</td>
<td>1.76</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>[1.298]</td>
<td>[1.634]</td>
</tr>
<tr>
<td>At-large *Endline Survey</td>
<td>0.25</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>[1.385]</td>
<td>[1.640]</td>
</tr>
<tr>
<td>Quadruple* Survey fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>4,494</td>
<td>3,886</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.10</td>
<td>0.10</td>
</tr>
</tbody>
</table>

|                                | Respondent is Satisfied with Work of Village Leaders | Household's Economic Situation Has Improved in Past 12 Months |
|                                | Male Respondents | Female Respondents | Male Respondents | Female Respondents |
| At-large *Midline Survey        | -1.87            | 0.39               | -0.78            | 4.45*             |
|                                | [1.834]          | [2.664]            | [2.473]          | [2.517]           |
| At-large *Endline Survey        | -0.83            | -3.67              | -1.18            | -1.59             |
|                                | [2.235]          | [2.567]            | [1.705]          | [1.949]           |
| Quadruple* Survey fixed effects | Yes              | Yes                | Yes              | Yes              |
| Observations                   | 4,270            | 3,931              | 4,493            | 3,997            |
| R-squared                      | 0.15             | 0.13               | 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%.
<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>4.04***</td>
</tr>
<tr>
<td></td>
<td>[1.43]</td>
</tr>
<tr>
<td>Ethnically Mixed Village*</td>
<td>6.11*</td>
</tr>
<tr>
<td>At-large elections</td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village</td>
<td>-8.17*</td>
</tr>
<tr>
<td></td>
<td>[4.49]</td>
</tr>
<tr>
<td>Large Village*</td>
<td></td>
</tr>
<tr>
<td>At-large elections</td>
<td></td>
</tr>
<tr>
<td>Large Village</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadruple fixed effects</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,716</td>
</tr>
<tr>
<td>R-squared</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 6. 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 A1. Effect of Electoral Rules on Educational Attainment of Council Members

<table>
<thead>
<tr>
<th></th>
<th>Council Member Has Finished Secondary school</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>At-large elections</td>
<td>3.43* (1.93)</td>
</tr>
<tr>
<td>Large Village* At-large elections</td>
<td>11.33** (5.49)</td>
</tr>
<tr>
<td>Large Village</td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village* At-large elections</td>
<td></td>
</tr>
<tr>
<td>Ethnically Mixed Village</td>
<td></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>0.19</td>
</tr>
</tbody>
</table>

Note: Male council members only. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.