Electoral Rules and Political Selection:

Theory and Evidence from a Field Experiment in Afghanistan[*]

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

November 10, 2014

Abstract. Voters commonly face a choice between competent candidates and those with policy preferences similar to their own. The paper explores how electoral rules, such as district magnitude, mediate this trade-off and affect the composition of representative bodies and policy outcomes. We show formally that anticipation of bargaining over policy causes voters in elections with multiple single-member districts to prefer candidates with polarized policy positions over more competent candidates. Results from a unique field experiment in Afghanistan are consistent with these predictions. Specifically, representatives elected in elections with a single multi-member district are better educated and exhibit less extreme policy preferences.

Keywords: electoral rules, political selection, district magnitude, quality of politicians, competence, polarization, local public good, legislative bargaining, field experiment, Afghanistan.

JEL Codes: D72, D78.

---

1 Introduction

Ideally, elections for representative bodies result in the selection of those candidates that both perfectly represent the preferences of the electorate and design effective policies. In the real world, however, candidates with the qualities necessary to design effective policies are scarce. As a result, voters face a trade-off between voting for candidates with policy positions closest to their own and candidates who are intrinsically more capable of designing effective policies. Whether voters choose more qualified candidates over those who ascribe to particular policy positions impacts the composition of representative bodies and, by extension, the representativeness and quality of public policy. While the trade-off between qualified candidates and those who best represent the wishes of the electorate is familiar to almost all who have voted, there is less appreciation of the role potentially served by electoral rules in mediating this trade-off and thereby producing differences in the composition of representative bodies and in the quality of public policy.

This paper seeks to identify the effects of differences in electoral systems – and district magnitude, in particular – on voter behavior, political selection, and policy outcomes. While numerous studies have posited links between electoral rules and policy outcomes (see Cox, 1997; Norris, 2004; Persson and Tabellini, 2000, 2005), exploration of the intermediating effects of voter incentives, political selection, and the qualities of elected representatives have been limited.[1] Moreover, difficulties in identifying exogenous sources of variation in electoral systems have generally precluded causal inferences in empirical studies, preventing the consolidation of insights of how electoral rules can affect policy outcomes. This paper overcomes these two issues, making a theoretical and empirical contribution to the understanding of the effects of differences in electoral rules and of the determinants of differences in policy outcomes. Specifically, the paper combines models of voter choice and legislative bargaining to generate predictions of how differences in district magnitude affect political selection and policy outcomes. These predictions are then tested using data from a unique experiment that induced randomized variation in electoral rules for local council elections across 250 villages in Afghanistan.

The paper focuses on two alternate and commonly used electoral rules: the district election

---

[1] The importance of political selection, though emphasized as equally important as political accountability as early as the Federalist papers (Madison, 1788), has been largely "neglected" (see Besley, 2005, p. 44). Theoretical work by Myerson (1993) provides a notable exception.

system and the at-large election system. District elections, which are used to elect the U.S. House of Representatives and the British House of Commons, are 'single-winner systems' whereby successful candidates must attain at least a plurality of votes cast in their assigned electorate. District elections thus necessitate the division of the area to be represented by the body into multiple geographic constituencies, with voters confined to casting votes for candidates competing to represent the voter's constituency. At-large elections systems, which are a common means by which parliaments are elected in continental Europe, broaden the scope of the constituency over which candidates are elected. As a result, multiple candidates are elected by each constituency. In the extreme, an at-large election may contain a single constituency spanning the area represented by the particular body (e.g., an entire country in the case of a national parliament).

The theoretical model developed in the paper examines how these two different electoral rules affect the composition of a village council. The model considers a linear village with a uniform distribution of villagers. The location of a villager's home corresponds to her ideal point for the location of a public good (e.g., a drinking well). The location of the public good is to be decided by a two-member council, elected by village residents, through a bargaining process. In line with the experiment, we consider a citizen-candidate model with no entry costs for the candidates, meaning that all residents are considered candidates. Under at-large elections, each resident has two votes and can vote for any two candidates.[2] Under district elections, the village is split into two geographically-defined districts and each citizen can only vote for a candidate residing in the same district. The number of residents capable of making effective policy decisions is assumed to be small and voters thereby trade off candidates' policy preferences with candidates' competence.[3] The model predicts that, in district elections, voters prefer candidates with more extreme policy preferences over more competent candidates, as they expect such candidates to achieve a better outcome in the bargaining game with candidates elected from other districts. In at-large elections, these strategic considerations are less pronounced, so voters are more willing to elect competent

---

[2]This is in line with Cox (1984), who was the first to study double member districts formally. The key difference is that in Cox (1984), voters vote sincerely and candidates choose positions strategically. In this paper, voters are strategic, but candidates' positions are fixed and tied to their location of residence.

[3]In a richer model, voters would face other trade-offs as well. For example, van Weelden (2013) demonstrates a trade-off between aggregating preferences (choosing politicians with preferences of the median voter) and providing them with incentives to exert effort. There, the median voter prefers politicians with biased preferences, because they may be effectively punished without hurting the median voter by electing their polar opposites. Our focus is purely on political selection.

candidates. The difference between the two electoral systems is magnified in villages with more heterogeneous policy preferences.

The predictions of the theoretical model are tested with data from a field experiment covering 250 villages across Afghanistan. The experiment induced randomized variation in the electoral rules governing the composition of village councils that were mandated to make decisions on the type and location of a package of local public goods. Of the 250 villages, 125 were randomly selected to compose councils by district elections. Per this procedure, villages were each divided into several single-member districts, with candidates elected from each district separately and exclusively by villagers residing in that same district. The other 125 villages were assigned to the at-large election system. Per this procedure, each village constituted one multi-member district, with villagers facing no restriction on which candidates in the village they could vote for and council members accordingly elected based on the number of votes garnered across the whole village. Under both electoral rules, all villagers were automatically considered candidates.

The results of the field experiment are consistent with the predictions of the theoretical model. Specifically, we find that at-large elections result in the election of more capable council members, as proxied by their level of education.[4] The effect is strong in heterogeneous villages (as measured by the divergence of villagers' ex-ante policy preferences, the geographic size of villages, and ethnic composition) and is absent in homogenous villages. We also find evidence that district elections result in the election of council members with more biased preferences over the location of local public goods, as proxied by the location of their houses. We show that the results are not consistent with a number of alternative explanations, such as restrictions on the number of qualified candidates that can be elected from the same district or by differences in incumbency advantage.

This paper contributes to an extensive literature analyzing the effects of electoral systems in general and district magnitude in particular. Previous studies have observed that the number of 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

---

[4]Education as a proxy for the quality of politicians has been previously used in Besley, Pande, and Rao (2005), Galasso and Nannicini (2011), Gagliarducci and Nannicini (2013), Folke, Persson, and Rickne (2014).

4

minority (Trebbi, Aghion, and Alesina, 2007). Studies generally find that single-member district systems result in better representation of geographically-concentrated interests, with high threshold proportional representation systems favoring geographically-dispersed interests (Ferree, Powell, and Scheiner, 2013).[5]

The determinants of the capability of elected representatives and government agents have invited extensive theoretical investigation (McKelvey and Reizman, 1992; Banks and Sundaram, 1998; Aragones and Palfrey, 2004; Caselli and Morelli, 2004; Mattozzi and Merlo, 2007; Acemoglu, Egorov, and Sonin, 2010; Egorov and Sonin, 2011). Empirical investigations have further established links between more capable elected officials and higher levels of intra-party competition (Besley, Folke, Persson, and Rickne, 2013; Folke, Persson, and Rickne, 2014), higher inter-party competition (Banerjee and Pande, 2007), higher wages (Ferraz and Finan, 2011; Dal Bó, Finan, and Rossi, 2013; Gagliarducci and Nannicini, 2013), smaller budgets (Brollo, Nannicini, Perotti, and Tabellini, 2013), and more democratic institutions (Besley and Reynal-Querol, 2011). However, only Myerson (1993) considers how the quality of elected representatives is affected by the type of electoral system, arguing that, when voting is strategic, small district magnitude increases the barriers to candidate entry, which has a negative effect on the capability.[6]

To the best of our knowledge, this paper is the first to provide experimental evidence on the causal effects of electoral rules. However, as with much other experimental work, internal validity is achieved within a specific setting, which raises the issue of the external validity of the results. A number of factors, though, suggest that this study's conclusions have broader applicability. First, the trade-off described in the theoretical model is general, so the predictions of the model are applicable to other representative bodies, both local and national, worldwide.[7] Second, the

---

[5] Persson and Tabellini (1999, 2000) and Persson, Tabellini, and Trebbi (2003) argue that proportional systems with high district magnitude lead to higher levels of political rent extraction. Such systems also favor bigger governments and higher levels of redistribution (Iversen and Soskice, 2006; Persson and Tabellini, 2004). Other works that look at the effect of electoral rules on the composition of government spending include Lizzeri and Persico (2001) and Milesi-Ferretti, Perotti and Rostagno (2002).

[6] See also Adams (1996), who considers the effect of a constitutional change in Illinois in 1980 on business friendliness of state legislators.

[7] For example, in the US context, our model would suggest that the electoral success of the Tea Party movement in Republican primaries and general elections in 2010-2012 need not imply that voters became biased to the right or more polarized. Instead, following the global financial crisis, deficit spending increased, and the salience of the taxation-deficit-spending issue increased. In this context, voters who are only moderately averse to deficit spending may have decided to elect representatives who are extremely averse to increasing the deficit, anticipating that their more extreme position will lead, in equilibrium, to a more favorable compromise, such as lower deficits. Though always present, this effect became more pronounced as the salience of the issue increased. This interpretation is consistent with findings in Ansolabehere, Rodden, and Snyder (2006), which suggests that, in the past three decades,

5

absence from the experiment of long-term considerations, such as over re-election, reputation, or accountability, and pre-election politicking, such as candidate selection, primaries, or electoral campaigns, enables identification of the effects of electoral rules on political selection and the exploration of mechanisms underlying such effects in a precise manner that would not be possible in a more complex political system.

Third, in contrast to many field experiments conducted in a single region of a country, the 250 villages that form the sample are drawn from five different regions of Afghanistan, which provides for a diverse set of contextual conditions that span numerous ethnicities, levels of social capital, economic structures, and openness to democratic practices. This suggests that our findings are not driven by specific local conditions of a particular region and are generalizable to other contexts. Finally, field experiments in Afghanistan have already been used in the literature to study corruption (Callen and Long, forthcoming), determinants of risk preferences (Callen, Sprenger, Isaqzadeh, and Long, 2014), the effect of school construction (Burde and Linden, 2013), and women's empowerment (Beath, Christia and Enikolopov, 2013). More generally, this paper is part of a growing literature that studies different countries of the world in detail in order to derive general implications.[8]

The rest of the paper is organized as follows: Section 2 presents the experimental design; Section 3 describes the theoretical model; Section 4 formulates empirical predictions from the model; Section 5 describes the data; Section 6 details the empirical results; Section 7 discusses the empirical and theoretical findings; and Section 8 concludes. The paper contains three web appendices: Appendix A contains extensions of the baseline model and explores the robustness of its predictions; Appendix B contains all the proofs; and Appendix C contains additional empirical results.

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 250 villages in Afghanistan. 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,

---

growing polarization of politicians has occurred despite stable electorate preferences.

[8]For example, see Olken (2007), Barron and Olken (2009), Martinez-Bravo (2014a) on Indonesia; Ferraz and Finan (2008, 2011) on Brazil; Dell (2010) on Peru; Casey, Glennerster, and Miguel (2012) and Acemoglu, Reed, and Robinson on Sierra Leone, among others.

Christia and Enikolopov, 2013b) and 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 induced across the 250 villages (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 National Solidarity Program

The National Solidarity Program (NSP) 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 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 (hereafter, "council"); and (ii) the disbursement of block grants to councils for implementation of village projects. The program is executed by the Ministry of Rural Rehabilitation and Development, facilitated by contracted NGOs, and funded by bilateral and multilateral donors.

NSP mandates the creation of gender-balanced village development councils through a secret-ballot, universal suffrage election.[9] 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,[10] 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 and 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 then oversee project implementation.

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

---

[9]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)

[10]The average block grant in the villages included in the sample was approximately $31,000.

elections for the council is also uncertain. Per NSP rules, villages are supposed to hold re-elections for council positions every four years, but since follow-up elections are not facilitated, it is unclear whether these actually occur.[11]

2.2 Local Governance in Afghanistan

Afghanistan's central government has historically lacked the resources to exercise local control or provide public goods in many parts of the country. As a result, local communities developed customary 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* or *jirga* members tend to be the elders of the village (Rahmani, 2006), although membership is ordinarily not fixed. *Shuras* or *jirgas* generally convene when there is an issue to resolve and reach decisions based on consensus (Boesen, 2004). In addition to *shuras* or *jirgas*, 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 generally be hidden from public observation – precludes female involvement in communal gatherings and thus from local governance.

Elected development councils are substantially different from customary governance institutions in the mode of selection and the respective accountability structure. While elected development 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 development councils, their authority in selecting, implementing, and managing NSP-funded projects provides them with control over what is, for many villages, an unprecedented volume of resources. Thus, the elected development council exists in parallel with customary governance structures as an institution vested with substantial authority.

---

[11]No such re-elections had occurred by early 2012, when the data collection for this evaluation was completed.

2.3 Electoral Rules

Every village resident, whether male or female, age eighteen or older, who has lived in the village for at least one year was eligible to vote and be elected to the council. NSP rules required that at least 60 percent of eligible voters must cast votes in the election for it to be valid. All eligible villagers were considered candidates and people voted by writing-in the name(s) of the candidates.[12] Villagers interested in being elected to the council were prohibited from campaigning in any way for the position. The council had 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 were segmented into geographically contiguous districts containing between 5 and 25 families, with each district having its own polling station. A village map with districts and enclosed dwellings was displayed in a public area in the village.

Council elections were conducted by secret-ballot according to one of two sets of electoral rules that differ primarily in terms of district magnitude:[13]

District Election: Voters were restricted to casting a ballot for a single candidate, who had to reside in the same district as the voter. In each district, 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-member, simple plurality election with multiple districts (Cox, 1997), similar to the first-past-the-post system.

At-large Election: Under this method, voters could cast their ballot for anyone residing in the village. The men and women receiving the most votes across the village were elected as council members. Voters could cast ballots for a maximum of three different people, who were not ranked.[14] 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: (i) higher district magnitude (multiple elected members instead of one) and (iii) number of votes cast (three instead of one).

---

[12] Illiterate villages could ask NGO workers to write the name for them. This arrangement could potentially violate the secrecy of the vote. However, the results of election monitoring (see below) indicate that 95 percent of voters were confident that their vote would remain secret.

[13] A detailed guide on the procedures is available at: http://www.nsp-ie.org/sti.html

[14] This means the system allows plumping, but not cummulation (Cox, 1997). Participating NGOs requested the casting of up to three votes be allowed in at-large elections as they 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 logistically feasible.

In all villages, council elections were organized and administered by "social organizers" employed by the contracted NGOs facilitating NSP. 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 contracted NGOs and that, in general, villagers exhibited a good understanding of the different electoral rules.[15] Monitoring results thus confirm high levels of compliance with the assigned treatment status.

2.4   Sample and Randomization

The randomization of electoral rules occurred in 250 villages that formed the treatment group for the randomized impact evaluation of NSP and thus were assigned to receive NSP. The 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.

The average population in our sample of villages is roughly one thousand people (see Table 1). There is notable variation in the geographic size of villages, with quite a few villages spanning 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 Pashtun, 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 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. The sample villages are also very poor: only forty-five percent of respondents indicate that they never or rarely have problems supplying food for their families.[16]

---

[15] A detailed description of the monitoring results can be found at: http://www.nsp-ie.org/reports/CDCE-MR.pdf

[16] An assessment of the demographic and economic characteristics of the sample villages reveals few substantive differences with those of a random sample of villages surveyed by the 2007-08 National Risk and Vulnerability Assessment.
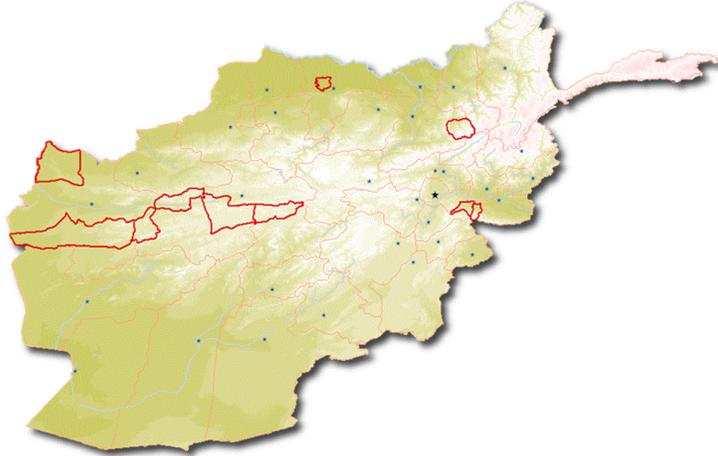
Figure 1: Ten sample districts.

Random assignment of electoral rules was made concurrently with the assignment of 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.[17] 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 the four combinations of council election rules and project selection procedures.[18]

This assignment procedure essentially represents randomization with stratification by quadruples. This 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 1 presents a comparison between the two groups of villages with regard to a number of pre-intervention characteristics,

---

[17]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). 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.

[18]Development projects were selected either through a secret-ballot referendum or at a village meeting. In either case, the village council was responsible for preparing a list of proposed projects and for implementing the selected project (see Beath, Christia and Enikolopov, 2013b, for more information). The project selection procedure assigned to a village was known to the NGOs. We check that there are no statistically significant differences on the effect of electoral rules between villages with different project selection procedures.

which shows that the differences between the two groups never exceed 13 percent of a 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 occurred between October 2007 and May 2008. Elections in 131 out of 250 villages were monitored, providing additional data on process. Project selection occurred between November 2007 and August 2008, whereas project implementation occurred between April 2008 and September 2011. The first follow-up survey (midline) was conducted between May and October of 2009. The second follow-up survey (endline) was conducted between May and October of 2011, at which time 99 percent of projects funded by NSP were complete.

3 Theory

In this Section, we introduce a simple model of elections where competent candidates are rare to study voters' choice between electing the most competent candidate versus voting for someone whose political positions they share. The model is deliberately simple, but in Appendix A, we consider multiple extensions on how electoral rules can affect the competence and political positions of elected candidates and show that our main findings are robust. Some features of the model, in particular the absence of entry costs for candidates or the absence of parties), were inspired by the setting of the experiment. However, the model was not backward engineered to match the empirical findings, and all the predictions except for 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 the 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 $u(p, b)$, 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: $a \in \{0, h\}$, with $h > 0$. 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 $a_i = h$. The results are most transparent when $N = 1$, which we assume for the rest of the paper; in Appendix A (Subsection A1) we show that the results go through if we allow N to be any number. We assume that the types of all individuals are known to all other individuals.[19] 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 in Subsection A4), 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 regardless of the policy. 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 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

[19] Given the context of the experiment, it is natural to assume that location of villagers' dwellings and education are observable to fellow villagers. But the theory is applicable in environments where political positions and/or competence are not readily observable by voters, as long as candidates are able to signal their political positions or competence prior to elections.

offers (Rubinstein, 1982) or a legislative bargaining game with random recognition (Baron and Ferejohn, 1989; Banks and Duggan, 2000) in the limit where offers are made very frequently.[20] We consider a bargaining game explicitly when we generalize the game to incorporate councils with more than two members and show that the results are robust to alternative assumptions regarding the bargaining process (see Appendix A). Here, slightly abusing notation, we assume that having a council with members (a_l, b_l) and (a_r, b_r) yields utility:

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

We compare two electoral procedures: at-large elections and district elections. In district elections, the society is divided into two districts: left district L, containing individuals with $b_i < 0$, and right district R, containing individuals with $b_i \geq 0$. This division is made according to the location of an individual's residence, so that 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, that is, every individual living in the district is considered a candidate.[21] Then, in each district, the individual who received 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. The two candidates who received the most votes are elected.[22] This setup assumes a homogenous population with no gender differentiation; in Appendix A (Subsection A7), we show that the predictions of the model hold if the council is formed from two populations (e.g. men and women) that have the same geographic distribution, which is more in line with the experiment.

---

[20]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 an explicit bargaining game in Appendix A (Subsection A2).

[21]Our model of elections thus falls into the category of citizen-candidate models with costless entry (Besley and Coate, 1998; see also Osborne and Slivinski, 1996, and Besley and Coate, 1997).

[22]We will prove that in at-large elections there is a Condorcet winner, that is, a pair of citizens such that there is no other pair that a majority would want to elect. The voting model we selected ensures that the Condorcet winner pair is elected. The assumption that voters cannot cast both their votes for the same candidate is important. If voters could cast both votes for the same candidate, there may be a continuum of election outcomes even if coalitional deviations are allowed, as in Definition 1 (we show that in Appendix B). We would run into the same problem if each citizen had only one vote. However, if citizens could cast more than two votes, or if they could cast votes for pairs of candidates rather than individuals, the same Condorcet winner pair would always be elected. This is an interesting observation per se: because citizens' preferences in at-large elections are defined over pairs of politicians, voting procedures that facilitate extracting this information are more successful in aggregating these preferences. (In district elections, each district has a standard majority voting rule and for any choice by the other district it elects the Condorcet winner.)

The strategy of each voter in district elections is therefore $\lambda(i)$, the identity of the individual in his district for whom he casts his vote (since only members of the same district may be elected, $b_{\lambda(i)} \in [-B, 0)$ if $b_i \in [-B, 0)$ and $b_{\lambda(i)} \in [0, B]$ if $b_i \in [0, B]$). The strategy of each voter in at-large elections is $\Lambda(i) = (\lambda_1(i), \lambda_2(i))$, 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:

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. For at-large elections, our equilibrium concept coincides with the Strong Nash.

3.2 Analysis

Analysis of the game is greatly simplified by the fact that the median voter theorem applies in both at-large and district elections. 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/she 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 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 (see Appendix B) 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 (h, q) elects either this citizen 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 which types of citizens get elected, consider at-large elections first. The median voter in the whole district, m_S, has the bliss point $b_{m_S} = 0$ and his ideal outcome is to elect two council members who negotiate and implement his ideal policy 0, while 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.

Next, 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) and 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:

$$w_{m_R}(a_l, b_l, a_r, b_r) = a_l + a_r - k \left(\frac{b_l + b_r}{2} - \frac{B}{2} \right)^2. \tag{1}$$

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 the citizens in the left district.[23] We thus see the following strategic delegation effect: even though the median voter m_R likes policies which are close to m_R, he/she prefers to elect an extreme council member because he/she would negotiate a better policy; this is similar to delegating bargaining to a more

---

[23] 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.

committed type.[24] 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:*

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

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 a 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 $\frac{h}{k} < \frac{3}{4}B^2$) and this difference is increasing in B and k;*

_____

[24]This effect is similar to strategic polarization in the case of divided government in Alesina and Rosenthal (2000) and in the case of party competition in Ortuño-Ortín (1997).

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, while in district elections, competence and distance from the median voter are negatively correlated.*

These results follow from Propositions 1 and 2 and are shown on Figure 2, which depicts 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 polarization result 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 the competence of a council member: as one can see from Figure 2, 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.[25]

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

---

[25]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 A2 shows that at-large elections lead to a worse outcome than district elections if offers are made infrequently and Subsection A6 shows that at-large result in more competent council members even if one member is elected at a time. The results are entirely driven by the bargaining over policies by the council members, which is anticipated by voters at the time of elections.
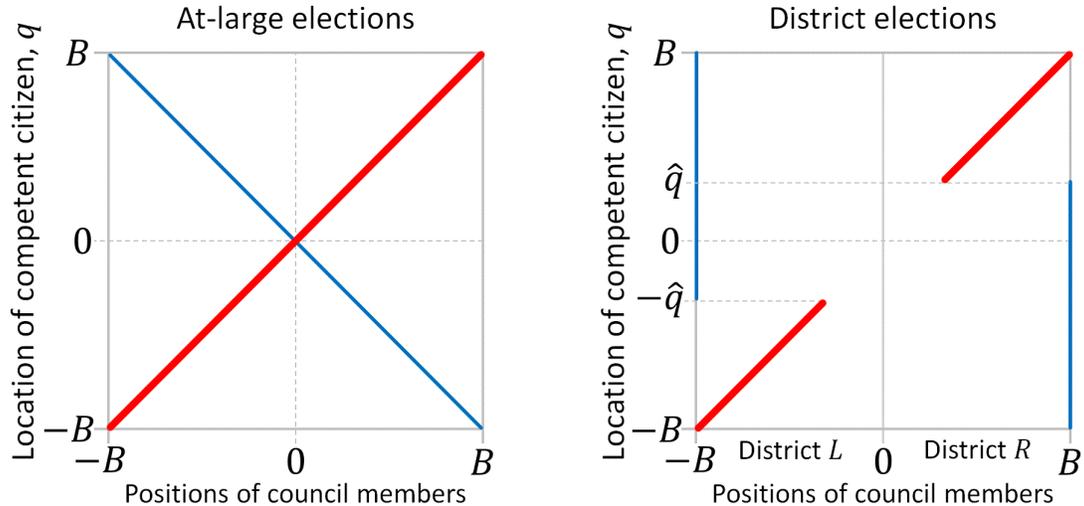
Figure 2: Types of council members (competent in red/thick, incompetent in blue/thin) as a function of the competent citizen's ideal point.

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 less certain and runs the risk of electing two incompetent citizens; both effects hurt every citizen equally.

4   Empirical Predictions

The theoretical model generates several empirical predictions, which we test using data from the field experiment. It should be emphasized that only empirical results for the first prediction were obtained before the model was formulated. All the remaining empirical predictions were first obtained from the model and only then tested empirically.

Based on the first statement of Proposition 3, we formulate the following empirical prediction:

1. The competence of elected candidates is higher in at-large elections than in district elections.

Following Besley, Pande, and Rao (2005), Galasso and Nannicini (2011), Gagliarducci and Nannicini (2013), and Folke, Persson, and Rickne (2014), we use educational attainment as a measure of candidate competence, since there is evidence that the leader's educational level has a positive effect on governance outcomes. In particular, better educated public officials are less likely

to use power opportunistically (Besley, Pande, and Rao, 2005), are more likely to promote higher economic growth (Besley, Montalvo, and Reynal-Querol, 2010; Congleton and Zhang, 2013), and are more likely to provide higher quality public goods (Martinez-Bravo, 2014b). We also check the robustness of our results to using occupation as a measure of candidate competence.[26]

The first part of Proposition 3 also asserts that the effect of the electoral system on the competence of politicians is stronger in communities with more diverse preferences (higher B) or where the choice of policy is relatively more important than the politicians' competence (higher k), whereas for communities with sufficiently homogenous preferences (where $\frac{3}{4}B^2 < \frac{h}{k}$) there should be no differences in the competence of candidates elected by the two electoral systems. Thus, we formulate the following two empirical predictions:

2. The difference in the competence of council members between at-large and district elections is higher in more heterogeneous villages.

3. In homogenous villages, the competence of council members does not depend on the type of elections.

In the empirical analysis, we use three alternative measures of heterogeneity: fractionalization of preferences over projects; ethnic heterogeneity; and geographic size of villages. These measures were chosen due to the nature of tasks performed by council members. As noted above, the main tasks of the elected council members were to guide the choice over development projects and then oversee the implementation of these projects. Thus, 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 and corresponds to higher B in the model. In addition, if preferences are diverse, policy choice becomes more important relative to the quality of implementation, so fractionalized preferences correspond to the case where policy is important relative to competence – that is, high k. Ethnic heterogeneity is often used in the literature to capture differences in tastes

---

[26]Other potential measures of candidate competence such as age or being a member of the pre-existing elite are not well-suited for the Afghan context, in which capture of the local institutions by pre-existing elites (including elders) is a serious concern.

(e.g. Alesina, Baqir, and Easterly, 1999) that might affect preferences over both project type[27] and location (if villages are ethnically segregated),[28] so ethnically heterogeneous villages are also likely to have higher B and k. Finally, preferences over project location are likely to be driven by the location of voters and candidates. The larger the village, the more diverse 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).

The second statement of Proposition 3 suggests that expected polarization is lower in at-large elections, leading to the following prediction:

4. Elected council members exhibit less biased policy preferences 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 project location, taking the assumption that the geographic location of elected candidates and voters reflects their policy preferences. Thus, we use the distance between 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.

The third statement of Proposition 3 corresponds to the following empirical prediction:

5. In district elections, there is a negative correlation between the competence of an elected representative and the bias of his preferences, while there is no such relationship in at-large elections.

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

According to Proposition 4, social welfare is higher in at-large elections as a result of the election of more qualified candidates. Although we do not have measures of social welfare, we can examine

---

[27]This 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).

[28]Unfortunately, the number of observations within villages is not large enough to test for the existence of segregation directly.

21

the effect of electoral rules on outcomes that are likely to be affected by the competence of council members and could serve as proxies for the welfare of villagers. Since the main task of the elected council is to oversee project implementation, one of the outcomes that could be affected by the competence of elected candidates is the speed of project implementation. In particular, council members of high competence would be expected to implement projects faster. Thus, we formulate the following empirical prediction:

6. Projects are implemented faster in villages that held at-large elections as compared to villages with district elections.

In particular, we look at the effect on two measures that are likely to affect the welfare of villagers: time to project start (which corresponds to the period of time until resources start flowing into the village) and time to project completion (which corresponds to the period of time when villagers can start enjoying the benefits of these projects).

Candidate quality might also affect the quality of project implementation and, as a result, the benefits that are delivered by the projects (access to clean water, electricity, irrigation etc.). Successful project implementation, in turn, should also improve the overall economic welfare and have a positive effect on the attitudes of the population towards their local leaders. Thus, we can formulate an additional empirical prediction:

7. Effectiveness of development projects, improvements in general economic welfare, and increases in support for local leaders and democratic processes are all higher in villages that held at-large elections as compared to villages with district elections.

To examine the effect of the electoral system on the effectiveness of development projects, we exploit data on access to clean drinking water, access to electricity, and agricultural productivity. These three indicators measure impacts of the three projects most commonly funded by NSP in the sample: drinking water wells, electrical generators, and irrigation facilities.[29] Measures of general economic welfare are provided by measures of subjective economic welfare, as well as measures of household income and expenditure. Finally, we look at villager attitudes toward local leaders and their support for democratic elections as a way of selecting public officials.[30]

---

[29] The fourth most popular type of project is the renovation of roads and bridges within the village, but unfortunately we do not have good measures to capture the benefits delivered by this type of project.

[30] Since the village council elections under study were the first local elections experienced by these villagers, we

5 Data

Indicators of educational attainment of council members, which are used to assess the competence of elected representatives, are given by three dummy variables indicating whether a member finished high school, middle school, or primary school. The data on members' educational attainment, as well as their gender, age, occupation, and district of residence and the election results for each village, was provided for 2,044 men and 2,015 women elected to councils for 241 villages by NGOs contracted to facilitate NSP in the 10 sample districts.[31] Overall, the level of education of council members is very low. Among male council members, only 9 percent had finished high school, 17 percent had finished middle school, and 33 percent had finished primary school (see Table 2). Among female representatives, only 0.8 percent had finished high school and less than 10 percent had finished primary school.

We use the distance between the location of council members' houses and village centers as a proxy for policy bias of elected representatives. To construct this measure, we first compute the location of the village center by averaging GPS coordinates of the houses of the baseline survey respondents. Then, we compute the distance between the village center and the locations of council members' houses. Information the locations of council members' houses is available only for a subset of villages, since some monitors failed to collect coordinates. [32] For cultural reasons, information was not collected on the location of the residences of female council members. The data provides coordinates of the residences of 1,104 male council members in 140 communities in nine out of ten evaluation districts. Although much data is missing, it is not correlated with the type of the election assigned to the village.

Measures of preference fractionalization and the size and ethnic heterogeneity of villages are derived from data collected by the baseline survey, which was administered in September 2007 to

---

expect satisfaction with particular candidates elected to the council to be readily translated into attitudes toward elections more generally.

[31]Of the 9 villages for which 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.

[32]In addition, some coordinates contained obvious mistakes and were excluded from the analysis. 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. We winsorize the distance at the 95 percent quintile to limit the influence of potential outliers.

2,387 randomly-selected male heads-of-household in the 250 sample villages. Data on preference fractionalization is given by responses to a question whereby respondents were asked to indicate their preference from a list of fifteen potential projects commonly-funded by NSP. 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 $f_v = 1 - \sum s_{vj}^2$. An indicator of preference fractionalization is then given by a dummy variable, which equals one if fractionalization is above the median in the sample and zero otherwise. An indicator of village size is given by a dummy variable which indicates if the average distance between respondents' households is above the sample median. Ethic heterogeneity of villages is given by a dummy variable that equals zero if all villagers belong to the same ethnicity.[33]

A measure of the speed of project implementation is derived from NSP administrative records, which indicate the type, budget, and start and end dates of 478 projects funded by NSP.[34] Project start dates range from April 2008 to February 2011 and completion dates range from July 2008 to September 2011. Two indicators are constructed, which measure the number of days between election of a council and the start of a project and the number of days between election of a council and the end of a project.[35]

Measures of the effectiveness of projects, of improvements in general economic welfare, and of increases in support for local leaders and democratic processes are derived from data collected by the midline and endline surveys. The midline survey was administered to 2,367 male and 2,141 female respondents between May and October 2009 and the endline survey was administered to 2,130 male and 1,858 female respondents between May and October 2011.[36] Project effectiveness is assessed by

[33]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.

[34]Six projects failed and were not terminated before completion. For some projects the time between the start and the end of the project was unrealistically high, so we trimmed all the observations for which the duration of the project implementation was above the 95th percentile. There is no significant difference in the number of trimmed observations between the two types of electoral rules.

[35]The dates of the elections do not depend on the type of electoral rules. The results are similar, but less precise if we look at the number of days from a specific date, common to all villages.

[36]Because of the deterioration in security conditions, we were not able to conduct surveys of male heads of household in 11 villages at midline and in 32 villages at endline. For female heads of household the surveys were not conducted in 33 villages at midline and in 59 villages at endline. 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 cases at midline and 44 percent of cases at endline. The predominant reason for enumerators not being able to interview baseline

five indicators: (i) number of seasons in past year primary source of drinking water was perceived to be contaminated; (ii) incidence of diarrhea among children in the past two weeks; (iii) hours of electricity available in past month; (iv) yield of most recent harvest; and (v) revenue from most recent harvest.[37] Indicators of improvements in general economic welfare are given by: (i) a dummy for whether the respondent perceived an improvement in their economic situation in the past year; (ii) a dummy for whether the respondent attributes positive economic change to local leaders and/or the council; (iii) annual household income; and (iv) annual household expenditure.[38] Measures of support for local leaders and democratic processes are given by: (i) whether respondent expressed desired to change some decisions made by local leaders in the last year; (ii) whether respondent expressed dissatisfaction with implementation of the development projects in the village in the past three years; (iii) whether the respondent had any complaints about corruption or nepotism related to the projects; (iv) whether the respondent prefers that a secret ballot election is used to select the village headman, in contrast to the customary hereditary practice; and (v) whether the respondent prefers that a secret ballot election is used to the select the President of Afghanistan.[39]

6   Empirical Results

The first empirical prediction of the model posits that at-large elections result in the election of more competent representatives. To the test this prediction, we estimate the following OLS model:

$$Education_{vi} = \alpha + \tau \cdot AL_v + \phi_q + \varepsilon_{vi}, \tag{3}$$

where $Education_{vi}$ is a dummy variable for whether council member i in village v has finished high school, AL_i is a dummy variable, which equals one if village v has been assigned at-large elections and zero if the village has been assigned district elections, ϕ_q is the quadruple fixed effect, and ε_{vi} is the error term. Standard errors are clustered at the village level. The first empirical prediction posits that $\tau > 0$.

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 electoral rules in attrition are not statistically significant.

[37] The latter three variables were winsorized at the 95th percentile and logged.

[38] The first two measures are available for female and male respondents separately. The latter two variables were winsorized at the 95th percentile and logged.

[39] Indicators (i) and (iv) are available for female and male respondents separately.

Under district elections, an average of 7 percent of male council members finished high school, whereas under at-large elections, the proportion is higher by 4 percentage points (the difference is statistically significant at the 1 percent level). Although modest in absolute magnitude, this effect constitutes a 57 percent increase in the share of male council members who finished high school. Given the overall low level of education of council members and the very low level of education and literacy in rural Afghanistan, the effect is substantively large. Thus, the results confirm the first empirical prediction of the model.

Due to the very low levels of variation in the competence of elected female representatives (which in turn are caused by the very low levels of female education in Afghanistan), the empirical analysis of the effects of electoral rules on education levels focuses on male council members. The-oretical analysis shows that the main predictions of the model regarding male candidates hold in the presence of a group with no such variation, but also predicts no effect of the electoral system on the competence of female candidates (see Subsection A7 of Appendix A).[40] Consistent with this prediction, the broadening of the sample to include both male and female council members does not substantively change the finding that at-large elections enhance the competence of elected representatives, but the effect is fully driven by male council members, with no effect for female council members (see Table C1 in Appendix C).[41]

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

$$Education_{vi} = \alpha + \gamma \cdot AL_v + \delta \cdot AL_v \cdot Heterogen_v + \mu \cdot Heterogen_v + \phi_q + \varepsilon_{vi}, \qquad (4)$$

[40] The extenstion of the model in Subsection A7 of Appendix A also predicts that female council members would have more extreme policy preferences if elected by district elections and that the extremity of such preferences would be higher than that of male council members. Unfortunately, we do not have information on the location of households for female council members, so cannot test this prediction.

[41] Another reason to focus on male council members is that in Afghan villages, women are traditionally excluded from community-level decision making and, consequently, female council members are not expected to have a significant effect on project selection and implementation. This is confirmed by Beath, Christia and Enikolopov (2013b), which finds that the preferences of even the most important female villagers do not have a significant effect on the choice of projects. However, despite their limited role in project-related decisions, female council member have played an important role in solving women-related issues and their election had an important effect on attitudes toward women (Beath. Christia, and Enikolopov, 2013a). However, these considerations are not related to the trade-off analyzed in this paper.

where $Heterogen_v$ is a measure of heterogeneity of village v and all other variables are the same as in equation (3). 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).

Consistent with the second empirical prediction, we find that the effect of variation in electoral rules is significantly stronger in more heterogeneous villages for all measures of heterogeneity (columns 2–4 in Table 3). Consistent with the third empirical prediction, we find no significant effect of variation in electoral rules on the competence of elected candidates in homogenous villages. The results also provide some evidence that increasing heterogeneity is associated with lower competence of elected candidates in villages with district elections. Although this result is not based on an exogenous source of variation, it provides evidence that is consistent with Proposition 2 and thus with the hypothesized mechanism for the effect.

The first two empirical predictions hold – but with smaller coefficients – if we define competence as the completion of middle school rather than high school (see Table C2 in Appendix C), but do not hold if defined as the completion of primary school. The first empirical prediction also holds if competence is defined by holding an occupation other than farming (see Table C3 in Appendix C).[42] The results for the second empirical prediction are no longer statistically significant if competence is defined by occupation, but the signs of the coefficients are aligned with the predictions of the model. The results in Table 3 are also robust to looking at village averages, rather than at individual council members, which indicates that the results are not driven by a few villages with a high fraction of educated council members.

The fourth empirical prediction posits that district elections will lead to the election of representatives that hold more extreme policy preferences. We test this prediction we use the following OLS model by estimating a model similar to (3) with the outcome variable provided by the logarithm of the distance between the house of an elected council member and the village center:

$$Distance_{vi} = \alpha + \chi \cdot AL_v + \upsilon \cdot AverageDistance_v + \phi_q + \varepsilon_{vi}, \tag{5}$$

where $Distance_{vi}$ is the logarithm of the distance between the house of council member i and the center of the village, $AverageDistance_v$ is the average distance between houses of baseline survey

---

[42]Farming is the occupation that is characterized by the lowest level of education. Of council members, 53 percent are farmers. The results also hold if we broaden the definition to include keepers of livestock and unskilled laborers.

respondents in village v (which reflects the size of the village), and other variables are the same as in the previous equations. The fourth empirical prediction posits that $\chi < 0$.

Consistent with the fourth empirical prediction, we find that the distance between the homes of elected officials and the center of the respective village is smaller in at-large elections (see Table 4).[43]

The fifth empirical prediction posits that, in district elections, there will exist a negative correlation between a representative's competence and the extremity of his/her policy preferences. We test this prediction by estimating the following model:

$$Education_{vi} = \alpha + \varphi \cdot AL_v + \rho \cdot AL_v \cdot Distance_{vi} + \kappa \cdot Distance_{vi} + \phi_q + \varepsilon_{vi}, \qquad (6)$$

where all the variables are the same as in the previous equations. The empirical prediction posits that $\kappa < 0$ and $\rho + \kappa = 0$.

The findings indicate that, in district elections, there is a negative, although not statistically significant, correlation between a representative's education and the distance between his house and the center of the village (see Table 5). In addition, the difference between the effects of distance in villages with at-large and district elections is positive and statistically significant in all specifications, confirming that $\rho > 0$. We also find that, in at-large elections, the correlation between a representative's educational attainment and the distance between the representative's house and the center of the village is positive (rather than zero, as implied by the model), although it is statistically significant only if village size is not controlled for.[44] As such, we observe only limited support for the fifth empirical prediction.

The sixth empirical prediction posits that the speed of project implementation will be higher in at-large elections. To test this hypothesis, we estimate the following OLS model:

$$Progress_{jv} = \alpha + \lambda \cdot AL_v + \phi_q + \varepsilon_{jv}, \qquad (7)$$

where $Progress_{jv}$ is either time to project start or time to project completion and all other variables

---

[43]Not surprisingly, the distance between the homes of elected officials and village centers is larger in villages with greater distance between the houses of baseline respondents.

[44]We also 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 their village.

are the same as in (3). According to the sixth empirical prediction $\lambda > 0$.

The results indicate that in villages with at-large elections, project implementation started earlier and, as a result, was completed faster (see Table 6). The results hold if we control for the number, cost, and type of projects and thus do not appear to be driven by an indirect effect of electoral rules on the type of projects selected.[45] The magnitude of the effect is substantial, with projects starting and finishing more than a month earlier in villages with at-large elections. A comparison of the distribution of dates for project commencement and completion by type of elections (Figures C1 and C2 in Appendix C) indicates that the difference in the speed of project completion reflects a shift in the distribution, rather than the effect of a few outliers. The results are thus consistent with the sixth empirical prediction.

The seventh and final empirical prediction posits that the effectiveness of development projects, improvements in general economic welfare, and increases in support for local leaders and democratic processes will be higher under at-large elections. To test this prediction, we estimate the following model:

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

where Y_{vit} is the outcome of interest for individual i in village v at time $t \in \{1, 2\}$ which corresponds to the midline and endline surveys respectively; τ_{1t} and τ_{2t} are dummies for the midline and endline surveys; all other variables are the same as in (3). Thus, the coefficients γ_1 and γ_2 measure the effect of at-large elections on the outcomes of interest at the time of the midline and endline surveys, respectively. Standard errors are clustered at the village level.

Estimates using data from the endline survey, which was administered approximately 4 years after the election of the village developments councils and 1–2 years following the completion of NSP-funded projects, indicate that villages assigned at-large elections exhibited lower levels of diarrhea among children (see Panel A of Table 7) and thus possessed water sources more likely to be free of contaminants.[46] We also observe that at-large elections induced a marginally significant increase

---

[45]There is some evidence that the number of projects is smaller and the cost of projects is higher if councils are populated by more educated representatives. This potentially suggests that more competent candidates are able to implement more complex projects and, in turn, may explain why there is no statistically significant effect of electoral rules on the duration of projects.

[46]Note that not all villages implemented drinking water projects. However, restricting the sample only to villages that implemented drinking water projects would potentially confound the estimation of the effects of electoral rules

in revenue derived from the harvest prior to the endline, which may reflect more successful implementation of irrigation projects. The signs of coefficients for other indicators of the effectiveness of development projects (reported water quality, hours of electricity available, yield of the most recent harvest) are consistent with at-large elections producing better project-related outcomes, although none of these attain conventional levels of statistical significance. Overall, there is weak evidence that the quality of implemented development projects was higher in villages assigned to at-large elections and no evidence that the faster implementation of projects in at-large villages came at the sacrifice off project quality.

Estimates using data from the midline survey, which was administered approximately 2 years after the election of the council, indicate that female respondents in villages assigned at-large elections were more likely to express positive perceptions of the economic situation and were more likely to attribute positive economic changes to local leaders (see Panel B of Table 7). However, there is no such effect for male respondents at midline or endline or for female respondents at endline. Data from the endline survey also indicates that there is a marginally significant positive effect of at-large elections on household expenditures. The effect of at-large elections on household income at endline is also positive, although not statistically significant. Overall, there is weak evidence that at-large elections improve general economic welfare.

Estimates using data from the endline survey indicate that, in villages assigned at-large elections, male respondents were less likely to express a desire to change decisions made by village leaders and were more supportive of using democratic procedures to select the President of Afghanistan (see Panel C of Table 7). Data from the midline survey indicate that male villagers in at-large villages were more likely to support election of the village headman, but the effect decreases in magnitude and loses statistical significance at endline. The results for other indicators are not statistically significant, although the signs of all coefficients for measures of support among male villagers for local leaders and democratic processes are consistent with the hypothesis that at-large elections improve villager attitudes. Overall, there is weak evidence that at-large elections resulted in improved attitudes toward village leaders and democratic processes.

On the whole, estimates using measures of villager welfare and opinion from the midline and endline surveys are weakly consistent with the seventh empirical prediction. However, as the data used

given the endogeneity of project choice.

to produce these estimates are subject to the imprecision associated with the capture of both subjective and objective information through household surveys and describe processes and outcomes over which representatives of the village development councils potentially have limited influence, these estimates are likely attenuated by measurement error. At a minimum, the results show that the faster speed by which projects in villages assigned to at-large elections are implemented does not come at the cost of lesser quality projects.

7 Discussion

According to our theoretical model, electoral rules affect the competence of elected representatives by changing voter incentives to support candidates with more extreme policy preferences over more qualified candidates. The results from the field experiment are consistent with the predictions of the theoretical model. In particular, we find that the competence of elected representatives is higher in villages assigned to at-large elections and that this difference is higher in more heterogeneous villages. There is also evidence that, in more heterogeneous villages, district elections lead to the election of representatives with more extreme policy preferences and lower levels of competence.

There are several possible alternative explanations for the positive effect of at-large elections on the quality of candidates. First, the effect may be a mechanical product of the geographical restrictions imposed by district elections on candidate selection, which preclude the election of multiple competent candidates residing in the same electoral district. The viability of this explanation is explored by examining the distribution of elected council members across districts. As prescribed, villages with district elections had exactly one male representative elected from each district. Although at-large elections imposed no formal restrictions on the geographical distribution of representatives, the distribution was similar to that in villages with district elections. Specifically, in villages with at-large elections, 93 percent of districts had a person residing in that district elected to the council. Only 37 out of 125 at-large villages had at least one district which did not have a resident council member and, of these, 25 villages had only one district that did not have a resident representative. Thus, the effect of at-large elections on the probability of a district not having a resident council member was very small.

To further test the viability of the explanation, we exclude from the sample representatives

31

from those districts that, in at-large elections, had more than one representative elected to the council.[47] Thus, we look only at the quality of candidates for whom the geographical restriction was not binding. While the results are to be interpreted with caution given the endogeneity of the restriction, we find that the sample restriction has no effect on estimates obtained by the benchmark specification (Table C4 in Appendix C). Finally, the restriction on candidates' residence in district elections should matter more 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, these results indicate that the effect of electoral rules on the quality of elected representatives is not affected by the geographical restriction on representatives imposed by district elections.

Another alternative explanation is that the larger district magnitude associated with at-large elections makes it harder for the incumbent to coordinate voting, which in turn reduces the incumbency advantage. Assuming that members of pre-existing elites are less qualified than potential challengers, this would increases the quality of elected representatives. The viability of this explanation is explored by examining the share of council members who were members of the de facto elite, for which data was collected from villagers in the baseline survey administered prior to the council elections.[48] The results indicate that electoral rules have no significant effect on the share of the council members that also belong to the group of pre-existing elites (see Table C5 in Appendix C). In fact, one measure provides a marginally significant positive effect of at-large elections on the share of council members who were named as members of the pre-existing elite. Similar results are obtained if we look at the proportion of pre-existing elite members who were subsequently elected to the council, with no significant difference between villages that used alternative electoral rules. Overall, the results indicate that electoral rules have no significant effect on incumbency advantage.

Yet another potential explanation is that educational attainment serves as a proxy for candidate competence if voters do not have precise information. 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 proxies such as educational attainment. In at-large elections, by contrast, the larger

---

[47]Ideally, we would exclude districts with more than one candidate of high quality, but since all villagers are considered candidates and we do not have a full census of the villages, we cannot implement this strategy.

[48]The share of council members that were members of the pre-existing elite, even if we use the most inclusive definition of pre-existing elites, is 39 percent, which indicates that pre-existing elites do not dominate the elected council.

size of districts leads to greater reliance on educational attainment as a proxy for candidate competence. While this interpretation could potentially explain the observation that at-large elections are associated with more educated representatives, it does a poor job of explaining why the results are stronger in more heterogeneous villages and why the educational attainment of elected representatives is on average lower in more heterogeneous villages. Specifically, in more heterogeneous villages, voters are likely to have less information on candidates, so according to this explanation voters should rely on formal education more. In addition, this explanation cannot account for the results regarding the location of the homes of elected representatives.

Differences in the competence of elected representatives may also reflect differences in 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 barriers to entry.

The observed correlation between at-large elections and competence of the representatives may also be driven by the advantage given by the at-large election system to competent candidates for reasons other than voting decisions by citizens. For instance, if candidates garner support through rallies (public speeches) and bribes (vote-buying) during the campaign and public speeches are more important in larger districts due to economies of scale, competent candidates will be advantaged in at-large elections if they have a comparative advantage in public speaking (and a comparative disadvantage at vote-buying).[49] This explanation, while attractive theoretically, cannot explain the results of the field experiment, as political campaigning was forbidden and election monitoring indicated no evidence of vote-buying. In addition, this story does not imply any correspondence between electoral rules and location of representatives'houses, which contradicts the experimental results.

8 Conclusion

This paper examines the effect of electoral rules on voting behavior and, by extension, on the composition of representative bodies and the quality of policy. The paper specifically compares two alternative electoral rules: district elections, with multiple single-member districts, and at-large

---

[49] 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.

elections, with a single multi-member district. Employing a theoretical model which builds on a citizen-candidate model with free entry of candidates and which assumes that voters value both competence and preference representation in electoral candidates, the paper posits that elections with a single multi-member district will result in the election of more competent representatives, with this effect becoming stronger in more heterogeneous communities. Empirical results from a field experiment conducted across 250 villages in Afghanistan are consistent with these predictions. Ancillary evidence indicates that these results are driven by strategic voting behavior informed by voter's anticipation of policy bargaining and not by alternative mechanisms such as incumbency advantage or comparative advantages in political campaigning.

The proposed theoretical model imposes a series of simplifying assumptions. In addition, the experiment used to test the model's theoretical predictions was conducted in a special institutional context that differed from those in which elections are ordinarily held, particularly in the absence of political parties, campaigns, or canvassing; the limited role of long-term considerations for candidates; and the limited mandate of the representative body. These assumptions and contextual peculiarities notwithstanding, the results of the paper provide a number of general implications.

First, the paper establishes the general result that electoral rules, by structuring the strategic incentives facing voters, condition voting behavior. The theoretical model demonstrates that rational voters will behave differently according to the magnitude of the districts of the election in which they are voting and the results of the experiment align with how voters should behave given the incentives presented to them by electoral rules.

Second, the paper shows that strategic voting results in the selection of candidates with relatively extreme policy preferences. As shown in the theoretical model, voters anticipate bargaining over policy and seek to elect candidates with relatively extreme policy preferences in order to ensure a policy outcome that is more likely to be favorable to the voter's specific interests. The results of the experiment indeed show that voters engage in such strategic behavior.

Third, the paper demonstrates that voters' preferences for candidates with relatively extreme policy preferences increase with heterogeneity over policy preferences of voters across electoral districts. That is, in cases where there exists high dispersion of policy preferences among voters, voters anticipate other voters' more extreme preferences and the deleterious consequences for the achievement of their policy ideal point, resulting in a preference for candidates with more extreme

34

policy preferences.

Fourth, the paper shows that elections with multiple single-member districts cause voters to focus on geographically-correlated aspects of policy preferences, such as the location of local public goods, and to elect candidates that can be expected to best represent the preferences of the district. In contrast, elections with single multi-member districts reduce the salience for voters of geographically-correlated policy preferences and cause voters to focus on other characteristics of candidates.

Fifth and finally, the theoretical model and empirical findings support the general implication that single multi-member district elections better enable the realization of voters' full set of preferences, while multiple single-member district elections unduly magnify the importance of preferences over geographically-correlated policies.

One limitation of the results of the paper, however, is the focus on policy aspects of representation. In addition to the task of bargaining over policy and facilitating policy implementation, elected representatives are also often tasked with providing constituents with various intangible services, such as assisting them in the navigation of bureaucratic processes and advocating on behalf of individual constituents or groups of constituents. It is certainly possible and, indeed, probable that the structure of direct representation provided by multiple single-member district elections may be more conducive to the provision of such services, as compared to the less direct systems of representation provided by single multi-member elections. To the extent to which citizens derive utility from receiving such services and to the extent to which there is ambiguity in the effects of the two election systems considered here on the provision of such services, the results of the paper cannot speak to welfare implications of the two election systems that extend beyond the policy function of elected representatives. An important area for future research is thus to extend the analysis to consider these aspects of representation. In addition, future research can shed light on how the mechanisms outlined in this paper play out in the settings with endogenous entry by candidates, repeated elections, active political campaigns, and political parties playing the role of intermediaries.

References

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin (2010) "Political Selection and Persistence of Bad Governments," *Quarterly Journal of Economics*, 125(4), 1511-1575.

Acemoglu, Daron, Tristan Reed, and James A. Robinson (forthcoming) "Chiefs: Elite Control of Civil society and Economic Development in Sierra Leone." *Journal of Political Economy.*

Adams, Greg (1996) "Legislative Effects of Single-Member vs. Multi-Member Districts," *American Journal of Political Science*, 40(1), 129-144.

Alesina, Alberto, Reza Baqir, and William Easterly (1999) "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114(4): 1243-1284.

Alesina, Alberto, and Howard Rosenthal (2000) "Polarized Platforms and Moderate Policies with Checks and Balances." *Journal of Public Economics*, 75(1), 1-20.

Angrist, Joshua D., and Jörn-Steffen Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder (2006) "Purple America," *Journal of Economic Perspectives*, 20(2): 97-118.

Aragones, Enriqueta, and Thomas Palfrey (2004) The Effect of Candidate Quality on Electoral Equilibrium: An Experimental Study, *American Political Science Review* 98 (1): 77-90.

Austen-Smith, David, and Jeffrey S. Banks (2005) *Positive Political Theory II Strategy and Structure*, The University of Michigan Press, Ann Arbor.

Banerjee, Abhijit, and Rohini Pande (2007) "Parochial Politics: Ethnic Preferences and Politician Corruption," CEPR Discussion Paper No. DP6381.

Banks, Jeffrey S., and John Duggan (2000) "A Bargaining Model of Collective Choice," *American Political Science Review*, 94(1): 73-88.

Banks, Jeffrey S., and Rangarajan K. Sundaram (1998) "Optimal Retention in Principal/Agent Models," *Journal of Economic Theory*, 82(2): 293-323.

Barfield, Thomas Jefferson (1984) *Weak Links on a Rusty Chain: Structural Weaknesses in Afghanistan's Provincial Government Administration.* Berkeley, CA: Institute of International Studies, University of California, Berkeley.

Baron, David. P., and Ferejohn, John.A. (1989) "Bargaining in legislatures," *American Political*

Science Review, 83(4): 1181-1206.

Barron, Patrick, and Benjamin Olken (2009) "The Simple Economics of Extortion: Evidence from Trucking in Aceh," *Journal of Political Economy* 117 (3):. 417-252,

Beath, Andrew, Fotini Christia and Ruben Enikolopov (2012) "Winning Hearts and Minds through Development Aid: Evidence from a Field Experiment in Afghanistan," MIT Working Paper.

Beath, Andrew, Fotini Christia and Ruben Enikolopov (2013a). "Empowering Women: Evidence from a Field Experiment in Afghanistan," *American Political Science Review*, 107(3): 540-557.

Beath, Andrew, Fotini Christia and Ruben Enikolopov (2013b) "Direct Democracy and Resource Allocation: Experimental Evidence from Afghanistan," MIT Working Paper.

Beath, Andrew, Fotini Christia and Ruben Enikolopov (2013c) "Do Elected Councils Improve Governance? Experimental Evidence on Local Institutions in Afghanistan," MIT Working Paper.

Besley, Timothy (2005) "Political Selection," *Journal of Economic Perspectives*, 19(3): 43-60.

Besley, Timothy, Rohini Pande, and Vijayendra Rao (2005) "Participatory Democracy in Action: Survey Evidence from India," *Journal of the European Economics Association Papers and Proceedings*, 3(2-3): 648-657.

Besley, Timothy and Stephen Coate (1997) "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112(1), 85-114.

Besley, Timothy and Stephen Coate (1998) "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review*, 88(1): 139-56.

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne (2013) "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden." Working paper.

Besley, Timothy, Jose Montalvo and Marta Reynal-Querol (2010) "Do educated leaders matter?" *Economic Journal*, 121, 205-227.

Besley, Timothy, Rohini Pande, and Vijayendra Rao (2005) "Political Selection and the Quality of Government: Evidence from South India." London School of Economics and Political Science, Working Paper.

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

Boesen Inger W. (2004) "From Subjects to Citizens: Local Participation in the National Solidarity Programme," AREU Working Paper Series.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini (2013) "The Political Resource Curse," *American Economic Review*, 103(5): 1759-96.

Bruhn, Miriam, and David McKenzie (2009) "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1(4): 200-232.

Burde, Dana, and Leigh L. Linden (2013) "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics*, 5(3): 27-40.

Callen, Michael and James D. Long (forthcoming) "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan," *American Economic Review.*

Callen, Michael, Charles Sprenger, Mohammad Isaqzadeh, and James D. Long (2014) "Violence and Risk Preference: Experimental Evidence from Afghanistan." *American Economic Review*, 104(1): 123-48.

Caselli, Francesco, and Massimo Morelli (2004) "Bad Politicians," *Journal of Public Economics*, 88(3-4), 759-782.

Casey, Katherine, Rachel Glennerster, and Edward Miguel (2012) "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan," *Quarterly Journal of Economics*, 127 (4): 1755-1812.

Congleton, Roger D., and Yongjing Zhang (2013) "Is it all about competence? The human capital of US presidents and economic performance." *Constitutional Political Economy* 24(2): 108-124.

Cox, Gary W. (1984) "Electoral Equilibrium in Double Member Districts," *Public Choice*, 44: 443-451.

Cox, Gary W. (1997) *Making Votes Count: Strategic Coordination in the World's Electoral Systems.* New York: Cambridge University Press.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi (2013) "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128(3): 1169-1218.

Dell, Melissa (2012) "The persistent effects of Peru's mining mita." *Econometrica* 78 (6): 1863-1903.

Duverger, Maurice (1954) *Political Parties: Their Organization and Activity in the Modern State.* Wiley, New York.

Egorov, Georgy, and Konstantin Sonin (2011) "Dictators and Their Viziers: Endogenizing the

Loyalty-Competence Trade-off," *Journal of the European Economic Association*, 9(5): 903-930.

Ferraz, Claudio and Frederico Finan (2008) "Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes," *Quarterly Journal of Economics*, 123(2): 703-745.

Ferraz, Claudio and Frederico Finan (2011a) "Electoral Accountability and Corruption in Local Governments: Evidence from Audit Reports," *American Economic Review,* 101(4): 1274-1311.

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

Ferree, Karen E., Bingham G. Powell, Jr., and Ethan Scheiner (2013) "How Context Shapes the Effects of Electoral Rules," in *APSA Report of the Task Force on Electoral Rules and Democratic Governance.*

Folke, Olle, Torsten Persson, and Johanna Rickne (2014) "Preferential Voting, Accountability and Promotions into Political Power: Evidence from Sweden." Working paper.

Gagliarducci, Stefano and Tommaso Nannicini (2013) "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association*, 11, 369-398.

Galasso, Vincenzo, and Tommaso Nannicini (2011) *American Political Science Review*, "Competing on good politicians." 105(1): 79-99.

Iversen, Torben, and David Soskice (2006) "Electoral Institutions and the Politics of Coalitions: Why Some Democracies Redistribute More than Others," *American Political Science Review*, 100(2), 165-181.

Kakar, Palwasha (2005) "Fine-Tuning the NSP: Discussions of Problem's and Solutions with Facilitating Partners," AREU Working Paper.

Lijphart, Arend (2004) "Constitutional Design for Divided Societies," *Journal of Democracy*, 15: 96-109.

Lizzeri, Alessandro, and Nicola Persico (2001) "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review*, 91: 225-245.

Madison, James (1788, 1961). *The Federalist Papers: A Collection of Essays in Support of the Constitution of the United States.* New York: Doubleday.

Martinez-Bravo, Monica (2014a) "The Role of Local Officials in New Democracies: Evidence from

Indonesia." *American Economic Review* 104 (4): 1244-1287.

Martinez-Bravo, Monica (2014b) "Educate to Lead? The Effect of Politicians' Education on Public Good Provision: Evidence from Java," working paper.

Mattozzi, Andrea, and Antonio Merlo (2007) "Mediocracy," NBER Working Paper 12920.

Mattozzi, Andrea, and Antonio Merlo (2008) "Political Careers or Career Politicians?" *Journal of Public Economics*, 92: 597-608.

Mattozzi, Andrea, and Erik Snowberg (2014) "The Right Type of Legislator," mimeo.

Milesi-Ferretti, Gian Maria, Roberto Perotti and Massimo Rostagno (2002) "Electoral Systems and the Composition of Public Spending." *Quarterly Journal of Economics*, 117, 609-657.

McKelvey, Richard, and Raymond Reizman (1992) "Seniority in Legislatures," *American Political Science Review*, 86, 597-608.

Moser, Robert G. (2008) "Electoral Systems and the Representation of Ethnic Minorities: Evidence from Russia," *Comparative Politics*, 40(3): 273-292.

Moser, Robert G., and Ethan Scheiner (2012) *Electoral Systems and Political Context: How the Effects of Rules Vary Across New and Established Democracies*. New York: Cambridge University Press.

Myerson, Roger (1993), "Effectiveness of Electoral Systems for Reducing Government Corruption: A Game-Theoretic Analysis," *Games and Economic Behavior*, 5: 118-132.

Nojumi, Neamat, Dyan E. Mazurana, and Elizabeth Stites (2004) Afghanistan's Systems of Justice: Formal, Traditional, and Customary. Working Paper.

Norris, Pippa (2004) *Electoral Engineering: Voting Rules and Political Behavior.* New York: Cambridge University Press

Olken, Benjamin (2007) "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 115 (2): 200-249

Ortuño-Ortín, Ignacio (1997) "A Spatial Model of Political Competition and Proportional Representation." *Social Choice and Welfare* 14(3), 427-438.

Osborne, Martin, and Al Slivinski (1996) "A Model of Political Competition with Citizen-Candidates, *Quarterly Journal of Economics*, 111(1), 65-96.

Pande, Rohini (2013) "Can informed voters enforce better governance? Experiments in low-income democracies," *Annual Review of Economics* 3(1): 215-237.

Persson, Torsten, and Guido Tabellini (1999) "The Size and Scope of Government: Comparative Politics with Rational Politicians, 1998 Alfred Marshall Lecture," *European Economic Review* 43, 699-735.

Persson, Torsten, and Guido Tabellini (2000) *Political economics: Explaining economic policy.* Cambridge, MA: MIT Press.

Persson, Torsten, and Guido Tabellini (2004) "Constitutional Rules and Fiscal Policy Outcomes," *American Economic Review*, 94, 25-45.

Persson, Torsten, and Guido Tabellini (2005) *The Economic Effects of Constitutions.* Cambridge, MA: MIT Press.

Persson, Torsten, Guido Tabellini, and Francesco Trebbi (2003) "Electoral rules and corruption," *Journal of the European Economic Association*, 1(4), 958-989.

Rahmani, Ahmad Idrees (2006) *The Role of Religious Institutions in Community Governance Affairs: How are Communities Governed Beyond the District Level?* Budapest, Hungary: Open Society Institute, Central European University Center for Policy Studies.

Rubinstein, Ariel (1982) "Perfect equilibrium in a bargaining model," *Econometrica* 50(1): 97-109.

Trebbi, Francesco, Philippe Aghion & Alberto Alesina (2007) "Electoral Rules and Minority Representation in U.S. Cities," *Quarterly Journal of Economics* 128(1): 325-358.

Van Weelden, Richard (2013) "Candidates, Credibility, and Re-election Incentives," *Review of Economic Studies*, 80(4): 1622-1651.

Table 1. Balance of Pre-treatment Covariates

	Mean (Full Sample)	Standard Deviation (Full Sample)	Observation	Mean in District Villages	Mean in At-large Villages	Standardized Difference
Number of Households in Village	118.43	115.71	2264	121.70	115.08	0.06
Household Members	9.79	5.00	2374	9.57	10.02	0.09
Household Members Under 15 Years	4.58	2.76	2374	4.50	4.66	0.06
Distance in Meters from the Respondent's House to Village Center	402	1011	2078	351	455	0.10
Primary Source of Household Income is Agriculture	0.66	0.47	2360	0.67	0.64	0.07
Age of Male Head-of-Household Respondent	43.81	13.30	2336	43.97	43.65	0.02
Male Head-of-Household Respondent Has No Formal Education	0.71	0.45	2387	0.73	0.69	0.09
Male Head-of-Household Respondent Finished Middle School	0.07	0.26	2387	0.07	0.08	0.02
Male Head-of-Household Respondent Finished High School	0.04	0.20	2387	0.04	0.05	0.03
First Language of Male Head-of-Household Respondent is Dari	0.7	0.46	2387	0.72	0.69	0.05
Village is Ethnically Mixed	0.24	0.43	250	0.21	0.26	0.13
Household Never or Rarely Faces Food Shortages	0.45	0.50	2387	0.43	0.47	0.09
Household's Main Source of Drinking Water is Unprotected Spring	0.27	0.44	2387	0.28	0.26	0.03
Household has Access to Electricity	0.15	0.35	2387	0.14	0.15	0.04
Household has a Mobile Phone	0.18	0.38	2387	0.19	0.17	0.04
Household has a Radio	0.75	0.43	2387	0.74	0.76	0.05
Household Expenditure on Food in Last 30 Days	3561	1982	2340	3524	3600	0.04
Household Received Loan in Past 12 Months	0.47	0.50	2387	0.48	0.46	0.05
Most Preferred Project of Male Respondents is Drinking Water	0.29	0.46	2387	0.30	0.28	0.05
Most Preferred Project of Male Respondents is Irrigation	0.13	0.33	2387	0.11	0.15	0.11
Most Preferred Project of Male Respondents is Electricity	0.06	0.24	2387	0.06	0.06	0.02
Most Preferred Project of Male Respondents is Road or Bridge	0.15	0.36	2387	0.16	0.14	0.04
Male Head-of-Household Respondent Attends *Shura* Meetings	0.32	0.47	2387	0.33	0.31	0.05

Note: Standardized difference reflects difference in means between district and at-large villages, divided by the standard deviation of district villages.

Table 2. Summary Statistics for Outcome Variables

	Mean	Standard Error	Obs.
Male Council Member Finished High School	8.68	28.16	2,016
Male Council Member Finished Middle School	16.82	37.41	2,016
Male Council Member Finished Primary School	32.64	46.90	2,016
Female Council Member Finished High School	0.80	8.92	1,995
Female Council Member Finished Middle School	2.81	16.52	1,995
Female Council Member Finished Primary School	9.62	29.50	1,995
Distance between Residences of Council Members and Village Center (Meters)	406	867	1,018
Days between election of the council and start of a project	381.38	368.2	1,239
Days between election of the council and end of a project	701.49	191.16	1,141

Table 3. Effect of Electoral Rules on Council Members' Competence

	Percent of Male Council Members who Finished High School			
	(1)	(2)	(3)	(4)
At-Large Elections	3.96***	0.07	1.95	-0.90
	[1.32]	[1.80]	[1.23]	[1.55]
Fractionalized Project Preferences		7.97***		
* At-Large Elections		[3.01]		
Fractionalized Project Preferences		-3.10		
		[1.99]		
Ethnically Mixed Village			7.96**	
* At-Large Elections			[3.21]	
Ethnically Mixed Village			-3.27	
			[2.21]	
Geographically Large Village				9.96***
* At-Large Elections				[3.01]
Geographically Large Village				-3.36*
				[1.96]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,016	2,016	2,016	2,016
R-squared	0.18	0.19	0.19	0.19

Note: The unit of observation is council member. The dependent variable is a dummy variable that equals 100 if a council member finished high school and zero otherwise. 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 Council Member Location

	Natural Log of Distance between Residences of Council Members and Village Center	
	(1)	(2)
At-Large Elections	-0.32***	-0.28***
	[0.11]	[0.08]
Natural Log of Median Distance between Residences of Villagers		0.75***
		[0.06]
Quadruple Fixed Effects	Yes	Yes
Observations	1,003	1,003
R-squared	0.21	0.36

Note: The unit of observation is council member. The sample includes only male council members. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5. Location of Council Member Residence and Educational Attainment

	Council Member Finished High School	
	(1)	(2)
Natural Log of Distance between Council Member Residence and Village Center	-0.65	-1.89
	[1.22]	[1.24]
At-Large Election * Natural Log of Distance between Council Member Residence and Village Center	4.45**	4.75**
	[2.24]	[2.20]
At-Large Elections	-16.16	-17.62
	[11.20]	[11.00]
Natural Log of Median Distance between Residence of Villagers		4.10**
		[1.70]
Quadruple Fixed Effects	Yes	Yes
p-Value for Effect of Distance in At-Large Villages	0.05	0.15
Observations	857	857
R-squared	0.21	0.21

Note: The unit of observation is council member. The dependent variable is a dummy that equals 100 if a council member finished high school and zero otherwise. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6. Effect of Electoral Rules on on the Speed of Project Implementation

	Days Between Elections and Project's Start				Days Between Elections and Project's Completion			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At-Large Elections	-40.05**	-42.91**	-39.25**	-37.40**	-33.94*	-35.23*	-32.13*	-32.74*
	[16.29]	[16.85]	[16.79]	[16.99]	[18.11]	[18.40]	[18.43]	[18.38]
Number of Projects in a Village			18.07*	10.45			15.29	17.77
			[10.72]	[11.65]			[10.86]	[11.24]
Log (Budget of the Project in USD)				-35.13***				11.45
				[13.34]				[13.35]
Type of Project Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	478	478	478	478	478	478	478	478
R-squared	0.34	0.35	0.36	0.38	0.38	0.40	0.40	0.40

Note: The unit of observation is project. The dependent variable in (1)-(4) measures the number of days between council election in a village and the start of a project. The dependent variable in (5)-(8) measures the number of days between the start of the project and the day project was finished. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7. Effect of Electoral Rules on Economic and Political Outcomes.

Dependent Variable:	At-large Elections* Midline	Standard Error	At-large Elections* Endline	Standard Error	Obs.	R-squared
Panel A. Benefits Delivered by Projects						
(1) Seasons in Past Year Water Was of Poor Quality	-0.074	[0.092]	-0.056	[0.091]	3,806	0.14
(2) Child Suffered Diarrhea in Past 2 Weeks			-0.057***	[0.019]	2,465	0.06
(3) Hours of Electricity in Past Month (winsorized and logged)	0.141	[0.177]	0.210	[0.189]	4,464	0.42
(4) Yield of Most Recent Harvest (tons, winsorized and logged)	0.003	[0.031]	0.048	[0.031]	2,954	0.42
(5) Revenue from Most Recent Harvest (winsorized and logged)	0.171	[0.165]	0.246*	[0.134]	2,926	0.40
Panel B. General Economic Outcomes						
(6) Economic Situation Has Improved in Past Year (Female Respondents)	0.045*	[0.026]	-0.016	[0.019]	3,997	0.15
(7) Economic Situation Has Improved in Past Year (Male Respondents)	-0.008	[0.024]	-0.012	[0.018]	4,493	0.19
(8) Respondent Attributes Positive Economic Change to Local Leaders or Council (Female Respondents)	0.013**	[0.006]			2,135	0.04
(9) Respondent Attributes Positive Economic Change to Local Leaders or Council (Male Respondents)	-0.003	[0.008]			2,355	0.08
(10) Income Earned in Past Year winsorized and logged	-0.015	[0.030]	0.005	[0.033]	4,439	0.20
(11) Annual Expenditure winsorized and logged	0.030	[0.035]	0.067*	[0.034]	4,172	0.17
Panel C. Political Outcomes						
(12) Desired Change in Decision of Influential Villagers in Past Year (Female Respondents)	0.003	[0.007]	0.004	[0.008]	3,980	0.07
(13) Desired Change in Decision of Influential Villagers in Past Year (Male Respondents)	0.009	[0.009]	-0.022**	[0.009]	4,495	0.05
(14) Complaint About Project			-0.023	[0.031]	1,716	0.21
(15) Complaint about Corruption or Nepotism Related to Projects			-0.008	[0.018]	1,722	0.16
(16) Prefers Election to Select Headman (Female Respondents)	-0.010	[0.028]	-0.020	[0.031]	3,578	0.12
(17) Prefers Election to Select Headman (Male Respondents)	0.051**	[0.025]	0.016	[0.022]	4,440	0.23
(18) Prefers Selection of President by Secret Ballot Election (Male Respondents)			0.025**	[0.012]	2,112	0.07

Note: Each row represents a separate regression. Measures are based on answers of villagers to surveys conducted approximately two years (midline) and four years (endline) after council elections. Some of the questions were asked only during one of the surveys, so the coefficients for the other survey is not available. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Online Appendix

Appendix A: Robustness

The model in Section 3 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. Subsection A1 considers an extension of the model in which we allow for multiple competent candidates. In subsection A2, we study an explicit legislative bargaining game. In subsection A3 we contrast joint decision-making with individual decision-making by the elected council members. In subsection A4 we consider the case of multiple electoral districts, and thus more than two council members. In subsection A5 we discuss where competence of council members affects their bargaining power. In subsection A6 we demonstrate that the results from the baseline model hold if council members are elected sequentially. Finally, in subsection A7 we show that the predictions of the model hold if citizens vote for male and female council members separately.

A1 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 ideal point, 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. The median voter theorem applies again, and for at-large elections, a pair of citizens that maximizes the utility of the median voter, $w_{m_X}(a_l, b_l, a_r, b_r)$, will 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 $\left(h, q_{(N)}\right)$, provided that there is a competent individual in the district $(q_{(N)} \geq 0)$, 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 $BR_R(b_l) \subset \left\{q_{(N)}, B\right\}$. 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 $q_{(N)} \in BR_R(b_l')$, then $q_{(N)} \in BR_R(b_l)$. 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 $BR_L(b_r) \in \left\{-B, q_{(1)}\right\}$. 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 does not hold, then $BR_R(b_l) = q_{(N)}$ for any b_l, and thus there is an equilibrium where R elects individual with type $\left(h, q_{(N)}\right)$ and L elects (a_l, b_l), where $b_l \in BR_L\left(q_{(N)}\right)$. Similarly, if the second inclusion fails, then there is an equilibrium where L elects $\left(h, q_{(1)}\right)$ and R elects (a_r, b_r) with $b_r \in BR_R\left(q_{(1)}\right)$. 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 $\frac{1}{4} + \frac{1}{4} - \left(\frac{\frac{1}{2} + \left(-\frac{1}{2}\right)}{2} - \frac{1}{2}\right)^2 = \frac{1}{4}$ 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 $\frac{1}{4} - \left(\frac{\frac{1}{2} + (-1)}{2} - \frac{1}{2}\right)^2 = -\frac{5}{16}$ 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 A1 *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, which is proven, along with other results, in Appendix B.

Lemma A1 *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*

$$
\begin{aligned}
P(N, z) &= \Pr\left(\exists j \in \{1, \ldots, N\} : 1 - q_j \leq z\right), \\
Q(N, z) &= \Pr\left(\exists i, j \in \{1, \ldots, N\}, i \neq j : |q_i + q_j| \leq z\right).
\end{aligned}
$$

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.

In what follows, assume that $\frac{h}{k} < \frac{1}{4}B^2$; 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, e.g., (h, q_1), and a corresponding 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 A1 (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 consider 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_l; then the the

median voter in district R will elect a competent citizen only if there is one with political position q_j such that $h - k \left(\frac{b_l + q_j}{2} - \frac{B}{2} \right)^2 \geq -k \left(\frac{b_l + B}{2} - \frac{B}{2} \right)^2$, i.e., if $q_j \geq B - b_l - \sqrt{4\frac{h}{k} + b_l^2}$. This is equivalent to $1 - \frac{q_j}{B} \leq \frac{b_l}{B} + \sqrt{4\frac{h}{kB^2} + \left(\frac{b_l}{B} \right)^2}$; therefore, by Lemma A1, for any given b_l, the probability that district R elects a competent citizen is $P(N, z(b_l))$, where $z(b_l) = \frac{b_l}{B} + \sqrt{4\frac{h}{kB^2} + \left(\frac{b_l}{B} \right)^2}$. 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, for any fixed council member from the left district who may be elected, 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 A2 *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.

A2 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 legislative bargaining game, as in Banks and Duggan (2000). 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. In this model, we assume $\frac{h}{k} < \frac{1}{2}B^2$; this is required to obtain strict results in Proposition A3 below, but it does not affect existence of equilibrium.

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 the 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 A2 *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 $\beta \geq \frac{P - k(b_r - b_l)^2}{P - \frac{k(b_r - b_l)^2}{2}}$, then l and r propose $\frac{b_l + b_r}{2} - \left(\sqrt{\frac{P}{k} + \left(\frac{b_r - b_l}{2(1 - \beta)} \right)^2 \beta(2 - \beta)} - \frac{b_r - b_l}{2(1 - \beta)} \right)$ and

$\frac{b_l + b_r}{2} + \left(\sqrt{\frac{P}{k} + \left(\frac{b_r - b_l}{2(1 - \beta)} \right)^2 \beta(2 - \beta)} - \frac{b_r - b_l}{2(1 - \beta)} \right)$, *respectively, and other council member is indifferent between accepting and rejecting these proposals;*

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

Lemma A2 says the following. If the discount factor β 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 β 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 β is close to 0 is also noteworthy. The following characterizes comparative statics in these extreme cases.

Proposition A3 *For any β, there exists an equilibrium. Moreover, 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 converge (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 A3 gives two important takeaways. First, if offers are made frequently and β 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 β 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 $k \left(q - \frac{B}{2} \right)^2 \leq h$ in district R and if $k \left(q + \frac{B}{2} \right)^2 \leq h$ 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.

A3 Joint and individual decisions

So far, we have assumed that the two council members 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.

Suppose, for example, that the legislative body makes decisions on a number of questions, and only share α requires a joint decision, while for $1 - \alpha$, a random council member is appointed to make a unilateral decision. The case considered in Section 3 corresponds to $\alpha = 1$, while $\alpha < 1$ may correspond to situations where some policy decisions are local, and the local council member has the sole responsibility of making the decision.

It turns out that our results remain intact for α sufficiently high, but as α becomes smaller, district elections will dominate at-large ones. 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 $h > k \left(B/2 \right)^2$, 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. Notice that this result is very similar to the prediction of Subsection A5 (Proposition A3): there, if β is low enough, the ideal points of council members would be picked with equal probability, which matches the case $\alpha = 0$.

This result once again confirms that our results are driven by the joint nature of decision-making in councils. At-large elections are preferred if council members make a joint decision. If they have

multiple policy questions which they split between themselves, then district elections should have an edge. Studying such trade-offs in more detail seems to be a fruitful area for future research.

A4 Multiple districts

In this subsection, we explore robustness of the results if there are multiple districts. Suppose that in district elections, the village is divided into M equally-sized contiguous districts, so for $j \in \{1, M\}$, district $D_j = \left[-B + \frac{2B}{M}(j-1), -B + \frac{2B}{M}j\right]$, 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 and can vote for M different citizens; this ensures existence of equilibrium, which will maximize the utility of the median voter. To generalize the decision-making in the council, we start with the case where council members play a bargaining game with random recognition as in Subsection A2. Namely, each council member is chosen randomly to make a proposal, and a proposal is accepted if sufficiently many council members support it. Let us focus on simple majority rules (which generalizes Subsection A2): a proposal is accepted if more than $\frac{M}{2}$ council members support it.

With this setup, one can easily show that decisions will be made by median voters in respective districts both in at-large and in district elections. To understand their incentives, consider the outcome of bargaining between four council members with political preferences $b_1 \leq \cdots \leq b_M$. Since this model is a particular case of Banks and Duggan (2000), it is characterized by an acceptance set, with each council member proposing ideal point from this set. It is not hard to show that if β is close to 1, then the acceptance set converges to a point (see Austen-Smith and Banks, 2005). Moreover, since the utility functions are symmetric (and quadratic, so all council members have the same preferences regarding the uncertainty of the outcome if the current one is rejected), 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. The intuition is very simple. If M is odd, then every council member, except for the median one $b_{\frac{M+1}{2}}$, chooses his proposal subject to the constraint that the median voter is indifferent between accepting and waiting; once this is true, all members lying on one of the sides would be in favor of accepting, which is enough for a majority. Thus, $b_{\frac{M+1}{2}}$ must always be in the acceptance set. If M is even, then an agenda-setter needs to get agreement from the two median voters, and then standard arguments

A-8

would imply convergence to the midpoint between them.

Let us take the limit $\beta \to 1$ and assume, for simplicity, that a council with an odd number of members chooses $b_{\frac{M+1}{2}}$, and a council with an even number of members chooses $\frac{1}{2}\left(b_{\frac{M}{2}} + b_{\frac{M}{2}+1}\right)$. In at-large elections, it is feasible to achieve the first best by electing the competent citizen (h, q) and complementing him with other members so that the council chooses policy 0. In district elections, only the median districts (one or two) have strategic incentives not to elect the competent citizen if he happens to reside there; other districts do not have an influence on policy in equilibrium, and therefore will elect the competent citizen if they can, or may pick a random citizen otherwise. Consequently, at-large elections are more likely to elect the competent citizen if he lives close to the center, but the difference disappears if he lives far; note, however, that this result relies on the assumption of majority voting.

A5   Education and bargaining power

We have assumed that competence directly affects citizens' utilities, and have demonstrated that education is indeed correlated with faster completion of projects (see Table 3). It is, however, also possible that education implies a higher bargaining power in the council. We may assume, for simplicity, that if a competent person bargains with an incompetent one, he is more likely to make a proposal. If so, then the equilibrium policy choice will be closer to the alternative that he prefers.

The median voter logic would still apply, but the incentives would be distorted. In at-large elections, the median voter would not always be able to get his ideal point 0 with a competent council member, if he resides sufficiently close to the border, because a more distant second member would be needed. Thus, it is now possible that the most competent person will not be elected; this will happen if the effect of competence on utility is small (h is small), but the distortion of bargaining power is substantial. At the same time, in district elections, the median voter would be more willing to elect the competent citizen, as he wants the equilibrium policy (0 if two extreme and incompetent citizens are elected) to be distorted towards his district.

Overall, if education is positively correlated with bargaining power, the effect that at-large elections lead to more competent council members will diminish. Intuitively, voters in district elections would prefer to elect competent council members, because this would help them distort the policy rather than hurt, as in the baseline model (see also Mattozzi and Snowberg, 2014, for

a similar effect). Yet, if the correlation between bargaining power and education is small, at-large elections would still lead to better councils, as in the baseline model.

A6 Electing one council member 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 A2, 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 chooses his ideal policy with equal probability (this happened if the discount factor β was low enough). This suggested that the results are driven by joint policy decisions rather than coordination. Similarly, in Subsection A3, if council members make decisions separately, the advantage of at-large elections disappears.

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, i.e., ability of voters to coordinate in at-large elections does not drive the results. (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 study the probability that the second council member would be competent. Suppose that the type of the existing council member 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.

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 bliss point of the competent individual at q. We focus on the case $q > 0$; if $q < 0$ (and in particular, if the competent citizen is already elected), then the question of comparing at-large elections and district elections becomes trivial, thus $q > 0$ is the interesting case.

Consider at-large elections first. The median voter is effectively choosing between mirroring the existing council member (thus electing someone with type $(0, -b_0)$ and getting utility $a_0 = 0$) and electing the competent citizen, thus getting utility $h - k \left(\frac{b_0 + q}{2} \right)^2$. He will choose the competent citizen if and only if $(b_0 + q)^2 \leq 4 \frac{h}{k}$, i.e., if q is in $2\sqrt{\frac{h}{k}}$-neighborhood of $-b_0$.

In district elections, the median voter is choosing between the most biased candidate (which will

A-10

give him utility $-k\left(\frac{b_0+B}{2}-\frac{B}{2}\right)^2=-k\left(\frac{b_0}{2}\right)^2$) and the competent one (which will give him utility $h-k\left(\frac{b_0+q}{2}-\frac{B}{2}\right)^2$. The competent candidate is elected if and only if $4\frac{h}{k}+(b_0)^2\geq(b_0-B+q)^2$, i.e., if q is in the $\sqrt{4\frac{h}{k}+(b_0)^2}$-neighborhood of $B-b_0$. Since $b_0<0$, this is true for $q\in\left[B+|b_0|-\sqrt{4\frac{h}{k}+(b_0)^2},B\right]$; the length of this interval is less than $2\sqrt{\frac{h}{k}}$. It is now clear that in expectation (taken over the value of b_0), at-large elections are still more likely to elect the 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 of the median voter for the new council member is $-b_0$, while in district elections, this point is $B-b_0$. Thus, in the former case, the induced ideal point is strictly in the interval of $[0,B]$, and in the latter case it is beyond this interval. This immediately implies 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 if one council member is to be elected, at-large elections produce superior results. 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).

A7 Two types of citizens

We have assumed so far the society is homogenous, and any composition of citizens can form a council. Suppose, instead, that the society consists of two parts of equal mass, men and women, with political preferences of each part distributed uniformly on $[-B,B]$. Suppose for now that only one citizen is competent. The electoral systems (at-large and district elections) are generalized in the following way. In district elections, the society is split into two districts, L and R, as before. Each citizen now has two votes: one must be cast for a man and the other must be cast for a woman. The man and the woman with the largest share of votes are elected into council. In at-large elections, the entire society comprises a district where each citizen needs to cast two votes for men and two votes for women; as before, assume that he cannot cast two votes for the same person. Here, two men and two women with the largest vote shares are elected into council.

The council now consists of four members, (a_i,b_i) for $i\in\{1,2,3,4\}$, where without loss of generality we assume that $\{b_i\}$ is nondecreasing. For simplicity, let us focus on a bargaining game

with a simple majority voting rule, where offers are made very frequently (similarly to the game considered earlier in Subsection A4). In the limit, the policy chosen by the council is halfway between the two median ideal points of council members, i.e., $p = \frac{1}{2}(b_2 + b_3)$. Thus, the utility of a citizen with ideal point b is now

$$w_i\left(\{a_j, b_j\}_{j=1,2,3,4}\right) = a_1 + a_2 + a_3 + a_4 - k\,(p - b)^2.$$

In what follows, we will look at both whether the competent person is elected (i.e., total competence $a_1 + a_2 + a_3 + a_4$) and at the polarization of the elected council, which we can, for simplicity, measure by the total bias, $|b_1| + |b_2| + |b_3| + |b_4|$.

Take at-large elections first. The median voter theorem still applies, again because of increasing differences consideration. The median voter prefers to elect the most competent person (with ideal point q) and three other people, such that the two median council members have the opposite ideal points. There are many ways to achieve this, but the following is true: in any equilibrium, the bliss points of the two median council members cannot exceed $|q|$ in absolute value; in other words, $|b_2|, |b_3| \leq |q|$. This follows from a simple argument: In equilibrium, $b_2 = -b_3$, so if both exceed $|q|$ in absolute value, then the competent person is one of the extreme council members. Moreover, in this case, bliss points of other council members are either above $|q|$ or below $-|q|$; in either case, $b_2 = -b_3$ cannot hold. This contradiction proves that the total bias of council members is limited from above by $2\,(|q| + B)$.

Now consider district elections. Suppose that district L contains the competent citizen, and suppose that the two citizens it elects have types (a_1, b_1) and (a_2, b_2); this may or may not include this competent citizen. As before, the median voter in district R would like b_3 to be as high as possible, and therefore it needs to elect two citizens with types $(0, B)$. District L can do one of the following: elect two extreme citizens of types $(0, -B)$, or elect one competent citizen with type (h, q) and another citizen; this other citizen may be chosen arbitrarily as long as his ideal point is less than q (if he is less biased than the competent citizen, this will affect the negotiated policy in a way that the median voter in district L would not like). In the first scenario, the society elects four citizens with maximum bias. In the second, two citizens have bias B, one has bias $|q|$, and one has bias between $|q|$ and B. In all cases, the total bias of council members is at least $2\,(|q| + B)$.

We have thus demonstrated the following: In at-large elections, the competent citizen is always elected, and in district elections this is not necessarily true. Also, for any equilibria played in at-large elections and in district elections, the total bias in at-large elections is at least as low as in district elections. This suggests that the implications of the theory in Section 3 are robust and the predictions remain the same.

It should also be noted that these results are not driven by the assumption that there is only one competent citizen, so one of the two groups of electorate has to contain low types only. Suppose, for example, that there are exactly two competent citizens, one male, with ideal point q_m, and one female, with ideal point q_f. In at-large elections, the society would still achieve the first best by electing the competent male and his opposite and the competent female and her opposite. In district elections, there again will be a trade-off between electing the competent citizen and the most biased one. Thus, at-large elections will lead to a weakly more competent council. In addition, as before, the total bias in at-large elections will be at most $|q_m| + |q_f| + 2B$, and the total bias in district elections will be at least $|q_m| + |q_f| + 2B$. Thus, total bias in at-large elections will not exceed total bias in district elections.

Appendix B: Proofs

B1 Proofs of main results

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 $b_r > b'_r$,

$$\mathbb{E}w_i\left(a_l, b_l, a_r, b_r\right) - \mathbb{E}w_i\left(a_l, b_l, a'_r, b'_r\right) > \mathbb{E}w_j\left(a_l, b_l, a_r, b_r\right) - \mathbb{E}w_j\left(a_l, b_l, a'_r, b'_r\right),$$

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

$$
\begin{aligned}
&\mathbb{E}w_i\left(a_l, b_l, a_r, b_r\right) - \mathbb{E}w_i\left(a_l, b_l, a'_r, b'_r\right) \\
=\ & \mathbb{E}a_l + a_r - \mathbb{E}k\left(\frac{b_l + b_r}{2} - b_i\right)^2 - \mathbb{E}a_l - a'_r + \mathbb{E}k\left(\frac{b_l + b'_r}{2} - b_i\right)^2 \\
=\ & (a_r - a'_r) + k\left(\frac{b_r - b'_r}{2}\right)\left(2b_i - \mathbb{E}b_l - \frac{b_r + b'_r}{2}\right),
\end{aligned}
$$

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 σ 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 σ. Then if $b' > \mathbb{E}b_l$, then all individuals with $b_i \geq -\frac{B}{2}$ prefer (a', b') because of increasing differences, and some of those with $b_i < -\frac{B}{2}$ 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 $b' < \mathbb{E}b_l$, whereas if $b' = \mathbb{E}b_l$, 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; 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

$$
\begin{aligned}
\mathbb{E}w_{m_L}(a_l, b_l, a_r, b_r) &= a_l + \mathbb{E}a_r - \mathbb{E}k\left(\frac{b_l + b_r}{2} + \frac{B}{2}\right)^2 \\
&= a_l + \mathbb{E}a_r - k\left(\frac{b_l + \mathbb{E}b_r}{2} + \frac{B}{2}\right)^2 - \frac{k}{4}\text{Var}(b_r),
\end{aligned}
$$

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 this citizen elects 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

$$
w_{m_L}(0, -B, 0, B) = w_{m_L}(h, q, 0, B).
$$

Since 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 $\frac{b_l + b_r}{2} > \frac{b'_l + b'_r}{2}$,

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

To see this, notice that

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

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 $(0, -q)$; let us show that there is no collective deviation that increases utility of all deviators. 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 equilibrium with the required properties.

Now, suppose that there is an equilibrium σ which induces some distribution over pairs of individuals $(a_l, b_l), (a_r, b_r)$ who get elected. Suppose first that $\mathbb{E}(b_l + b_r) = 0$. 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 σ is an equilibrium stipulated by the Proposition. If (h, q) is not part of the pair with a positive probability, then $\mathbb{E}(a_l + a_r) < h$. 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 $\mathbb{E}(b_l + b_r) \neq 0$; without loss of generality, $\mathbb{E}(b_l + b_r) < 0$. Consider coalition X of citizens with $b_i > \frac{\mathbb{E}(b_l+b_r)}{4}$; each of them prefers policy 0 to policy $\frac{\mathbb{E}(b_l+b_r)}{2}$, 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 $w_{m_L}(0, -B, 0, B) < w_{m_L}(h, q, 0, B)$, 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 $\frac{3}{4}B^2 > \frac{h}{k}$. When this is true, the probability that the competent citizen is elected is

$$R = 1 - \frac{\hat{q} - (-\hat{q})}{2B} = 1 - \frac{\hat{q}}{B} = \sqrt{4\frac{h}{kB^2} + 1} - 1. \tag{B1}$$

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 (B1)). Thus, $C_a \geq C_d$, because $P \leq 1$, and the inequality is strict whenever $R < 1$, which may be simplified to $\frac{h}{k} < \frac{3}{4}B^2$. The difference is $C_a - C_d = (1 - P)\frac{h}{2} = \left(2 - \sqrt{4\frac{h}{kB^2} + 1}\right)\frac{h}{2}$, 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 $P_a = \frac{1}{B}\int_0^B \frac{1}{B}qdq = \frac{1}{2}$. In district elections, it equals $P_d = \frac{1}{B}\left(\int_0^{\hat{q}} \frac{1}{B}Bdq + \int_{\hat{q}}^B \frac{1}{B}\left(\frac{q+B}{2}\right)dq\right) = \frac{1}{4}\left(3 - \frac{\hat{q}}{B}\right)\left(1 + \frac{\hat{q}}{B}\right)$, 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 $\frac{\hat{q}}{B} = 2 - \sqrt{4\frac{h}{kB^2} + 1}$, and thus is increasing in k and B.

Part 3. In at-large elections, for a council member (a, b), $\Pr\left(\frac{|b|}{B} < x \mid a = h\right) = \Pr\left(\frac{|b|}{B} < x \mid a = 0\right) = x$ (for $x \in [0, 1]$). Therefore, 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 $\Pr\left(\frac{|b|}{B} < x \mid a = h\right) = \frac{x - \hat{q}/B}{\hat{q}/B}$ 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, $\mathbb{E}U_a(q, b_i) = h - kb_i^2$, and

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

Thus,

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

This completes the proof. ∎

B2   Proofs of results from Subsection A1

Proof of Proposition A1. Existence (and generic uniqueness) of equilibrium in at-large elections is proven similarly to the corresponding part of Proposition 1; this proof is omitted. Existence of equilibrium in the case of district elections was proven in the text. ■

Proof of Lemma A1. 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,

$$
\begin{aligned}
P(N, z) &= \Pr\left(1 - q_{(N)} \le z\right) = \Pr\left(q_{(N)} \ge 1 - z\right) \\
&= 1 - \Pr\left(q_{(N)} \le 1 - z\right) = 1 - F^N(1 - z) \\
&= 1 - \left(\frac{1 - z + 1}{2}\right)^N = 1 - \left(\frac{2 - z}{2}\right)^N.
\end{aligned}
$$

We prove that $Q(N, z) \ge P(N, z)$ (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

$$Q(2, z) = \Pr\left(|q_1 + q_2| \le z\right) = \Pr\left(-z \le q_1 + q_2 \le z\right) = 2\Pr\left(0 \le q_1 + q_2 \le z\right),$$

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 $|x| \le 2$, and thus

$$
\begin{aligned}
Q(2, z) &= 2\Pr\left(0 \le q_1 + q_2 \le z\right) = 2\int_0^z \frac{2 - x}{4} dx = \frac{z(4 - z)}{4} \\
&= 1 - \left(\frac{2 - z}{2}\right)^2 = P(2, z).
\end{aligned}
$$

Now take $N \geq 4$

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

which proves the result for even N.

Consider the case of odd N. Suppose $N = 3$. Then we have

$$
\begin{aligned}
Q(3, z) &= \Pr\left(|q_1 + q_2| \leq z \vee |q_1 + q_3| \leq z \vee |q_2 + q_3| \leq z\right) \\
&> \Pr\left(|q_1 + q_2| \leq z \vee |q_1 + q_3| \leq z\right) \\
&= 1 - \Pr\left(|q_1 + q_2| \geq z, |q_1 + q_3| \geq z\right) \\
&= 1 - \Pr\left(q_1 + q_2 \leq -z \vee q_1 + q_2 \geq z, |q_1 + q_3| \geq z\right) \\
&= 1 - \Pr\left(|q_1 + q_3| \geq z\right) \Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z\right) \\
&= 1 - (1 - Q(2, z)) \Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z\right) \\
&= 1 - \left(\frac{2-z}{2}\right)^2 \Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z\right).
\end{aligned}
$$

It therefore suffices to prove that $\Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1 \mid |q_1 + q_3| \geq z\right) \leq \frac{2-z}{2}$. For that, it suffices to prove that $\Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1\right)$ for any $q_1 \in [-1, 1]$. To prove this, consider the case $q_1 \geq 0$ (the case $q_1 \leq 0$ is symmetric and may be considered similarly). If $z + q_1 \leq 1$, then $\Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1\right) = \frac{-z - q_1 - (-1)}{2} + \frac{1 - (z - q_1)}{2} = 1 - z \leq \frac{2-z}{2}$. If $z + q_1 > 1$, then $\Pr\left(q_2 \leq -z - q_1 \vee q_2 \geq z - q_1\right) = \Pr\left(q_2 \geq z - q_1\right) = \frac{1 - (z - q_1)}{2} < \frac{2-z}{2}$. 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

$$
\begin{aligned}
Q(N, z) \quad &= \quad \Pr\left(\exists i, j \in \{1, \ldots, N\}, i \neq j : |q_i + q_j| \leq z\right) \\
&> \quad \Pr\left(|q_1 + q_2| \leq z \vee \cdots \vee |q_{N-2} + q_{N-1}| \leq z \vee |q_{N-2} + q_N| \leq z \vee |q_{N-1} + q_N| \leq z\right) \\
&= \quad 1 - \Pr\left(|q_1 + q_2| \geq z \wedge \cdots \wedge |q_{N-2} + q_{N-1}| \geq z \wedge |q_{N-2} + q_N| \geq z \wedge |q_{N-1} + q_N| \geq z\right) \\
&= \quad 1 - \Pr\left(|q_1 + q_2| \geq z\right) \times \cdots \times \Pr\left(|q_{N-4} + q_{N-3}| \geq z\right) \\
&\quad \times \Pr\left(|q_{N-2} + q_{N-1}| \geq z \wedge |q_{N-2} + q_N| \geq z \wedge |q_{N-1} + q_N| \geq z\right) \\
&= \quad 1 - (1 - Q(2, z))^{\frac{N-3}{2}} \times (1 - Q(3, z)) = 1 - \left(\left(\frac{2-z}{2}\right)^2\right)^{\frac{N-3}{2}} \left(\frac{2-z}{2}\right)^3 = P(N, z).
\end{aligned}
$$

This completes the proof. ∎

Proof of Proposition A2. The result will follow from the following argument. Let C_a and C_d be random variables corresponding to total competences of councils and at-large and in district elections, respectively (where uncertainty is in locations of competent agents). We need to prove that $\mathbb{E}(C_a) > \mathbb{E}(C_d)$. For this, it suffices to prove that C_a first-order stochastically dominates C_d. Since the support of both distributions involves only three points, $0, 1, 2$, it suffices to prove that $\Pr(C_a = 0) < \Pr(C_d = 0)$ and $\Pr(C_a \leq 1) < \Pr(C_d \leq 1)$. The first is true, because $\Pr(C_a = 0) = 0$ (in at-large elections, one council member will always be competent, because electing some competent member (h, q_i) and an incompetent person $(0, -q_i)$ is always better for the median voter than two incompetent members); at the same time, $\Pr(C_d = 0) > 0$ (e.g., if all competent citizens are located close to 0, $|q_i| < \varepsilon$ for $1 \leq i \leq N$, then two extreme agents will be elected). It therefore suffices to prove that $\Pr(C_a = 2) > \Pr(C_d = 2)$.

Consider at-large elections. As argued in the text, two competent citizens will be elected if and only if for some q_i, q_j (where $i \neq j$), $|q_i + q_j| \leq 2\sqrt{\frac{h}{k}}$. If we take N random variable $r_i = \frac{q_i}{B}$, $1 \leq i \leq N$, they are independent and distributed uniformly on $[-1, 1]$. Consequently, $\Pr(C_a = 2) = Q(N, z)$ for $z = \frac{2}{B}\sqrt{\frac{h}{k}} < 1$.

Now consider district elections. We have $\Pr(C_d = 2) = \Pr(a_l = a_r = h) < \Pr(a_r = h)$. The latter probability is shown in the text not to exceed $P(N, z)$. Therefore, $\Pr(C_a = 2) = Q(N, z) \geq P(n, z) > \Pr(C_d = 2)$. This inequality shows that C_a first-order stochastically dominates C_d, which completes the proof. ∎

B3 Proofs of results from Subsection A2

Proof of Lemma A2. This bargaining model is a particular case of Banks and Duggan (2000), with unanimity voting rule. Theorem 1 in that paper shows that there exists a no-delay equilibrium and, moreover, every stationary equilibrium is a no-delay equilibrium; Theorem 2 implies that any such equilibrium is in pure strategies. Finding the explicit formulas and showing uniqueness reduces to a simple exercise, which is omitted. ■

Proof of Proposition A3. The utility of any agent with ideal point b from a council with types $(a_l, b_l), (a_r, b_r)$, is

$$a_l + a_r - k \left(\frac{b_l + b_r}{2} - b \right)^2 - kV(b_l, b_r), \tag{B2}$$

where $V(b_l, b_r)$ is the variance of the proposals by the two council members. As $\beta \to 1$, the equilibrium proposals of any two council members converge, uniformly, to $\frac{b_l + b_r}{2}$. Therefore, the variance of $V(b_l, b_r)$ uniformly converges to 0. Moreover, one can easily check that $\left| \frac{\partial}{\partial b_l} V(b_l, b_r) \right|$ and $\left| \frac{\partial}{\partial b_r} V(b_l, b_r) \right|$ are bounded for all values of b_l, b_r, and the maximization problem (B2) is concave in b_l and concave in b_r. This ensures existence of equilibrium.

Consider at-large elections. For β sufficiently high, the utility of the median voter of electing (h, q) and $(0, -q)$ exceeds any other option (in particular, electing two council members of type $(0, 0)$); therefore, the competent type will be elected. The other council member may have ideal point other than $-q$, but it is determined uniquely because of concavity of (B2). Since $V(b_l, b_r)$ uniformly converges to 0, the type of the other council member must be arbitrarily close to $-q$ for β high enough.

Now consider district elections, and suppose that the competent citizen resides in district L. For β close to 1, district R will elect a council member with types exactly $(0, B)$. District L, following the logic of at-large elections, will either elect the competent citizen (h, q) or the extreme one, $(0, -B)$. As $\beta \to 1$, this two-way problem of the median voter m_L will converge to the problem he faces in the case where midpoint is selected automatically. This proves convergence in distribution. Furthermore, for β high enough, at-large elections will always result in election of the most competent citizen, while in district elections, this is not always the case (provided that $\frac{h}{k} < \frac{3}{4} B^2$, as in Proposition 3).

Now observe that for β sufficiently close to 0, council members of any type propose their

ideal points in equilibrium; this follows from Lemma A2, given that $P > 4kB^2$, which holds by assumption. Consequently, the utility of a citizen with ideal point b from a council with types $(a_l, b_l), (a_r, b_r)$ is

$$a_l + a_r - \frac{1}{2}k\left(b_l - b\right)^2 - \frac{1}{2}k\left(b_r - b\right)^2.$$

In at-large elections, one elected council member will have bliss point 0, and the competent citizen (h, q) will be elected if and only if $h \geq \frac{1}{2}kq^2$. In district elections, the problems of both districts are independent, and district L elects the competent citizen if and only if $h \geq \frac{1}{2}k\left(q + \frac{B}{2}\right)^2$; similarly, district R elects the competent citizen if and only if $h \geq \frac{1}{2}k\left(q - \frac{B}{2}\right)^2$. Therefore, the probability of electing the competent citizen in at-large elections is $\min\left(\sqrt{\frac{2h}{kB^2}}, 1\right)$; the corresponding probability in case of district elections is $\min\left(2\sqrt{\frac{2h}{kB^2}}, 1\right)$. The former is weakly less, and it is strictly less if $\frac{h}{k} < \frac{1}{2}B^2$. This completes the proof. ∎

B4 Proof of auxiliary results claimed in Footnote 22

Proof that if in at-large elections each voter may cast two votes for the same candidate, there may be multiple equilibria. This fact trivially follows from the result that we prove next. Indeed, suppose that parameter values are such that if citizens can cast only one vote, there are multiple equilibria. Take any such equilibrium σ, and consider strategy profile $\tilde{\sigma}$ where each citizen casts both votes for the same candidate he voted for under profile σ. Then $\tilde{\sigma}$ is an equilibrium in the game where two votes which may be cast for the same candidate. ∎

Proof that if in at-large elections each voter may cast only one vote, there may be multiple equilibria. Let us prove that for some parameter values, there are several equilibria. Suppose that the competent voter has bliss point q, and suppose that h is high enough (namely, $h \geq \frac{16}{9}kB^2$) Let us show that any pair of council members (h, q) and $(0, b)$ may be elected in equilibrium, provided that $|q + b| < \frac{2B}{3}$.

Consider an equilibrium where share ε of voters (where $\varepsilon > 0$ is small) vote for the competent citizen (h, q), and the rest vote for $(0, b)$; these two are then elected. The equilibrium policy in this case is $\frac{q+b}{2}$. The condition on h ensures that nobody wants to jeopardize election of a competent citizen. Indeed, a citizen with ideal point b_i gets $h - k\left(\frac{q+b}{2} - b_i\right)^2$; if a deviation prevents the

competent citizen from being elected, he will get at most 0. Since $\left|\frac{q+b}{2}\right| < \frac{B}{3}$ and $|b_i| \leq B$, $h - k\left(\frac{q+b}{2} - b_i\right)^2 > h - k\left(\frac{4}{3}B\right)^2 > 0$, and thus such deviation is not profitable.

It remains to consider the case where a coalition that plans to deviate and prevent $(0, b)$ from being elected must also ensure that it gives enough votes to the competent candidate (h, q) so that he is still elected. This implies that at least two-thirds of citizens must prefer electing of another incompetent citizen b' so that policy is $\frac{q+b'}{2}$ rather than $\frac{q+b}{2}$. This is only possible if $\frac{q+b}{2}$ lies outside of the interval $\left[-\frac{B}{3}, \frac{B}{3}\right]$ (otherwise no alternative is preferred by two-thirds). Therefore, if $|q + b| < \frac{2B}{3}$, no coalition will have a profitable deviation. This proves that there is a voting profile which constitutes an equilibrium, provided that $|q + b| < \frac{2B}{3}$, which completes the proof. ∎

Proof that if in at-large elections each voter has more than two votes which must be the cast for different candidates, there is a unique equilibrium. This proof is similar to the proof of Proposition 1 and is omitted. ∎

Proof that if in at-large elections voters vote for pairs of candidates, there is a unique equilibrium. It is trivial to show that a strategy profile where everyone votes for a pair of candidates $((h, q), (0, -q))$ is an equilibrium, because no majority has a profitable deviation (this follows from that this pair is a Condorcet winner). At the same time, if any other pair is elected in equilibrium, then there is a majority willing to deviate and cast all its votes for $((h, q), (0, -q))$. The proof of the latter fact is similar to the proof of Proposition 1 and is omitted. ∎

Appendix C: Additional Empirical Results

Table C1. Effect of Electoral Rules on Council Member Competence
(Including Female Council Members)

	Percent of Council Members who Finished High School							
	All Council Members				Female Council Members			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At-Large Elections	1.60**	-0.20	0.84	-0.95	-0.68	-0.38	-0.26	-1.02
	[0.74]	[0.99]	[0.65]	[0.95]	[0.49]	[0.63]	[0.40]	[0.83]
Fractionalized Project Preferences		3.70**				-0.62		
* At-Large Elections		[1.69]				[1.09]		
Fractionalized Project Preferences		-1.57				-0.11		
		[1.21]				[0.85]		
Ethnically Mixed Village			2.94				-1.94	
* At-Large Elections			[2.03]				[1.79]	
Ethnically Mixed Village			-0.90				1.50	
			[1.49]				[1.63]	
Geographically Large Village				5.19***				0.65
* At-Large Elections				[1.76]				[1.22]
Geographically Large Village				-1.98				-0.79
				[1.25]				[1.08]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,011	4,011	4,011	4,011	1,995	1,995	1,995	1,995
R-squared	0.11	0.11	0.11	0.11	0.08	0.08	0.08	0.08

Note: The unit of observation is council member. The dependent variable is a dummy variable that equals 100 if a council member finished high school and zero otherwise. Results in (1)-(4) based on a sample that includes both male and female council members. Results in (5)-(8) based on a sample that includes only female council members. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C2. Effect of Electoral Rules on Male Council Members' Competence

	Percent of Male Council Members who Finished Middle School				Percent of Male Council Members who Finished Primary School			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At-Large Elections	3.43*	-2.48	0.60	-0.17	2.43	-2.45	-2.19	-0.11
	[1.93]	[2.76]	[1.98]	[2.55]	[2.83]	[3.66]	[3.00]	[4.15]
Fractionalized Project Preferences		12.17***				10.03		
* At-Large Elections		[4.27]				[6.23]		
Fractionalized Project Preferences		-3.16				-3.84		
		[3.05]				[4.32]		
Ethnically Mixed Village			11.33**				18.57**	
* At-Large Elections			[5.49]				[7.75]	
Ethnically Mixed Village			-4.98				-8.27	
			[4.26]				[5.63]	
Geographically Large Village				7.24*				5.31
* At-Large Elections				[4.11]				[6.45]
Geographically Large Village				-4.38				-0.48
				[2.80]				[4.38]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
R-squared	0.19	0.19	0.19	0.19	0.26	0.26	0.26	0.26

Note: The unit of observation is male council member. The dependent variable is a dummy variable that equals 100 if a council member finished middle (primary) school and zero otherwise. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C3. Effect of Electoral Rules on Council Members' Competence
(as Measured by Occupation)

	Percent of Male Council Members who are Not Farmers			
	(1)	(2)	(3)	(4)
At-Large Elections	13.29***	8.63*	12.34***	11.91***
	[3.22]	[4.47]	[3.96]	[4.54]
Fractionalized Project Preferences		9.62		
* At-Large Elections		[6.97]		
Fractionalized Project Preferences		-6.75		
		[10.88]		
Ethnically Mixed Village			5.17	
* At-Large Elections			[6.66]	
Ethnically Mixed Village			-11.04	
			[11.04]	
Geographically Large Village				2.71
* At-Large Elections				[6.77]
Geographically Large Village				-5.76
				[1.96]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,044	2,044	2,044	2,044
R-squared	0.19	0.20	0.19	0.19

Note: The unit of observation is council member. The dependent variable is a dummy variable that equals 100 if the main occupation of the council member is farmer. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C4. Effect of Electoral Rules on Council Member Competence
(Excluding Districts with More than One Member Elected)

	Percent of Council Members who Finished High School			
	(1)	(2)	(3)	(4)
At-Large Election	4.04***	0.36	2.46*	-1.60
	[1.43]	[1.94]	[1.40]	[1.56]
Fractionalized Project Preferences		7.75**		
* At-Large Elections		[3.51]		
Fractionalized Project Preferences		-2.58		
		[2.02]		
Ethnically Mixed Village			6.11*	
* At-Large Elections			[3.21]	
Ethnically Mixed Village			-2.06	
			[2.15]	
Geographically Large Village				11.98***
* At-Large Elections				[3.15]
Geographically Large Village				-3.25
				[2.02]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,716	1,716	1,716	1,716
R-squared	0.20	0.20	0.20	0.21

Note: The unit of observation is council member. The dependent variable is a dummy variable that equals 100 if a council member finished middle school and zero otherwise. The sample excludes observations from districts in which more than one candidate was elected to the council. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C5. Electoral Rules and Incumbency Advantage

Definition of Elite	Percent of Male Council Members Considered As Pre-Existing Elite		Percent of Pre-Existing Elite Elected to Council	
	Mean in District Elections	Difference between At-large and District	Mean in District Elections	Difference between At-large and District
Member of Baseline Focus Group (including Non-Attendees)	31.9	2.43	25.8	1.55
		[2.23]		[1.89]
Observations in District Villages	1055		1,301	
Observations in At-large Villages	1003		1,293	
Decision-Maker According to Male Focus Group	13.2	-0.54	43.5	-3.40
		[1.27]		[3.79]
Observations in District Villages	1055		317	
Observations in At-large Villages	1003		308	
Decision-Maker According to Male Head-of-Household Survey	20.7	3.24*	27.8	1.75
		[1.86]		[2.24]
Observations in District Villages	1055		784	
Observations in At-large Villages	1003		868	
Decision-Maker According to Female Individual Survey	14.9	-0.66	30.2	-2.13
		[1.53]		[2.93]
Observations in District Villages	1055		523	
Observations in At-large Villages	1003		501	
Either of the Four Above	38.9	3.2	21.1	1.21
		[2.39]		[1.47]
Observations in District Villages	1055		1935	
Observations in At-large Villages	1003		1969	

Note: The difference between district and at-large elections estimated using the same model as in regression (1). Only male council members are considered. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Figure C1. Cumulative Distribution Function for the Days between Elections
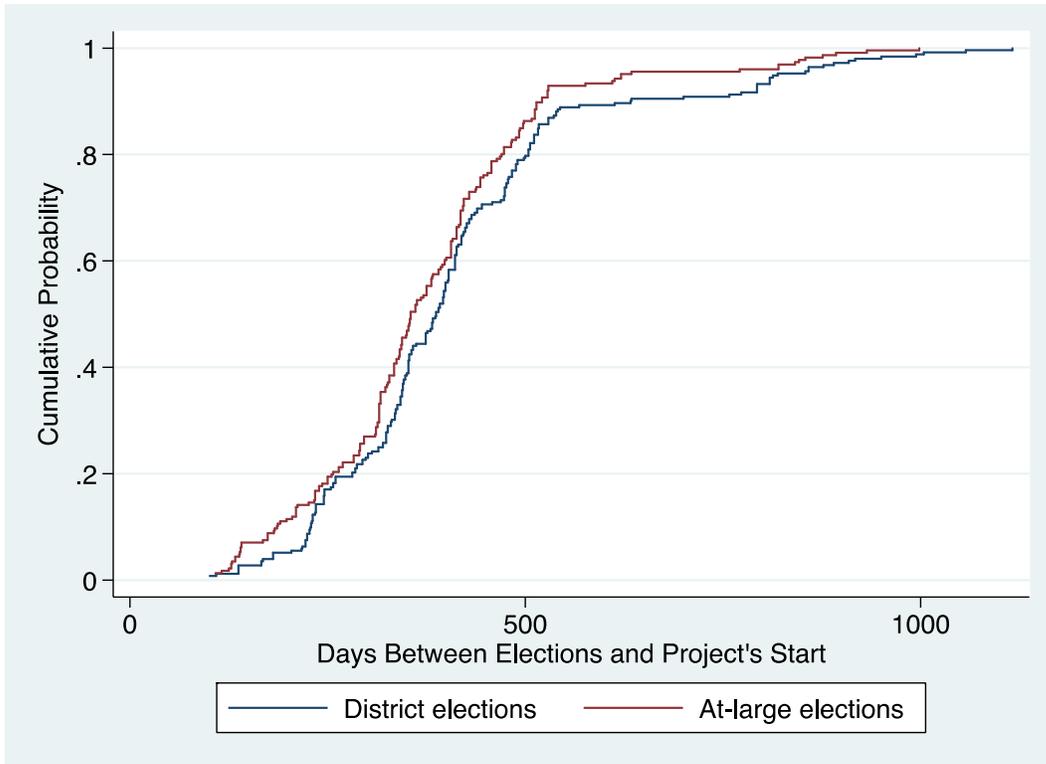and Project's Start by Type of Elections.

Figure C2. Cumulative Distribution Function for the Days between Elections
and Project's Completion by Type of Elections.