

NBER WORKING PAPER SERIES

ELECTORAL RULES AND THE QUALITY OF POLITICIANS:
THEORY AND EVIDENCE FROM A FIELD EXPERIMENT IN AFGHANISTAN

Andrew Beath
Fotini Christia
Georgy Egorov
Ruben Enikolopov

Working Paper 20082
<http://www.nber.org/papers/w20082>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2014

We thank Ernesto Dal Bó, Hülya Eraslan, Horacio Larreguy, Tommaso Nannicini, Mattias Polborn, Mohamed Saleh, and participants of the NBER Summer Institute Political Economy Public Finance meeting, Priorat Workshop on Bargaining and Politics, Political Economy in the Chicago Area conference and Elections and Electoral Institutions conference in Toulouse for helpful comments. The authors would like to acknowledge the generous cooperation and assistance provided by H. E. Wais Barmak, Tariq Ismati, and Abdul Rahman Ayubi of the Ministry of Rural Rehabilitation of the Government of the Islamic Republic of Afghanistan; Ehsan Zia; staff of AfghanAid, C.H.A., InterCooperation, IRC, NPO/RRRAA, Oxfam UK, and People-in-Need; and Philippe Dongier, Susanne Holste, Qazi Azmat Isa, Zishan Karim, Norman Piccioni, and Mio Takada of the World Bank. We thank Anna Belong and Denis Shishkin for excellent research assistance and Hamidullah Gharibzada, Shahim Kabuli, and Maiwand Siddiqi for excellent oversight of data collection. Data collection was supported by the Ministry of Rural Rehabilitation and Development of the Government of the Islamic Republic of Afghanistan through the National Solidarity Programme. Additional financial and logistical support for the study was provided by the Food and Agriculture Organisation of the United Nations and the World Bank. Fotini Christia acknowledges support from ARO MURI award No. W911NF-12-1-0509. Ruben Enikolopov acknowledges financial support from the Center for the Study of Diversity and Social Interactions at NES and financial support through Deutsche Bank Membership during his stay at IAS. Andrew Beath was a consultant for the World Bank during the data collection process and was a staff member of the World Bank thereafter. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors, the countries they represent, or the National Bureau of Economic Research. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Andrew Beath, Fotini Christia, Georgy Egorov, and Ruben Enikolopov. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

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

Andrew Beath, Fotini Christia, Georgy Egorov, and Ruben Enikolopov

NBER Working Paper No. 20082

May 2014

JEL No. D72,D78

ABSTRACT

We examine the effect of electoral rules on the quality of elected officials using a unique field experiment which induced randomized variation in the method of council elections in 250 villages in Afghanistan. In particular, we compare at-large elections, with a single multi-member district, to district elections, with multiple single member districts. We propose a theoretical model where the difference in the quality of elected officials between the two electoral systems occurs because elected legislators have to bargain over policy, which induces citizens in district elections to vote strategically for candidates with more polarized policy positions even at the expense of candidates' competence. Consistent with the predictions of the model, we find that elected officials in at-large elections are more educated than those in district elections and that this effect is stronger in more heterogeneous villages. We also find evidence that elected officials in district elections have more biased preferences.

Andrew Beath
World Bank
One Global Place, 26th Floor
5th Ave. corner 25th St.,
Bonifacio Global City
Taguig City 1634
Philippines
abeath@worldbank.org

Fotini Christia
Department of Political Science
Massachusetts Institute of Technology
fotini.christia@gmail.com

Georgy Egorov
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
g-egorov@kellogg.northwestern.edu

Ruben Enikolopov
Barcelona IPEG
Universitat Pompeu Fabra
Barcelona GSE, New Economic School
Edif. Mercè Rodoreda 23.105 (IPEG)
C/ Ramon Trias Fargas, 25-27
08005 Barcelona, Spain
ruben.enikolopov@upf.edu

1 Introduction

Electoral systems constitute a fundamental aspect of constitutional arrangements. To date, numerous studies have assessed the effect of electoral systems on electoral and policy outcomes (see, e.g., Cox, 1997; Norris, 2004; Persson and Tabellini, 2000, 2005). Particular attention has been focused on how electoral systems alter political incentives and representation. However, the difficulty of identifying exogenous sources of variation in electoral systems has limited the ability of studies to make causal inferences. This paper contributes to the literature by exploiting a field experiment to causally identify the effect of electoral systems, specifically focusing on the understudied question of how differences in systems affect the quality of elected officials.¹

The paper combines a theoretical model with a field experiment that induced randomized variation in electoral rules for local council elections across 250 villages in Afghanistan. Specifically, councils were elected using one of two alternative systems with large differences in district magnitude. Per the “district election” system, the village is divided into several single-member districts and candidates are elected from each district separately, with villagers only voting for those candidates who live within their assigned district. Per the “at-large election” system, the whole village constitutes one multi-member district, so that council members are elected based on the number of votes garnered across the whole village, with villagers having no restrictions for whom they can vote. Under both electoral rules, all villagers are automatically considered candidates.

The theoretical model considers a linear village with a uniform distribution of villagers. The location of a villager’s home corresponds to his/her ideal point for the location of a public good (e.g., a drinking well). In line with the experiment, we consider a citizen-candidate model with no entry costs for the candidates. Voters elect a two-member council in which officials bargain over a joint policy decision. In at-large elections, each citizen has two votes and can vote for any two candidates.² In district elections, the village is split geographically into two districts, and each citizen can vote for a candidate residing in the same district. The number of competent citizens is assumed to be small and voters thereby trade-off candidates’ policy preferences with candidates’

¹See theoretical work by Myerson (1993) for a notable exception. 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).

²This is in line with Cox (1984), who was the first to study double member districts formally. The difference is that in Cox (1984), voters vote sincerely and candidates choose positions strategically; here, voters are strategic, but candidates’ positions are fixed (and tied to their locations).

quality. The model predicts that, in district elections, voters prefer candidates with more extreme preferences over competent candidates, as they expect them to achieve a better outcome in the bargaining game against candidates elected from other districts. In at-large elections, voters are less bound by strategic bias considerations and are thus more willing to elect competent candidates. The difference between the two electoral systems becomes more pronounced in villages with more heterogeneous policy preferences.

The predictions of the theoretical model prove to be consistent with the empirical evidence provided by the field experiment. In particular, we find that at-large elections result in the election of council members with higher levels of education.³ 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 extreme policy preferences, as proxied by the location of their houses. We show that the results are not consistent with a number of alternative explanations. In particular, they are not driven by 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 minority (Trebbi, Aghion, and Alesina, 2007). Studies generally argue 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).⁴

³Existing literature suggests that education is an appropriate measure for the quality of politicians, as highly 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), and provide higher quality of public goods (Martinez-Bravo, 2013).

⁴Persson 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

While there is an extensive theoretical literature that examines the determinants of the quality of politicians (McKelvey and Reizman, 1992; Banks and Sundaram, 1998; Aragonés and Palfrey, 2004; Caselli and Morelli, 2004; Mattozzi and Merlo, 2007; Acemoglu, Egorov, and Sonin, 2010; Egorov and Sonin 2011), only Myerson (1993) considers the effect of electoral system, arguing that small district magnitude together with strategic voting increases the barriers to entry in the electoral system, which in turn has a negative effect on the quality of politicians.⁵ The empirical literature on the determinants of the quality of public officials finds that it is positively affected by increasing intra-party competition (Besley, Folke, Persson, and Rickne, 2013; Folke, Persson, and Rickne, 2014) as well as inter-party competition (Banerjee and Pande, 2007), and by higher wages (Ferraz and Finan, 2011; Dal Bó, Finan, and Rossi, 2013; Gagliarducci and Nannicini, 2013), whereas higher budgets attract politicians of lower quality (Brollo, Nannicini, Perotti, and Tabellini, 2013). However, except for the finding that democracies are more likely to select highly educated leaders (Besley and Reynal-Querol, 2011), there is no evidence on the effect of political institutions on the quality of politicians.

In contrast with previous works, this paper combines an original theoretical model with a large-sample field experiment. The model makes a theoretical contribution by showing that electoral rules can affect the quality of elected officials even in the absence of entry barriers, party politics, or specific assumptions about the sizes and preferences of the minority groups. The paper also provides an empirical contribution by providing what is, to the best of our knowledge, the first experimental evidence on the causal effects of electoral rules.

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 robustness of predictions; Appendix B contains all the proofs; and Appendix C contains additional empirical results.

at the effect of electoral rules on the composition of government spending include Lizzeri and Persico (2001) and Milesi-Ferretti, Perotti and Rostagno (2002).

⁵See also Adams (1996), who considers the effect of a constitutional change in Illinois in 1980 on business friendliness of state legislators.

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, Christia and Enikolopov, 2013b), as well as the program itself (Beath, Christia, and Enikolopov, 2012, 2013a). This section provides further details on NSP (subsection 2.1), describes the variation in electoral rules 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 Setting

Elections of village development councils are mandated by the National Solidarity Program (NSP), which was devised in 2002 by the Government of Afghanistan to deliver services and infrastructure to the country’s rural population and build representative institutions for village governance. NSP has been implemented in over 32,000 villages in all of Afghanistan’s 34 provinces and has disbursed over \$1.1 billion, making it the largest development program in Afghanistan. The program is structured around two interventions: (i) the creation of an elected Community Development Council (CDC); and (ii) the disbursement of block grants to councils for implementation of village projects. The program is executed by the 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.⁶ 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,⁷ with villages required to contribute at least 10 percent of project costs, which they largely do in the form of labor. Projects are selected by the council in consultation with the village community. Projects are ordinarily focused on either the construction or rehabilitation of infrastructure, such

⁶Note 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)

⁷The average block grant in the villages included in the sample was approximately \$31,000.

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 to 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 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 the follow-up elections are not facilitated, it is unclear whether these actually occur.⁸ Given this, and the general uncertainty that accompanies planned future development activity in Afghanistan, villagers perceive NSP as a one-shot event, which does not provide strong re-election incentives to council members.

The average population in our sample of villages is roughly one thousand people (see Table 1). There is 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. These villages are also notably poor: only forty-five percent of respondents indicate that they never or rarely have problems supplying food for their families.

2.2 Local Governance in Afghanistan

Afghanistan’s central government has historically lacked the strength and resources to exercise local control or provide public goods in many parts of the country. As a result, local communities have had their own structures of governance and accountability (Barfield, 1984). The foundation of governance in rural Afghanistan is the local *jirga* or *shura*, a participatory council that has traditionally managed local public goods and adjudicated disputes (Nojumi, Mazurana and Stites,

⁸No such elections had occurred by early 2012, when the data collection for this evaluation was completed.

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 their decisions based on consensus (Boesen, 2004). In addition to councils, villages ordinarily have a headman (termed a *malik*, *arbab*, or *qariyadar*) – usually a large landowner – who serves as liaison between the village and the central government (Kakar, 2005). Traditional leadership in rural Afghanistan consists almost exclusively of males, as the principle of *pardah* – which stipulates that women should generally be hidden from public observation – precludes female involvement in communal gatherings and thus from local governance.

A key contrast between elected development councils and customary governance institutions is the mode of selection and 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, although the creation of an elected development council does not directly usurp the major administrative tasks undertaken by the headman or other customary village institutions, the elected development council exists in parallel with customary governance structures as an institution vested with substantial authority.

2.3 Electoral Rules

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

District Election: Voters were restricted to casting a ballot for a single candidate, who had to reside in the same district. In each district, 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 regardless of district. 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.¹⁰ The at-large election method represents a multi-member election under a plurality rule with a single district and multiple non-transferable votes. Thus, the two main differences from district elections are: (i) higher district magnitude (multiple elected members instead of one) and (iii) number of votes cast (three instead of one).

In all villages, council elections were organized and administered by “social organizers” employed by the contracted NGOs facilitating NSP in the district. 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.¹¹ 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

⁹A detailed guide on the procedures is available at: <http://www.nsp-ie.org/sti.html>

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

¹¹A detailed description of the monitoring results can be found at: <http://www.nsp-ie.org/reports/CDCE-MR.pdf>

(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.¹²

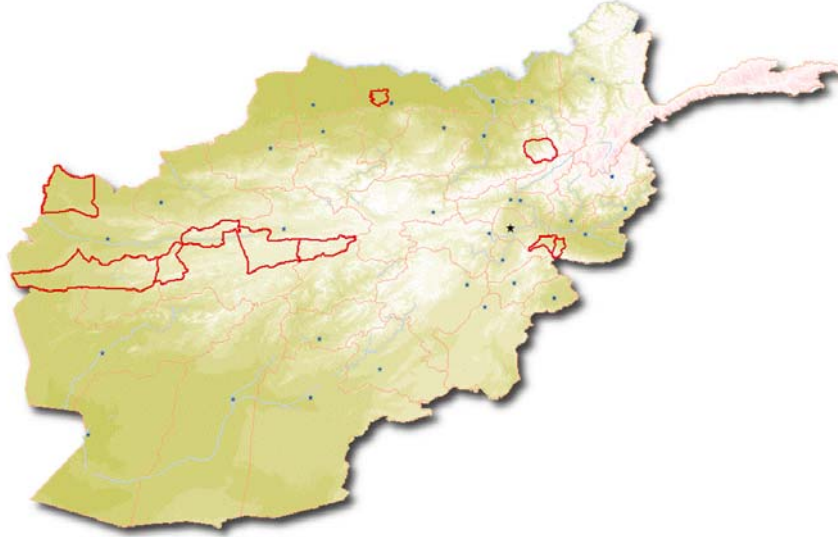


Figure 1: Ten sample districts.

Random assignment of electoral rules was made concurrently with the assignment of the two alternative project selection procedures.¹³ 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.¹⁴ 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.

This assignment procedure ensures that each village in the sample had an equal probability

¹²An assessment of the demographic and economic characteristics of the sample villages reveal few substantive differences with those of a random sample of villages surveyed by the 2007-08 National Risk and Vulnerability Assessment.

¹³Development projects were selected either through a secret-ballot referenda or at a village meetings. For more on those results, see Beath, Christia and Enikolopov (2013b).

¹⁴These 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.

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, 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 the processes. Project selection occurred between November 2007 and August 2008, whereas project implementation occurred between April 2008 and September 2011.

3 Theory

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

3.1 Setup

The society consists of a continuum of individuals distributed uniformly on a compact $S = [-B, B]$. The policy space that these agents care about coincides with 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.¹⁵ 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, e.g., by providing the public good better or earlier, which is consistent with the data. If the two council members have types (a_l, b_l) and (a_r, b_r) and implement policy p , then individual i will get utility:

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

To simplify exposition, we assume that the council members pick a policy that maximizes their joint utility: $p = \frac{b_l + b_r}{2}$. Notice that this policy will be the outcome of a bargaining game with alternating offers (Rubinstein, 1982) or a legislative bargaining game with random recognition (Baron and

¹⁵Given the context of the experiment, it is natural to assume that location of villagers’ dwellings and education are observable to fellow villagers.

Ferejohn, 1989; Banks and Duggan, 2000) in the limit where offers are made very frequently.¹⁶ We consider such 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, i.e., every individual living in the district is considered a candidate, again in line with the experimental setup.¹⁷ Then in each district the individual who got the largest share of votes is elected, and in the case of a draw, a random person among those who received the most votes is chosen. In at-large elections, the entire society comprises a single district, and each individual casts two votes for two (different) citizens. The two candidates who received the most votes are elected.¹⁸

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

¹⁶If 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).

¹⁷Our model of elections thus fall 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).

¹⁸We will prove that in at-large elections there is a Condorcet winner, i.e., a pair of citizens such that there is no other pair that a majority would prefer to be elected. The voting model we picked ensures that the Condorcet winner pair is elected. The assumption that voters cannot cast both their votes for the same candidate is important here; if they could, 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 aggregation of 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.)

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 Strong Nash.

3.2 Analysis

Our analysis of the game is greatly simplified by the fact that the median voter theorem applies in both at-large and in 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 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 (below) fills in the details.

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

1. *In district elections, the district without the competent citizen elects the most biased individual (with $b_i = \pm B$), and the district with the competent citizen (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 types of citizens who 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, 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:

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

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.¹⁹ 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 committed type.²⁰ 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:*

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

¹⁹This 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.

²⁰This 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).

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 ;
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 they are especially easy to see on Figure 2, where we depicted the outcome of elections (types of council members) for different realizations of

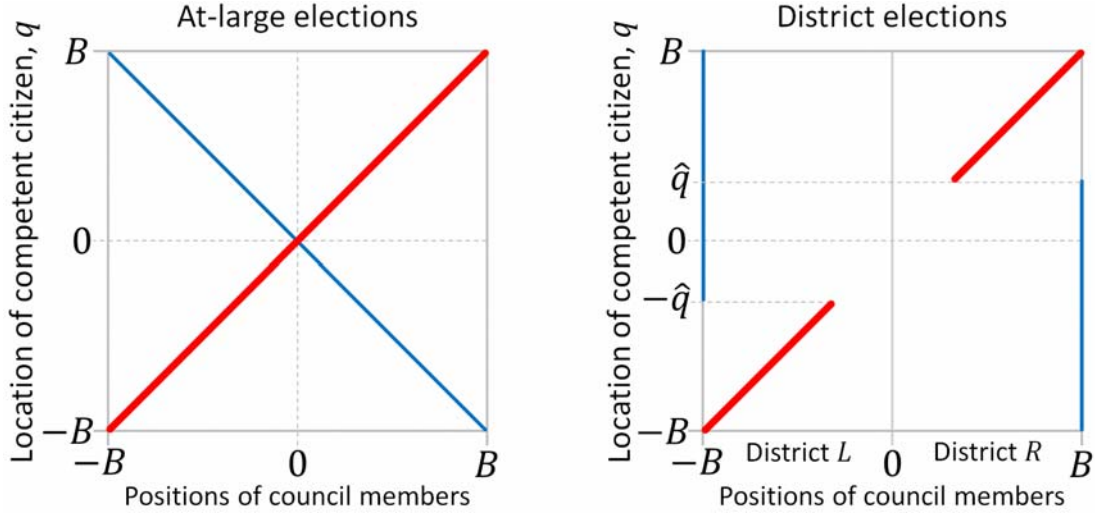


Figure 2: Types of council members (competent in red/thick, incompetent in blue/thin) as a function of the competent citizen’s ideal point.

q . In at-large elections, the competent individual is always elected, and this is not true in district elections, provided that $\hat{q} > 0$. Interestingly, if h is high or k is low or B is low, the competent individual will be elected in both cases, and the difference between the two types of elections disappears. The result on polarization is easy to see from the following consideration: the two council members elected in at-large elections are as far from the median as the competent one, while in district elections, one or both districts elect individuals who are further from the median than the competent individual. In addition, if $\hat{q} > 0$, then in district elections, the most moderate types will never be elected. Finally, in at-large elections, there is no correlation between preferences and competence of a council member: as one can see from Figure 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.²¹

Apart from the empirical predictions about competence and polarization, our model has clear welfare implications.

²¹The 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.

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

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

4 Empirical Predictions

The theoretical results generate several empirical predictions, which we test using data from the field experiment. It should be emphasized that only empirical results for the first prediction had been 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 can formulate the following empirical prediction:

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

Following Besley, Pande, and Rao (2005), we use educational attainment as a measure of candidate quality, since there is evidence that leaders' level of education has a positive effect on governance outcomes. In Section 6 we show that educational attainment is associated with better council performance in the context of our experiment.

The first statement of Proposition 3 also asserts that the effect of the electoral system on the quality of politicians is stronger in communities with more diverse preferences (higher B) 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 between electoral systems. Thus, we can formulate the following two empirical predictions:

2. The difference in the quality of council members between at-large and district elections is higher in more heterogeneous villages.
3. In homogenous villages, the quality of council members does not depend on the electoral method.

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

Fractionalization of preferences over projects directly measures heterogeneity of interests in terms of project type. More diverse villagers’ preferences regarding the project type correspond to higher B in the model. In addition, quality of implementation is likely to be more important if preferences are aligned and people get their preferred project; at the same time, if preferences are diverse, then policy choice becomes more important. Thus, fractionalized preferences correspond to the case where policy is important relative to competence, i.e., high k . Ethnic heterogeneity is often used in the literature to capture differences in tastes (e.g. Alesina, Baqir, and Easterly, 1999) that might affect preferences over both project type²² and location (if villages are ethnically segregated),²³ 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).

According to the model, the quality of elected candidates improves outcomes for any policy choice. In the context of the experiment, the quality of elected candidates is likely to affect project

²²This 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).

²³Unfortunately, the number of observations within villages is not large enough to test for the existence of segregation directly.

implementation regardless of their type and location. In particular, council members of high quality are likely to implement projects faster. Thus, we can formulate the following empirical prediction:

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

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

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

In the model, polarization corresponds to the distance between council member preferences and the median voter. In the empirical analysis, we focus on candidate preferences with respect to project location, taking the assumption that the geographic location of elected candidates and voters reflects their policy preferences literally. 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.

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

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

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

5 Data

The data used in the empirical analysis come from several different sources. Information on the characteristics of elected council members was supplied by the contracted NGOs. Specifically, the data contains information on electoral results and personal characteristics of all council members, including their gender, age, educational attainment, and occupation, as well as their district of

residence. Data was provided for 2,044 male council members from 241 villages.²⁴ We construct a dummy variable that equals one if the person has full secondary education (i.e., finished high school) and zero otherwise and a similar measure for getting lower secondary education (i.e., finishing middle school). Among male council members only 9% have finished high school, and 17% have finished middle school (see Table 2).

Information on the geographic distribution of elected council members is based on GPS coordinates of their residences collected during the monitoring of program implementation. This information is available only for a subset of villages, since some monitors failed to follow the instructions and did not collect the required information. In addition, some coordinates contained obvious mistakes and were excluded from the analysis.²⁵ The resulting database contains coordinates of the residences of 1,104 male council members in 140 communities in nine out of ten evaluation districts. Although sample attrition is substantial, there are no reasons to believe that it was systematic in a way that can affect the corresponding results.

We restrict our analysis to male council members for two reasons. First, in Afghan villages, as discussed above, women are traditionally excluded from community-level decision making. The female council members elected through the gender quota are therefore not expected to have a significant effect on project selection and implementation.²⁶ Second, the average level of education of female council members was so low that there was almost no variation in educational attainment. In particular, more than 90% of female candidates did not have any formal education and only 0.8% had finished high school.

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

²⁴Of the 9 villages for which the data was not received, 7 villages did not comply with the NSP treatment assignment, an error driven primarily by confusion between villages with similarly-sounding names. For the remaining 2 villages, the relevant NGO did not provide the necessary information. In both cases, attrition was not correlated with the assigned electoral rules.

²⁵These 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.

²⁶This is confirmed by the results in Beath, Christia and Enikolopov (2013b) which indicate that the preferences of even the most important female villagers do not have a significant effect on the choice of projects, whereas preferences of the male villagers are highly correlated with the subsequent choice of development projects.

socio-economic characteristics for 2,387 male heads of household.

To construct a measure of heterogeneity of preferences with regard to the types of development projects, we use a question from the baseline survey in which respondents were asked to indicate – from a list of fifteen potential projects that correspond to the types of projects sponsored by NSP – the project that they believed should be selected if the village was provided with a \$60,000 grant. In each village v for each type of project j we calculate the share s_{vj} of respondents that indicated this project and calculate fractionalization of preferences in the village $f_v = 1 - \sum s_{vj}^2$. Next, we construct a dummy variable, which equals one if fractionalization is above the median in the sample and zero otherwise. We also calculate the average distance between the households of the respondents based on the GPS coordinates of their houses and construct a dummy variable for large villages which equals one if the average distance is above the median in the sample and zero otherwise. As a measure of ethnic heterogeneity of villages, we use a dummy variable that equals zero if all villagers belong to one ethnicity and one otherwise.²⁷

Information on project implementation comes from the administrative records of NSP’s central office. The data contains information for 595 projects that were implemented in the villages as part of NSP and includes the dates when implementation of each project started and when it ended. The start dates range from April 2, 2008 to February 22, 2011; the completion dates range from July 23, 2008 to September 24, 2011.²⁸ Based on this information, we construct five dummy variables for whether a project started before the beginning of the fourth quarter of 2008 and each of the quarters in 2009. Similarly we construct five dummy variables for whether a project started before the beginning of each of the quarters in 2009 and the first quarter of 2010.

6 Empirical Results

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

²⁷The 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.

²⁸Six projects failed and were not completed.

model:²⁹

$$Progress_{jv} = \alpha + \beta \cdot EducationCouncil_v + \gamma \cdot EducationVillagers_v + \phi_q + \varepsilon_{jv}, \quad (3)$$

where $Progress_{jv}$ is an indicator for whether project j in village v has started or finished by a specific date, $EducationCouncil_v$ is an indicator for whether at least one council member has finished high school (which is true for 32% of the villages), $EducationVillagers_v$ is an indicator for whether at least one of the villagers in the baseline survey has finished high school (which is true for 26% of the villages), ϕ_q is the quadruple fixed effect, and ε_{jv} is the error term.

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

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

$$Education_{vi} = \alpha + \tau \cdot AL_v + \phi_q + \varepsilon_{vi}, \quad (4)$$

where $Education_{vi}$ is a dummy variable for whether council member i in village v has finished high school,³⁰ 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 $\tau > 0$.

Results of the estimation confirm that male council members elected through at-large elections are better educated as compared with male council members elected through district elections (see column 1 in Table 3). The share of council members who finished high school among the council

²⁹In all specifications with binary dependent variables we use linear models, since our specifications are close to a saturated model (Angrist and Pischke 2008), while linear specification makes interpretation of the coefficients easier and avoids the incidental parameter problem caused by the inclusion of quadruple fixed-effects. The results are robust to using a logit model instead.

³⁰The results are robust to using a dummy variable for finishing middle school (see Table C1 in Appendix C).

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

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

$$Education_{vi} = \alpha + \gamma \cdot AL_v + \delta \cdot AL_v \cdot Heterogen_v + \mu \cdot Heterogen_v + \phi_q + \varepsilon_{vi}, \quad (5)$$

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

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

According to the fourth empirical prediction, the speed of project implementation should be higher in at-large elections as compared with district elections. To test this hypothesis we estimate the following OLS model:

$$Progress_{jv} = \alpha + \lambda \cdot AL_v + \phi_q + \varepsilon_{jv}, \quad (6)$$

where all the variables are the same as in (3) and (4). According to the fourth empirical prediction $\lambda > 0$. The results indicate that in villages with at-large elections, project implementation was more likely to start earlier and be completed faster (see Table 5). The magnitude of the effect is noticeable, with projects being 9 percent more likely to have started before April 2009 and 5 percent more likely to have finished before July 2009. Figures C1 and C2 in Appendix C show the distribution of dates for project commencement and completion by type of elections.³¹ The results are robust to controlling for the type of implemented projects, so that the effect is not driven by the choice of different types of projects in villages that held at-large elections. This is consistent with the finding that there is no significant effect of the electoral system on the type of selected projects, which conforms with the prediction of the model that, on average, there is no difference in policies implemented under at-large and district elections. Thus, the results are in accordance with the fourth empirical prediction.

If the electoral system affects project implementation only through its effect on the education of council members, one can estimate the causal effect of council members education on the speed of project implementation using an IV specification in which electoral system is used as an instrument for the education of council members. The results of such an estimation are generally higher in magnitude than the result of the respective OLS estimation (see Table C2 in Appendix C for the results). These results, however, should be taken with extreme caution, as the exclusion restriction is unlikely to hold, since the electoral system can affect project implementation through other channels. For instance, greater difference in preferences of elected candidates in district elections, as predicted by the theoretical model, may cause a gridlock, which would delay the implementation of projects.

The fifth empirical prediction posits that district elections will lead to the election of candidates with more biased preferences. We test this prediction by estimating a model similar to (4) with the logarithm of the distance between the house of an elected council member and the village center (calculated by averaging the coordinates of the houses of baseline survey participants) as the outcome variable.³² To account for differences in the size of the village, we can include a measure of

³¹There is no effect on the duration of implementation of projects, which might be driven by the fact that more educated council members choose more complicated projects, which is partially confirmed by the fact that villages with more educated council members tend to implement a smaller number of more expensive projects.

³²We winsorize the distance at the 95 percent quintile to limit the influence of potential outliers.

the distance between the villagers’ homes in the baseline survey as an additional control. The results indicate that consistent with the empirical prediction, the distance between the homes of elected officials and the village centers is smaller in at-large elections (see Table 6). Not surprisingly, the distance between the homes of elected officials and village centers is larger in villages with greater distance between the houses of baseline respondents.

Finally, according to the sixth empirical prediction, we should observe a negative correlation between a candidate’s quality and the extent of his bias. We test this prediction by estimating the following model:

$$Education_{vi} = \alpha + \varphi \cdot AL_v + \rho \cdot AL_v \cdot Distance_{vi} + \kappa \cdot Distance_{vi} + \phi_q + \varepsilon_{vi}, \quad (7)$$

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

The results indicate that in district elections there is indeed a negative, although not statistically significant, correlation between an elected candidate’s education and the distance between his house and the center of the village (see Table 7). At the same time, the difference between the effects of distance in villages with at-large and district elections is indeed positive and statistically significant in all specifications, confirming that $\rho > 0$. We also document that the correlation between education and distance to the center of the village in at large elections is positive (rather than zero, as implied by the model), although it is statistically significant only if we do not control for village size.³³ Thus, we find limited support for the sixth empirical prediction.

7 Discussion

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

³³We also test and find that for the random sample of villagers in the baseline survey there is no significant relationship between their education and the distance from their houses to the center of a village.

that in more heterogeneous villages, district elections lead to the election of candidates with more biased preferences at the expense of their quality.

There are several possible alternative explanations for the positive effect of at-large elections on the quality of candidates. First of all, this effect can be driven by the restriction on the residence of candidates in district elections. If there are two high-quality candidates that live in the same district, only one of them can be elected under district elections, but both of them can be elected in at-large elections. This restriction can have a negative effect on the quality of elected candidates in a context where candidates of good quality are scarce. To see if the difference in the quality of the elected council members is driven by this restriction, we look at the distribution of elected council members across districts. Villages with district elections had exactly one male candidate elected to the council from each district. Although in at-large elections there were no formal restrictions on the distribution of candidates across districts, the distribution turned out to be not much different from that in villages with district elections. Specifically, in villages with at-large elections, 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. Of these, in 25 villages there was only one district that did not have a resident in the council. Thus, the negative effect of at-large elections on the probability of a district having a resident council member was very small.

To further address this point, we exclude from the sample candidates in at-large elections from districts that had more than one candidate elected to the council. Thus, we look only at the quality of candidates for whom the restriction on the number of candidates from the same district was not binding. Although this restriction is endogenous and these results should be interpreted with caution, it provides some evidence regarding the robustness of our results to this alternative explanation. The results in Table C3 in Appendix C indicate that the effects obtained in the benchmark specification are robust to such a sample restriction. Finally, the restriction on candidates' residence in district elections should 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, the empirical results suggest that the effect of electoral rules on the quality of elected candidates is not driven by the restriction on candidates in district elections.

Another potential explanation is that an increase in district magnitude in at-large elections makes it harder for the incumbent to coordinate voting, which should reduce the incumbency advantage of the members of the pre-existing elites and increase the quality of candidates. To examine the potential role of incumbency, we look at the share of council members who were identified as pre-existing elite members. Specifically, as the results in Table C4 in Appendix C indicate, 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. While sizable, this result suggests that pre-existing elites do not dominate the elected council. Importantly, electoral rules have no significant effect on the share of pre-existing elites among council members. The only marginally significant result is that the share of council members who were named ex-ante as the main decision makers is somewhat higher in villages with at-large elections. We obtain similar results if we use an alternative measure of incumbency advantage and look at the proportion of pre-existing elite members who subsequently were elected to the council. The share of elite members elected to the council varies between 19 and 44 percent depending on the measure, but again there is no significant difference between villages that used alternative electoral rules. Overall, results indicate that electoral rules have no significant effect on incumbency advantage.

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

In general, the difference in the quality of elected representatives can also reflect barriers to entry, which are higher in electoral systems with small district magnitudes (Myerson, 1993). However,

this explanation is not relevant in our context, as all citizens are considered candidates and there are no entry barriers.

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

It is also true that the consequences of having high quality council members, especially in at-large elections, may go beyond better implementation of development projects. More educated council members may also improve villagers' attitudes toward elected local leaders, making them more likely to support elections as the preferable method of selecting local leaders. In Table C5 in Appendix C, we provide some evidence that at-large elections have a positive effect on villagers' attitudes toward local leaders and elections as a method of selection of the village head. These results are broadly consistent with the model's prediction that at-large elections lead to higher social welfare.

8 Conclusion

In this paper, we examine the effect of electoral rules on the quality of elected representatives. We consider two alternative electoral rules – district and at-large elections. We provide a theoretical model in the tradition of citizen-candidate models with free entry of candidates, and show that at-large elections should lead to better educated council members, and that this effect is stronger in

more heterogeneous communities. Moreover, the model predicts that the candidates elected under district elections should be more biased relative to those elected under at-large elections. In the model, these effects are driven by the voters' strategic incentive to elect more biased politicians in district elections, in anticipation of the legislative bargaining process in the council. Our empirical results from a field experiment in 250 villages in Afghanistan are consistent with the theoretical predictions of the model.

The suggested mechanism is obviously not the only way through which electoral rules may affect the quality of elected officials and, in particular, their level of education. There is, however, substantial evidence that the mechanism we describe is driving the identified results. First, the design of the experiment rules out alternative possible channels such as barriers to entry or comparative advantage in political campaigning. More importantly, we find a direct link between the type of election and location of council members' residences, which at least implies that geographical considerations played a role in voting decisions. Given the decisions that elected officials had to make on the type and location of a public project, the strategic voting / delegation mechanism implied by the model appears to be the simplest and most natural mechanism consistent with the available empirical evidence.

The unique conditions of our experiment (first-time local elections; absence of political parties, campaigns or canvassing; a single question that the council had to decide) allow us to not only establish the causal link between the type of elections and quality of elected representatives, but also to identify a specific mechanism. However, these unique conditions may also give rise to questions of external validity. Such questions are reasonable and it is feasible that the identified mechanism may not operate in a similar manner in contexts with different characteristics. At the same time, our paper obtains two results we believe to be important in other contexts. First, the quality of elected politicians depends on the electoral rule. Second, voters may vote strategically, taking into account not only their delegate's position, but also anticipating collective bargaining in a decision-making body. Understanding the impact of electoral rules on election outcomes and, in particular, the quality of politicians, in more general settings appears to be a fruitful avenue for future work.

References

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

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

- 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.
- 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 (2011) "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.
- 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 Problems 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 (2013) “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: Feinstein International Famine Center, Youth and Community Program, Tufts University.
- Norris, Pippa (2004) *Electoral Engineering: Voting Rules and Political Behavior*. New York: Cambridge University Press.
- 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.
- 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.

Table 1. Balance of Pre-treatment Covariates

	Mean (Full Sample)	Standard Deviation (Full Sample)	Observation (Full Sample)	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	0.70	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 <i>Shura</i> Meetings	0.32	0.47	2387	0.33	0.31	0.05

Table 2. Summary Statistics for Outcome Variables

	Mean	Standard Error	Obs.
Council Member Finished High School	8.7	28.2	2,016
Council Member Finished Middle School	16.8	37.4	2,016
Distance between Residences of Council Members and Village Center (Meters)	406	867	1,018
Project Implementation Started Before:			
October 2008	0.13	0.34	1317
January 2009	0.41	0.49	1317
April 2009	0.76	0.43	1317
July 2009	0.82	0.39	1317
October 2009	0.87	0.34	1317
Project Implementation Finished Before:			
January 2009	0.41	0.49	1317
April 2009	0.82	0.39	1317
July 2009	0.87	0.34	1317
October 2009	0.90	0.31	1317
January 2010	0.91	0.29	1317

Table 3. Effect of Council Members' Educational Attainment on Project Implementation

Panel A	Project Implementation Started Before:				
	October 2008 (1)	January 2009 (2)	April 2009 (3)	July 2009 (4)	October 2009 (5)
At Least One Council Member Finished High School	0.016 [0.049]	0.033 [0.051]	0.249*** [0.065]	0.304*** [0.061]	0.298*** [0.056]
At Least One Male Baseline Survey Respondent Finished High School	0.092* [0.053]	0.171*** [0.058]	0.195** [0.078]	0.211*** [0.066]	0.189*** [0.068]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	595	595	595	595	595
R-squared	0.29	0.48	0.41	0.36	0.28

Panel B	Project Implementation Completed Before:				
	January 2009 (1)	April 2009 (2)	July 2009 (3)	October 2009 (4)	January 2010 (5)
At Least One Council Member Finished High School	0.033 [0.051]	0.304*** [0.061]	0.281*** [0.057]	0.270*** [0.049]	0.245*** [0.048]
At Least One Male Baseline Survey Respondent Finished High School	0.171*** [0.058]	0.211*** [0.066]	0.190*** [0.069]	0.252*** [0.054]	0.267*** [0.052]
Quadruple fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	595	595	595	595	595
R-squared	0.48	0.36	0.27	0.34	0.39

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

Table 4. Effect of Electoral Rules on Council Members' Competence

	Council Member Finished High School			
	(1)	(2)	(3)	(4)
At-Large Election	3.96*** [1.32]	0.07 [1.80]	1.95 [1.23]	-0.90 [1.55]
Fractionalized Project Preferences * At-Large Elections		7.97*** [3.01]		
Fractionalized Project Preferences		-3.10 [1.99]		
Ethnically Mixed Village * At-Large Elections			7.96** [3.21]	
Ethnically Mixed Village			-3.27 [2.21]	
Geographically Large Village * At-Large Elections				9.96*** [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: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5. Effect of Electoral Rules on Project Implementation

Panel A	Project Implementation Started Before:				
	October 2008	January 2009	April 2009	July 2009	October 2009
	(1)	(2)	(3)	(4)	(5)
At-Large Election	-0.001 [0.043]	0.032 [0.045]	0.086** [0.035]	0.047 [0.031]	0.058*** [0.018]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	506	506
R-squared	0.39	0.56	0.53	0.40	0.22

Panel B	Project Implementation Completed Before:				
	January 2009	April 2009	July 2009	October 2009	January 2010
	(1)	(2)	(3)	(4)	(5)
At-Large Election	0.032 [0.045]	0.047 [0.031]	0.051*** [0.018]	0.019* [0.010]	0.005 [0.005]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	506	506
R-squared	0.56	0.40	0.21	0.21	0.09

Note: 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 Council Members' Location

	Natural Log of Distance between Residences of Council Members and Village Center	
	(1)	(2)
At-Large Election	-0.32*** [0.11]	-0.28*** [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: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Location of Council Members' Residences and their Educational Attainment

	Council Member Finished High School	
	(1)	(2)
Natural Log of Distance between Council Members' Residences and Village Center	-0.65 [1.22]	-1.89 [1.24]
At-Large Election * Natural Log of Distance between Council Members' Residences and Village Center	4.45** [2.24]	4.75** [2.20]
At-Large Election	-16.16 [11.20]	-17.62 [11.00]
Natural Log of Median Distance between Residences 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: 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. Finally, in subsection A6 we demonstrate that the results from the baseline model hold if council members are elected sequentially.

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, \dots, q_N are competent, where q_1, \dots, q_N are independent random variables distributed uniformly on $[-B, B]$ (as usual, we will denote the order statistics by $q_{(1)} \leq \dots \leq q_{(N)}$). 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, \dots, 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 $(h, q_{(N)})$, 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 $BR_R(b_l) \subset \{q_{(N)}, B\}$. Moreover, this best-response function is monotone: if $b'_l < b_l$ and $B \in BR_R(b_l)$, then $B \in BR_R(b'_l)$, and if $q_{(N)} \in BR_R(b'_l)$, 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 $BR_L(b_r) \in \{-B, q_{(1)}\}$. It also satisfies monotonicity: if $b'_r > b_r$ and $-B \in BR_L(b_r)$, then $-B \in BR_L(b'_r)$, and if $q_{(1)} \in BR_L(b'_r)$, then $q_{(1)} \in BR_L(b_r)$. This monotonicity of best responses already implies existence. Obviously, if $B \in BR_R(-B)$ and $-B \in BR_R(B)$, then there is an equilibrium where $(a_l, b_l) = (0, -B)$ and $(a_r, b_r) = (0, B)$ are elected. If the first inclusion 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 $(h, q_{(N)})$ and L elects (a_l, b_l) , where $b_l \in BR_L(q_{(N)})$. Similarly, if the second inclusion fails, then there is an equilibrium where L elects $(h, q_{(1)})$ and R elects (a_r, b_r) with $b_r \in BR_R(q_{(1)})$. In any case, there is an equilibrium in pure strategies. The argument above applies, with obvious modifications, to $N = 0$ as well.

Notice, however, that it in the case of district elections, the equilibrium need not be unique (even in terms of elected types). For example, take $N = 2$, $B = 1$, $k = 1$, $h = \frac{1}{4}$, and suppose $q_1 = -\frac{1}{2}$, $q_2 = \frac{1}{2}$. Then there is an equilibrium where (h, q_1) and (h, q_2) are elected: indeed, the median voter in district R gets $\frac{1}{4} + \frac{1}{4} - \left(\frac{\frac{1}{2} + (-\frac{1}{2})}{2} - \frac{1}{2}\right)^2 = \frac{1}{4}$ by electing the competent citizen, but only $\frac{1}{4} - \left(\frac{1 - \frac{1}{2}}{2} - \frac{1}{2}\right)^2 = \frac{3}{16}$ by electing the extreme one, and thus does not want to deviate (and the calculation for district L is symmetric). At the same time, there is an equilibrium where $(0, -B)$ and $(0, B)$ are elected: in this case, the median voter in district R gets $-\left(\frac{1 + (-1)}{2} - \frac{1}{2}\right)^2 = -\frac{1}{4}$ by electing the most extreme one, but only $\frac{1}{4} - \left(\frac{\frac{1}{2} + (-1)}{2} - \frac{1}{2}\right)^2 = -\frac{5}{16}$ 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, \dots, q_N are independent and uniformly distributed on $[-1, 1]$. Fix any real number $z \in (0, 1)$. Let*

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

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 $\beta \rightarrow 1$, 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 $k(q - \frac{B}{2})^2 \leq h$ in district R and if $k(q + \frac{B}{2})^2 \leq h$ in district L , and otherwise it elects the median voter in that district. The incentives are also changed in at-large elections. Now, the median voter prefers to elect one council member with $b_i = 0$, and also the competent person, provided that $kq^2 \leq h$. Thus, to get elected, the competent person needs to be within $\sqrt{\frac{h}{k}}$ distance from 0 in at-large elections, and within such distance from either $-\frac{B}{2}$ or $\frac{B}{2}$ in district elections, and the second is clearly more likely.

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

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 $h > k(B/2)^2$, the most competent candidate is guaranteed to be elected. In at-large elections, the median voter (at 0) would be quite a bit averse to competent but biased candidates; in this case, we can only guarantee that the competent citizen will be elected if $h > kB^2$, which is a stronger condition. 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 $D_j = [-B + \frac{2B}{M}(j-1), -B + \frac{2B}{M}j]$, and each district needs to elect one council member. In at-large elections, the entire village elects M council members (to keep the model similar to the previous case, it is natural to assume that each citizen has M votes 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 $b_1 \leq \dots \leq b_M$. 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

would imply convergence to the midpoint between them.

Let us take the limit $\beta \rightarrow 1$ 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

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).

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 $b_r > b'_r$,

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

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

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

which is again increasing in b_i . Obviously, a similar increasing differences condition holds for elections in district L , holding the distribution in 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 $\frac{b_l + b_r}{2} > \frac{b'_l + b'_r}{2}$,

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

To see this, notice that

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

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

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

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 $P_a = \frac{1}{B} \int_0^B \frac{1}{B} q dq = \frac{1}{2}$. In district elections, it equals $P_d = \frac{1}{B} \left(\int_0^{\hat{q}} \frac{1}{B} B dq + \int_{\hat{q}}^B \frac{1}{B} \left(\frac{q+B}{2}\right) dq \right) = \frac{1}{4} \left(3 - \frac{\hat{q}}{B}\right) \left(1 + \frac{\hat{q}}{B}\right)$, 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 $\Pr\left(\frac{|b|}{B} < x \mid a = h\right) = \frac{x - \hat{q}/B}{1 - \hat{q}/B}$ 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,

$$\mathbb{E}\left(\frac{|b|}{B} \mid a = h\right) = \frac{1}{2} \left(1 + \frac{\hat{q}}{B}\right) < 1 = \mathbb{E}\left(\frac{|b|}{B} \mid a = 0\right),$$

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,

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

We prove that $Q(N, z) \geq P(N, z)$ (with equality only if $N = 2$) by induction by N , separately for even and odd N . We start with even N .

Suppose $N = 2$. Then

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

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

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

Now take $N \geq 4$

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

which proves the result for even N .

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

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

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

$$Q(3, z) > 1 - \left(\frac{2-z}{2} \right)^2 \frac{2-z}{2} = P(3, z).$$

Now suppose $N \geq 5$. We have

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

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

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

where $V(b_l, b_r)$ is the variance of the proposals by the two council members. As $\beta \rightarrow 1$, 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 $\beta \rightarrow 1$, 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

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

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 18

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 $[-\frac{B}{3}, \frac{B}{3}]$ (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 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

In Appendix C, we present some results on the effect of election type on villagers' satisfaction. This Appendix also contains tables for auxiliary empirical results referred to in the main text.

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

In particular, information on villagers' attitudes comes from two follow-up surveys. The midline survey was conducted in May–October 2009, at which point all complying villages had elected councils and selected projects, but only 18 percent of projects had been completed. By the time the endline survey was conducted in May–October 2011, virtually all projects were complete. The survey was designed to be administered to ten randomly selected households surveyed at baseline, with separate questionnaires for male household heads and a senior woman in the household. The data provide information on 2,367 male respondents and 2,141 female respondents in the midline survey and 2,130 male and 1,858 female respondents in the endline survey.³⁴ To measure villagers' satisfaction with local economic and governance outcomes, we use four perception-based binary indicators from male and female household surveys at midline and endline: (i) whether respondent thinks that village headman should be elected; (ii) whether the respondent attributes positive economic changes to actions of the village leadership;³⁵ (iii) whether the respondent is satisfied with the work of the village leaders; and (iv) whether the respondent perceives that the household is better off than it was last year.

³⁴Because of the deterioration in security conditions, we were not able to conduct surveys of male heads of household in 11 villages and of female heads of household in 33 villages. Enumerators administering the male household questionnaire were instructed to locate and interview the same households and, whenever possible, the same villagers who participated in the baseline survey. Enumerators were able to successfully locate such respondents in 65 percent of households in which male respondents were interviewed during the baseline survey. The predominant reason for enumerators not being able to interview baseline respondents was that the person was away from home on the day that the survey team visited the village, as it was the time of harvest. Differences between villages with different procedures of project selection in individual-level attrition are not statistically significant. We also check that the effect on attrition of such characteristics of respondents as age, income, assets, size of household, education, and ethnicity are similar in villages with different project selection rules.

³⁵This indicator is available only for the midline survey.

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. The results of the analysis of the effect of electoral rules on villagers' attitudes toward local leaders and their perception of their economic situation are presented in Table C5. They indicate that male villagers are more likely to support election of the village headman in the midline survey, but this effect decreases in magnitude and loses statistical significance in the endline survey. Female villagers in the midline survey are more likely to attribute positive economic changes to village leaders and to perceive an improvement in the economic situation of their household in villages with at-large elections, although there is no significant difference in their levels of satisfaction with the work of village leaders or in the instances of disagreement with the decisions of village leaders. There is also no evidence that electoral rules affect attitudes of male or female respondents in the endline survey.

Overall, the results are weakly consistent with the model's predictions, as there is some evidence that attitudes of the villagers are more positive in villages with at-large elections in the midline survey. We do not observe any effect of electoral rules on villager's satisfaction in the endline survey, but this is not surprising, given the timing of the endline survey. The survey was conducted four years after the selection of the councils and on average a year after all development projects were finished, i.e., at the time where all activities for which the council was responsible were over, and thus the quality of elected candidates is unlikely to affect the satisfaction of the villagers at this time.

Table C1. Effect of Electoral Rules on Council Members' Educational Attainment

	Council Member Finished Middle School			
	(1)	(2)	(3)	(4)
At-Large Election	3.43*	-2.48	0.6	-0.17
	[1.93]	[2.76]	[1.98]	[2.55]
Fractionalized Project Preferences		12.17***		
* At-Large Elections		[4.27]		
Fractionalized Project Preferences		-3.16		
		[3.05]		
Ethnically Mixed Village			11.33**	
* At-Large Elections			[5.49]	
Ethnically Mixed Village			-4.98	
			[4.26]	
Geographically Large Village				7.24*
* At-Large Elections				[4.11]
Geographically Large Village				-4.38
				[2.80]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,016	2,016	2,016	2,016
R-squared	0.19	0.19	0.19	0.19

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

Table C2. Council Members' Educational Attainment and Project Implementation (IV results)

Panel A	Project Implementation Started Before:				
	October 2008 (1)	January 2009 (2)	April 2009 (3)	July 2009 (4)	October 2009 (5)
At Least One Council Member Finished High school	-0.006 [0.265]	0.212 [0.291]	0.574** [0.262]	0.312 [0.195]	0.388** [0.151]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	506	506

Panel B	Project Implementation Completed Before:				
	January 2009 (1)	April 2009 (2)	July 2009 (3)	October 2009 (4)	January 2010 (5)
At Least One Council Member Finished High school	0.212 [0.291]	0.312 [0.195]	0.339** [0.145]	0.125* [0.067]	0.033 [0.033]
Project Type Fixed Effects	Yes	Yes	Yes	Yes	Yes
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	506	506	506	506	506

Note: Results of the instrumental variables estimation in which the variable “At Least One Council Member Has Finished High school” is instrumented with the dummy variable for at-large elections. F-statistics for the exclusion of instrument from the first-stage regression is 10.95. 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' Educational Attainment
(Excluding Districts with More than One Member Elected)

	Council Member 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: Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C4. Electoral Rules and Incumbency Advantage

	Percent of Male Council who are Pre-Existing Elite		Percent of Pre-Existing Elite Elected to Council	
	Mean in District Elections	Difference	Mean in District Elections	Difference
Definition of Elite				
Member Of Baseline Focus Group (including Non-Attendees)	31.9	2.43 [2.23]	25.8	1.55 [1.89]
Observations	1055	2058	1,301	2,594
Decisions-Maker According To Male Focus Group	13.2	-0.54 [1.27]	43.8	-3.73 [3.81]
Observations	1055	2058	320	628
Decisions-Maker According To Male Head-of-Household Survey	20.7	3.24* [1.86]	27.9	1.58 [2.25]
Observations	1055	2058	790	1,658
Decisions-Maker According To Female Individual Survey	14.9	-0.66 [1.53]	30.2	-2.05 [2.95]
Observations	1055	2058	526	1,027
Either Of The Four Above	38.9	3.20 [2.39]	19.3	1.25 [1.23]
Observations	1055	2058	2,761	5,530

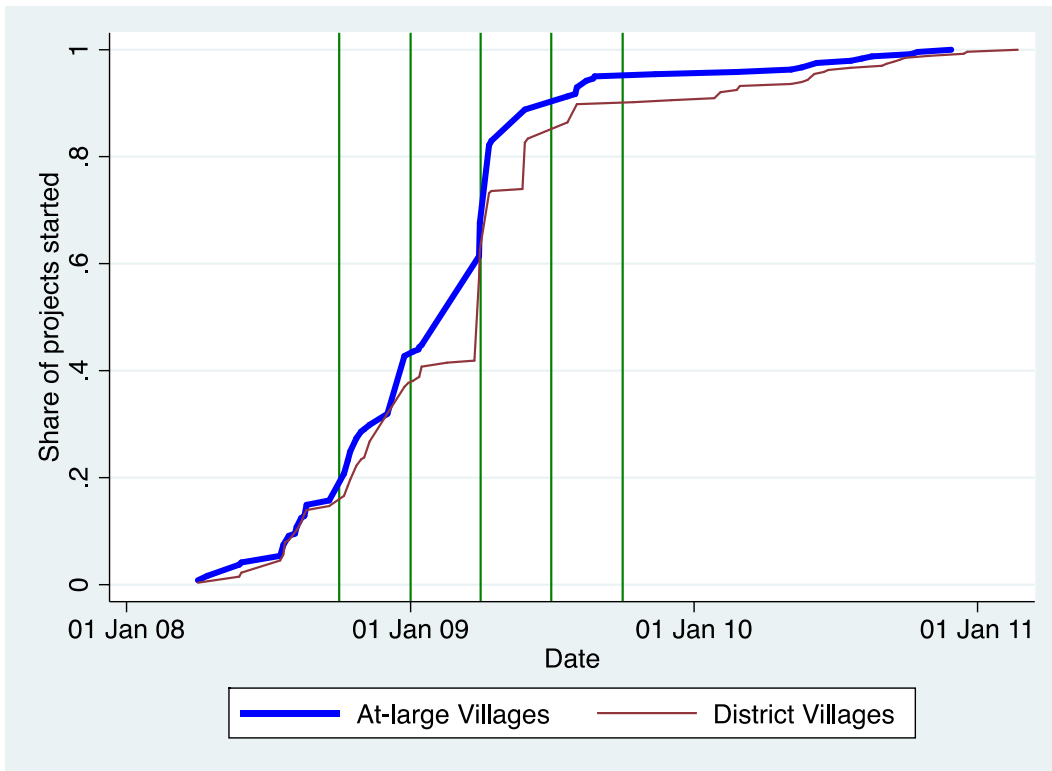
Note: The difference between district and at-large elections estimated using the same model as in regression (1). Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C5: Effect of Electoral Rules on Voter Attitudes

	Respondent Prefers to Elect Village Head		Respondent is Satisfied with Village Leaders	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
At-Large Election * Midline Survey	0.05** [0.02]	-0.01 [0.03]	-1.87 [1.834]	1.27** [0.557]
At-Large Election * Endline Survey	0.02 [0.02]	-0.02 [0.03]	-0.83 [2.235]	
Quadruple * Survey Fixed Effects	Yes	Yes	Yes	Yes
Observations	4,440	3,578	4,270	2,135
R-squared	0.23	0.12	0.15	0.04
	Respondent Attributes Positive Economic Change to Village Leaders		Household's Economic Situation Has Improved in Past 12 Months	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
At-Large Election * Midline Survey	-0.28 [0.832]	1.27** [0.557]	-0.78 [2.473]	4.45* [2.517]
At-Large Election * Endline Survey			-1.18 [1.705]	-1.59 [1.949]
Quadruple * Survey Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,355	2,135	4,493	3,997
R-squared	0.08	0.04	0.19	0.15
	Respondent Complains About Project Implementation		Respondent Complains About Corruption or Nepotism in Project	
	Male		Male	
At-Large Election * Midline Survey				
At-Large Election * Endline Survey	-2.26 [3.115]		-0.79 [1.758]	
Quadruple * Survey Fixed Effects	Yes		Yes	
Observations	1,716		1,722	
R-squared	0.21		0.16	

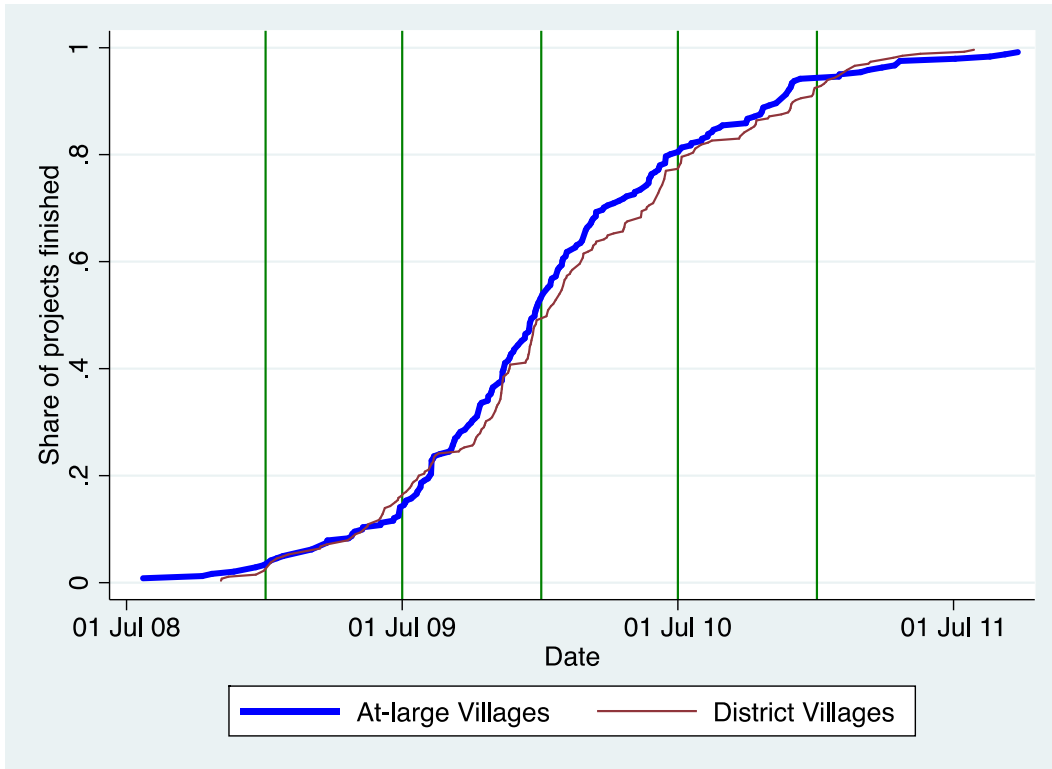
Note: Outcomes measured in percent. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Figure C1. Effect of Electoral Rules on Distribution of Project Start



Note: Green/vertical lines show dates of points used to measure project starts in the empirical analysis.

Figure C2. Effect of Electoral Rules on Distribution of Project Completion



Note: Green/vertical lines show dates of points used to measure project completion in the empirical analysis.