Online Appendix

Appendix A: Robustness

The model in Section 3 is simple and makes clear predictions. In Appendix A, we show that these predictions are not due to excess simplification of the environment and that our results are robust. Subsection A1 considers an extension of the model in which we allow for multiple competent candidates. In subsection A2, we study an explicit legislative bargaining game. In subsection A3 we contrast joint decision-making with individual decision-making by the elected council members. In subsection A4 we consider the case of multiple electoral districts, and thus more than two council members. In subsection A5 we discuss where competence of council members affects their bargaining power. In subsection A6 we demonstrate that the results from the baseline model hold if council members are elected sequentially. In subsection A7 we show that the predictions of the model hold if citizens vote for male and female council members separately. Finally, in subsection A8 we show that with multiple competent candidates, the results hold regardless of whether abilities of council members are complements or substitutes.

A1 Several competent individuals

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

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

We start by showing that for any N and any realization of $q_1, \ldots, q_N$, there exists an equilibrium in pure strategies, both in district and in at-large elections. The median voter theorem applies again, and for at-large elections, a pair of citizens that maximizes the utility of the median voter,

$w_{m_X}(a_l, b_l, a_r, b_r)$, will be elected in an equilibrium. Notice that the median voter only needs to consider $N(N-1)/2$ pairs of competent citizens plus a combination of one competent citizen with type, say, (h, q_1) and his political antipode $(0, -q_1)$; since he only needs to choose among a finite number of pairs, the maximum is attained at some pair.

The argument is only slightly more involved in the case of district elections. Suppose that in some pure strategy equilibrium the left district L elects a citizen (a_l, b_l). The best response by the right district's median voter is either to elect the most extreme of the competent individuals $\left(h, q_{(N)}\right)$, provided that there is a competent individual in the district ($q_{(N)} \geq 0$), or to elect the most extreme individual $(0, B)$; this only depends on b_l. Thus, the political preferences of the best-response individual is $BR_R(b_l) \subset \left\{q_{(N)}, B\right\}$. Moreover, this best-response function is monotone: if $b_l' < b_l$ and $B \in BR_R(b_l)$, then $B \in BR_R(b_l')$, and if $q_{(N)} \in BR_R(b_l')$, then $q_{(N)} \in BR_R(b_l)$. Similarly, if the right district elects a citizen (a_r, b_r), the political preferences of the best-response individual in the left district L is $BR_L(b_r) \in \left\{-B, q_{(1)}\right\}$. It also satisfies monotonicity: if $b_r' > b_r$ and $-B \in BR_L(b_r)$, then $-B \in BR_L(b_r')$, and if $q_{(1)} \in BR_L(b_r')$, then $q_{(1)} \in BR_L(b_r)$. This monotonicity of best responses already implies existence. Obviously, if $B \in BR_R(-B)$ and $-B \in BR_R(B)$, then there is an equilibrium where $(a_l, b_l) = (0, -B)$ and $(a_r, b_r) = (0, B)$ are elected. If the first inclusion does not hold, then $BR_R(b_l) = q_{(N)}$ for any b_l, and thus there is an equilibrium where R elects individual with type $\left(h, q_{(N)}\right)$ and L elects (a_l, b_l), where $b_l \in BR_L\left(q_{(N)}\right)$. Similarly, if the second inclusion fails, then there is an equilibrium where L elects $\left(h, q_{(1)}\right)$ and R elects (a_r, b_r) with $b_r \in BR_R\left(q_{(1)}\right)$. In any case, there is an equilibrium in pure strategies. The argument above applies, with obvious modifications, to $N = 0$ as well.

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

We thus have the following result.

Proposition A1 *Suppose that there are N competent individuals, where N is a non-negative integer. Then for any realization of their political preferences there exists an equilibrium.*

To proceed further, we need the following technical lemma, which is proven, along with other results, in Appendix B.

Lemma A1 *Suppose that $N \geq 2$ random variables $q_1, \ldots, q_N$ are independent and uniformly distributed on $[-1, 1]$. Fix any real number $z \in (0, 1)$. Let*

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

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

In what follows, assume that $\frac{h}{k} < \frac{1}{4}B^2$; this assumption means that the political dimension is sufficiently important. It says that any citizen prefers his ideal point implemented by an incompetent council member to a point at distance $B/2$ implemented by a competent one. The assumption guarantees that in within-district elections, a competent citizen with b_i close to 0 will not be elected, so there is a real competence-vs.-bias trade-off in district elections.

Consider at-large elections. The median voter can always guarantee himself utility h by electing any competent citizen, e.g., (h, q_1), and a corresponding incompetent citizen $(0, -q_1)$. However, he could do better if there were two competent citizens q_i and q_j with $k\left(\frac{q_i+q_j}{2}\right)^2 < h$; in this case, he would get $2h - k\left(\frac{q_i+q_j}{2}\right)^2$. Therefore, in at-large elections, if there are two competent citizens with $|q_i + q_j| \leq 2\sqrt{\frac{h}{k}}$, then the council will consist of two competent citizens, and otherwise will contain one competent and one incompetent one. By Lemma A1 (which we may apply with an appropriate normalization), the probability that both members are competent equals $Q(N, z)$, where we denoted $z = \frac{2}{B}\sqrt{\frac{h}{k}} < 1$.

Now consider district elections. Two competent citizens will be elected only if both districts elect competent citizens. Suppose district L elected a citizen with political position b_l; then the the median voter in district R will elect a competent citizen only if there is one with political position q_j such that $h - k\left(\frac{b_l+q_j}{2} - \frac{B}{2}\right)^2 \geq -k\left(\frac{b_l+B}{2} - \frac{B}{2}\right)^2$, i.e., if $q_j \geq B - b_l - \sqrt{4\frac{h}{k} + b_l^2}$. This is equivalent to $1 - \frac{q_j}{B} \leq \frac{b_l}{B} + \sqrt{4\frac{h}{kB^2} + \left(\frac{b_l}{B}\right)^2}$; therefore, by Lemma A1, for any given b_l, the probability that district R elects a competent citizen is $P(N, z(b_l))$, where $z(b_l) = \frac{b_l}{B} + \sqrt{4\frac{h}{kB^2} + \left(\frac{b_l}{B}\right)^2}$. Notice that $z(b_l) \leq z$ for all $b_l \in [-B, 0]$, and the inequality is strict for $b_l \neq 0$.

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

Proposition A2 *For any number of competent citizens $N \geq 1$, the expected quality of council members under at-large elections is higher than under district elections. Moreover, the number of competent council member under at-large elections first-order stochastically dominates that under district elections.*

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

A2 Legislative bargaining game

In this subsection, we modify the game from Section 3 by assuming that the two elected council members do not automatically choose the policy midway between their ideal points, but rather participate in a legislative bargaining game, as in Banks and Duggan (2000). Namely, the two council members, l with type (a_l, b_l) and r with type (a_r, b_r) (where $b_l < b_r$) play the following game.

There are an infinite number of periods, starting with period 0. In each period, each of the council members becomes agenda-setter with probability $\frac{1}{2}$. The agenda-setter proposes policy p, and the other member either accepts or rejects it. If p is accepted in period t, then each citizen i (including the two council members) get $u(p, b_i) = -k(p - b_i)^2$ in each subsequent period. In every period before a policy is accepted, all citizens suffer a penalty $-P$, where $P > 4kB^2$ (since the payoff from policy $u(p, b_i)$ is non-positive, we need to assume that the payoff without any policy is even worse, even if the distance to that policy is $2B$). All citizens maximize their discounted expected payoff, and $\beta \in (0,1)$ is a common discount factor. In this model, we assume $\frac{h}{k} < \frac{1}{2}B^2$; this is required to obtain strict results in Proposition A3 below, but it does not affect existence of equilibrium.

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

Lemma A2 *Suppose that two council members, l and r, have ideal points (a_l, b_l) and (a_r, b_r) with $b_l < b_r$. Then in equilibrium:*

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

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

Lemma A2 says the following. If the discount factor β is sufficiently high, then the acceptance set A is sufficiently narrow; it lies strictly between the ideal positions of the two council members, and each agenda-setter proposes the policy at the extreme of the acceptance set. If the discount factor is sufficiently low, then the acceptance set is wide, as the politicians are too impatient and are willing to accept policies that are far from their ideal point. This allows each politician to insist on their ideal policy in equilibrium. It is easy to see that if punishment P is very high, then the acceptance set is likely to be large, and politicians will propose their ideal policy.

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

Proposition A3 *For any β, there exists an equilibrium. Moreover, there exist $0 < \beta_1 < \beta_2 < 1$ such that:*

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

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

Proposition A3 gives two important takeaways. First, if offers are made frequently and β is close to 1, the outcomes of elections are similar to the outcomes of the game studies in Section 3, and this implies robustness of those results. Second, if offers are made rarely, the results are overturned, and district elections lead to more competent council members. This goes in contrast to the previous results; to see the intuition, it is helpful to observe that if β is low enough, then each council member will propose his ideal point. This creates very different incentives to voters in district elections: instead of electing a very biased council member in hope that his influence would moderate the council member from the other district, the median voter in a district would prefer to elect someone with ideal point close to him. Indeed, this median voter has no hope of influencing the offer made by the council member elected by the other district, and instead he wants to get higher utility from offers made by his own delegate. As a result, a district which lacks a competent individual elects his median voter to the council, whereas a district with the competent person elects him if he is close enough to the median voter, i.e., if $k\left(q - \frac{B}{2}\right)^2 \leq h$ in district R and if $k\left(q + \frac{B}{2}\right)^2 \leq h$ in district L, and otherwise it elects the median voter in that district. The incentives are also changed in at-large elections. Now, the median voter prefers to elect one council member with $b_i = 0$, and also the competent person, provided that $kq^2 \leq h$. Thus, to get elected, the competent person needs to be within $\sqrt{\frac{h}{k}}$ distance from 0 in at-large elections, and within such distance from either $-\frac{B}{2}$ or $\frac{B}{2}$ in district elections, and the second is clearly more likely.

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

A3 Joint and individual decisions

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

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

It turns out that our results remain intact for α sufficiently high, but as α becomes smaller, district elections will dominate at-large ones. To see why, consider the extreme, $\alpha = 0$, and notice that in this case the median voter in district elections does not have a strategic reason for voting for biased candidates. His ideal candidate has the same ideal point as he does ($-B/2$ or $B/2$), and moreover, the problems of the two districts are independent. Now, the reason why district elections would lead to more competent candidates is clear: the median voter in the district is not too averse to any of the candidates in this district; for example, if $h > k\left(B/2\right)^2$, the most competent candidate is guaranteed to be elected. In at-large elections, the median voter (at 0) would be quite a bit averse to competent but biased candidates; in this case, we can only guarantee that the competent citizen will be elected if $h > kB^2$, which is a stronger condition. Notice that this result is very similar to the prediction of Subsection A5 (Proposition A3): there, if β is low enough, the ideal points of council members would be picked with equal probability, which matches the case $\alpha = 0$.

This result once again confirms that our results are driven by the joint nature of decision-making in councils. At-large elections are preferred if council members make a joint decision. If they have multiple policy questions which they split between themselves, then district elections should have an edge. Studying such trade-offs in more detail seems to be a fruitful area for future research.

A4 Multiple districts

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

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

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

A5 Education and bargaining power

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

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

Overall, if education is positively correlated with bargaining power, the effect that at-large elections lead to more competent council members will diminish. Intuitively, voters in district elections would prefer to elect competent council members, because this would help them distort the policy rather than hurt, as in the baseline model (see also Mattozzi and Snowberg, 2015, 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).

A7 Two types of citizens

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

The council now consists of four members, (a_i,b_i) for $i\in\{1,2,3,4\}$, where without loss of generality we assume that $\{b_i\}$ is nondecreasing. For simplicity, let us focus on a bargaining game with a simple majority voting rule, where offers are made very frequently (similarly to the game considered earlier in Subsection A4). In the limit, the policy chosen by the council is halfway between the two median ideal points of council members, i.e., $p = \frac{1}{2}(b_2 + b_3)$. Thus, the utility of a citizen with ideal point b is now

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

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

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

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

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

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

A8 Nonadditive impact of competent citizens

In Subsection A1 above, we studied the impact of multiple competent citizens in the village. There, we assumed that the villagers' utility from electing a certain council in additive in council members' abilities. This need not be true in practice. For example, the performance of a council may critically depends on whether there is at least one competent member, in which case the additional benefit of having another one is minimal. It is also possible that having two competent members is much more important than having one, because they can work together as a team. In this Subsection, we show that our results are robust, regardless of whether the effect of council members' competences on the villagers' welfare is additive, subadditive, or superadditive.

More precisely, let us go back to the setup of Subsection A1, with two districts, L and R. The society elects two council members; suppose that all individuals value one competent official as h (as in the main model), and two competent officials as h'. Assume $h' > h$, but not necessarily equal to $2h$. For simplicity, consider the case where the number of competent citizen is exactly 2; generalizing the reasoning below to an arbitrary N may be done similarly to Subsection A1.

Let us first show that a pure-strategy equilibrium exists for any realization of q_1 and q_2, the locations of the competent individuals (without loss of generality, assume $q_1 < q_2$). In at-large elections, the median voter theorem applies, and a pair of citizens that maximizes the utility of median voter is elected. As in Subsection A1, the median voter is effectively maximizing over a finite number of pairs (two competent citizens, or one competent citizen and an incompetent antipode). Thus, the maximum is attained at some pair, and an equilibrium exists.

Now consider district elections; here, again, in each district, the outcome is decided by the median voter in that district. First, consider the case where both competent citizens reside in the same district (e.g., district L). Then district R elects the most extreme individual regardless of the choice made in district L. The median voter in district L will in this case choose between two options: the most extreme citizen $(0, -B)$ and the more extreme of the competent citizens, (h, q_1). Clearly, there is an equilibrium.

Finally, consider the case where q_1 resides in district L and q_2 resides in district R. Without loss of generality, suppose that q_1 is at least as extreme as q_2: $q_1 - (-B) \leq B - q_2$. If it is an equilibrium for both districts to elect their competent residents, then we are done. Suppose that it is not an equilibrium, which means that at least one of the districts prefers to elect its extreme citizen, conditional on the other district electing the competent one. Obviously, the "temptation to deviate" is stronger in district R, because its competent citizen is relatively less extreme, thus, if either district has a profitable deviation, then district R does. Consider a candidate equilibrium where L elects the competent citizen (h, q_1), while R elects $(0, B)$. District R does not have a profitable deviation here (otherwise, electing two competent citizens would be an equilibrium). If the best response of district L is to elect (h, q_1), as opposed to $(0, -B)$, then we have found an equilibrium. Otherwise, if its best response is to elect $(0, -B)$, then consider the candidate equilibrium where both districts elect the most extreme citizens, $(0, -B)$ and $(0, B)$. Here, district L cannot have a profitable deviation, since otherwise we would have found an equilibrium in the previous case. Suppose district R has a profitable deviation to (h, q_2). But q_2 is relatively farther away from B than q_1 is from $-B$; thus, if R has a profitable deviation, then L also does, to (h, q_1). However, we already proved that this is not the case, which shows that if neither (h, q_1) and (h, q_2) nor (h, q_1) and $(0, B)$ constitute an equilibrium, then $(0, -B)$ and $(0, B)$ does. This proves existence of a pure-strategy equilibrium; notice that this argument goes through for any h'.

Let us now show that the number of competent council members under at-large elections first-order stochastically dominates that under district elections. For simplicity, assume (similarly to Subsection A1) that $\frac{\max(h,h'-h)}{k} < \frac{1}{4}B^2$; here, this guarantees that in within-district election, a competent citizen with b_i close to 0 will not be elected, regardless of whether the other district elected a competent citizen (so that the incremental utlility is $h' - h$) or incompetent one (so the incremental utility is h).

More precisely, we need to show that under at-large elections, electing zero competent members is less likely than under district elections, and electing two competent members is more likely than under district elections. The first part is obviously true, because under at-large elections, at least one competent citizen is always elected, whereas under district elections this is not always the case (e.g., this will not happen if both competent members have preferences close to 0, because of the assumption we made). It remains to prove the second part. Under at-large elections, the median voter elects two competent citizens if and only if $h' - k\left(\frac{q_1+q_2}{2}\right)^2 > h$, i.e., if $|q_1 + q_2| \leq 2\sqrt{\frac{h'-h}{k}}$. In the notation of Lemma A1, the probability that this is the case equals $Q\left(2, z'\right)$, where $z' = \frac{2}{B}\sqrt{\frac{h'-h}{k}}$. Under district elections, two competent individuals are elected only if both districts elect such a citizen, which is only possible if the two competent citizens reside in different districts. Using the same logic as in Subsection A1, we can show that the conditional probability that district R elects competent citizen if district L elected a competent citizen equals $P\left(2, z'\left(q_1\right)\right)$, where $z'\left(q_1\right) = \frac{q_1}{B} + \sqrt{4\frac{h'-h}{kB^2} + \left(\frac{q_1}{B}\right)^2}$. We have $z'\left(q_1\right) \leq z'$ for all $q_1 \in [-B, 0]$, with strict inequality for $q_1 \neq 0$, which implies that the conditional probability that R elects a competent citizen if L did so equals $P\left(2, z'\left(q_1\right)\right) < P\left(2, z'\right) \leq Q\left(2, z'\right)$; here, the first inequality is strict because a competent citizen with q_1 close to 0 would never be elected in L. This means that the probability that both districts elect competent citizens is also strictly less than $Q\left(2, z'\right)$, the corresponding probability under district elections, even if they reside in different districts.

Thus, the number of competent council member under at-large elections first-order stochastically dominates that under district elections, and in particular, the expected quality of council members under at-large elections is higher than under district elections. While the reasoning above deals with two competent citizens, the result is true for any number of competent citizens N. This result suggests that additivity of council members' competences is not a critical assumption, but rather a simplifying one.

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 is district R fixed.

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

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

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

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. Consider at-large elections and take 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 $\frac{b_l+b_r}{2} > \frac{b_l'+b_r'}{2}$. We have

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

which is increasing in b_i. This establishes the following increasing differences property:,

$$w_i(a_l, b_l, a_r, b_r) - w_i(a_l', b_l', a_r', b_r') > w_j(a_l, b_l, a_r, b_r) - w_j(a_l', b_l', a_r', b_r').$$

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

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

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

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

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

Proof of Proposition 3. Part 1. In at-large elections, one council member is competent and the other is not, thus expected competence is $C_a=\frac{h}{2}$. In district elections, the expected competence is $C_d=R\frac{h}{2}$ (where R is given by (B1)). Thus, $C_a\geq C_d$, because $P\leq 1$, and the inequality is strict whenever $R<1$, which may be simplified to $\frac{h}{k}<\frac{3}{4}B^2$. The difference is $C_a-C_d=(1-P)\frac{h}{2}=\left(2-\sqrt{4\frac{h}{kB^2}+1}\right)\frac{h}{2}$, which is increasing in B and k.

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

Part 3. In at-large elections, for a council member (a,b), $\Pr\left(\frac{|b|}{B}<x\mid a=h\right)=\Pr\left(\frac{|b|}{B}<x\mid a=0\right)=x$ (for $x\in[0,1]$). Therefore, in elected council members, competence and bias are independent and thus uncorrelated. In district elections, if $a=h$, the conditional distri-

bution is uniform on $\left[\frac{\hat{q}}{B},1\right]$, so $\Pr\left(\frac{|b|}{B}<x\mid a=h\right)=\frac{x-\hat{q}/B}{\hat{q}/B}$ for $x\in\left[\frac{\hat{q}}{B},1\right]$. At the same time, if $a=0$, the conditional distribution is an atom at 1: $\Pr\left(\frac{|b|}{B}=1\mid a=0\right)=1$. Hence,

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

because $\hat{q}<B$. Consequently, in district elections, a and b are negatively correlated. This completes the proof. ■

Proof of Proposition 4. In district elections, if there is a competent candidate, a council member with $a=0$ may only have $|b|=B$; this follows from Proposition 1. Reasoning similar to that used in the proof of Proposition 1 suggests that in the absence of competent candidates, only council members with $|b|=B$ will be elected. Thus, in district elections, neither presence of a competent candidate nor the fact that one is elected affects the political bias of incompetent candidates.

In the case of at-large elections, suppose that the council members play a bargaining game as in Subsection A2. Proposition A3 shows that as the discount factor β tends to 1, proposals made by the two council members tend to the midpoint between their political positions. It follows that in the limit, the ideal pair of council members from the median voters' perspective converges, to an equilibrium in the game from Section 3, where the council members chose the midpoint automatically. If a competent candidate is present and his political position is $b=q$, then for β sufficiently close to 1 he will be elected, and the political position of the other candidate will tend to $-q$ as $\beta\to 1$. If a competent candidate is absent, then from Lemma A2 it follows that for any $\beta<1$, the median voter will prefer to elect two council members with the same political position $b=0$. In the limit as $\beta\to 1$, we have that if a competent candidate is absent, then the incompetent council members have zero bias, and if one is present, then this bias is almost always non-zero. Since we showed that in at-large elections, if β is close to 1, then a competent candidate is elected if and only if he is present, the result follows. ■

Proof of Proposition 5. 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

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

In case of district elections, it equals

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

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

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

Thus,

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

This completes the proof. ■

B2 Proofs of results from Subsection A1

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

Proof of Lemma A1. The fact that $P(N,z)$ and $Q(N,z)$ are strictly increasing in both variables is trivial. Denote the c.d.f. of each of q_j by $F(x)$; then $F(x) = \frac{x+1}{2}$ for $x \in [-1,1]$. Let us first show that

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

Indeed,

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

We prove that $Q(N,z) \geq P(N,z)$ (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 $N \geq 4$

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

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

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

Now suppose $N \ge 5$. We have

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

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 $\Pr(C_a \le 1) < \Pr(C_d \le 1)$. 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 $1 \le i \le N$, then two extreme agents will be elected). It therefore suffices to prove that $\Pr(C_a = 2) > \Pr(C_d = 2)$.

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

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

B3 Proofs of results from Subsection A2

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

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

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

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

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

Now consider district elections, and suppose that the competent citizen resides in district L.

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

Now observe that for β sufficiently close to 0, council members of any type propose their ideal points in equilibrium; this follows from Lemma A2, given that $P > 4kB^2$, which holds by assumption. Consequently, the utility of a citizen with ideal point b from a council with types $(a_l, b_l), (a_r, b_r)$ is

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

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 26

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

Proof that if in at-large elections each voter may cast only one vote, there may be multiple equilibria. Let us prove that for some parameter values, there are several equilibria.

Suppose that the competent voter has bliss point q, and suppose that h is high enough (namely, $h \geq \frac{16}{9}kB^2$) Let us show that any pair of council members (h, q) and $(0, b)$ may be elected in equilibrium, provided that $|q+b| < \frac{2B}{3}$.

Consider an equilibrium where share ε of voters (where $\varepsilon > 0$ is small) vote for the competent citizen (h, q), and the rest vote for $(0, b)$; these two are then elected. The equilibrium policy in this case is $\frac{q+b}{2}$. The condition on h ensures that nobody wants to jeopardize election of a competent citizen. Indeed, a citizen with ideal point b_i gets $h - k\left(\frac{q+b}{2} - b_i\right)^2$; if a deviation prevents the competent citizen from being elected, he will get at most 0. Since $\left|\frac{q+b}{2}\right| < \frac{B}{3}$ and $|b_i| \leq B$, $h - k\left(\frac{q+b}{2} - b_i\right)^2 > h - k\left(\frac{4}{3}B\right)^2 > 0$, and thus such deviation is not profitable.

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

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

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

Appendix C: Additional Empirical Results

Table C1. Spatial Correlation of Preferences.

	Villagers Prefer the Same Type of Project		
Natural Log of Distance between Residences of Two Villagers in the Same Village	-0.013***		
	[0.004]		
Natural Log of Median Distance between Residences of Villagers	-0.016		
	[0.012]		
Distance between Residences of Two Villagers is Below Median		0.065***	
		[0.009]	
Distance between Residences of Two Villagers is in the First Tercile			0.073***
			[0.011]
Distance between Residences of Two Villagers is in the Second Tercile			0.021**
			[0.009]
Observations	20,930	20,930	20,930
R-squared	0.002	0.005	0.004

Note: The unit of observation is a pair of villagers within the same village. The dependent variable is a dummy variable that equals 1 if both villagers indicated the same type of project as the most preferred one in the baseline survey and zero otherwise. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C2. Effect of Electoral Rules on Council Member Competence (Including Female Council Members).

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

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

Table C3. Effect of Electoral Rules on Male Council Member Competence with Additional Controls.

	Percent of Male Council Members who Finished High School			Percent of Male Council Members who Finished Middle School			Percent of Male Council Members who Finished Primary School		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
At-Large Elections	3.96	7.48	6.03	-4.84	0.78	2.22	-14.91	-14.50	-8.88
	[7.85]	[7.36]	[6.98]	[11.21]	[10.95]	[10.55]	[15.96]	[14.97]	[14.93]
Fractionalized Project Preferences	8.52***			12.31***			10.09		
* At-Large Elections	[2.86]			[4.34]			[6.24]		
Fractionalized Project Preferences	-3.51*			-3.81			-5.81		
	[1.95]			[2.89]			[4.21]		
Ethnically Mixed Village		7.23**			9.67*			18.82**	
* At-Large Elections		[3.64]			[5.58]			[7.42]	
Ethnically Mixed Village		-4.41*			-5.10			-10.19*	
		[2.39]			[4.11]			[5.57]	
Geographically Large Village			9.62***			5.42			0.93
* At-Large Elections			[3.18]			[4.33]			[5.76]
Geographically Large Village			-2.76			-3.66			0.19
			[1.96]			[2.92]			[4.22]
Quadruple Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
R-squared	0.19	0.19	0.19	0.19	0.19	0.19	0.26	0.26	0.26

Note: The unit of observation is male council member. The dependent variable is a dummy variable that equals 100 if a council member finished high, middle or primary school respecitvely and zero otherwise. Additional controls include population of a village, average size of the household, average age of male villagers, household expenditure on food in last 30 days, share of households for which the primary source of household income is agriculture, shaer of male villagers who finished high school, as well as interactions of all these variables with an indicator for at-large elections. Standard errors clustered at the village level in parentheses. *significant at 10%; ** significant at 5%; *** significant at 1%.

Table C4. Effect of Electoral Rules on Male Council Member Competence.

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

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

Table C5. Effect of Electoral Rules on Council Member' Competence (as Measured by Occupation).

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

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

Table C6. Effect of Electoral Rules on Council Member Competence
(Excluding Districts with More than One Member Elected).

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

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

Table C7. Electoral Rules and Incumbent Advantage.

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

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

Table C8. Lee (2009) Bounds for the Attrition Bias.

	Percent of Male Council Members who Finished High School	Natural Log of Distance between Residences of Council Members and Village Center	Days Between Elections and Project Start	Days Between Elections and Project Completion	Summary Index of Economic and Political Outcomes, Midline	Summary Index of Economic and Political Outcomes, Endline
Effect of At-Large Elections						
Lower Bound	2.999**	-0.222***	-35.148**	-29.551**	0.012	0.023*
	[1.235]	[0.074]	[13.947]	[13.455]	[0.011]	[0.013]
Upper Bound	5.198***	-0.189***	-34.653***	-29.272**	0.053***	0.048***
	[1.402]	[0.069]	[13.268]	[13.250]	[0.011]	[0.014]
Observations	2,116	2,349	512	512	2,500	2,500

Note: The table shows lower and upper bounds for the effect of at-large elections that accounts for nonrandom attrition, as proposed by Lee (2009). Bootstraped standard errors based on 1000 repetitions in brackets.

Table C9. Location of the Projects.

	Natural Log of Distance from the Location of the Project to the Center of the Village		
At-Large Elections	-0.153	-0.157	-0.119
	[0.221]	[0.224]	[0.309]
At-large Elections*At Least One Council Member Finished High School			-0.188
			[0.443]
At Least One Council Member Finished High School			0.406
			[0.491]
Natural Log of Median Distance between Residences of Villagers		0.346**	0.357**
		[0.163]	[0.164]
Quadruple Fixed Effects	Yes	Yes	Yes
Observations	335	335	333
R-squared	0.533	0.546	0.548
Number of villages	144	144	143

Note: The unit of observation is project.Standard errors clustered at the village level in parentheses.
*significant at 10%; ** significant at 5%; *** significant at 1%.

`

Figure C1. Size of the Council by Type of Elections.

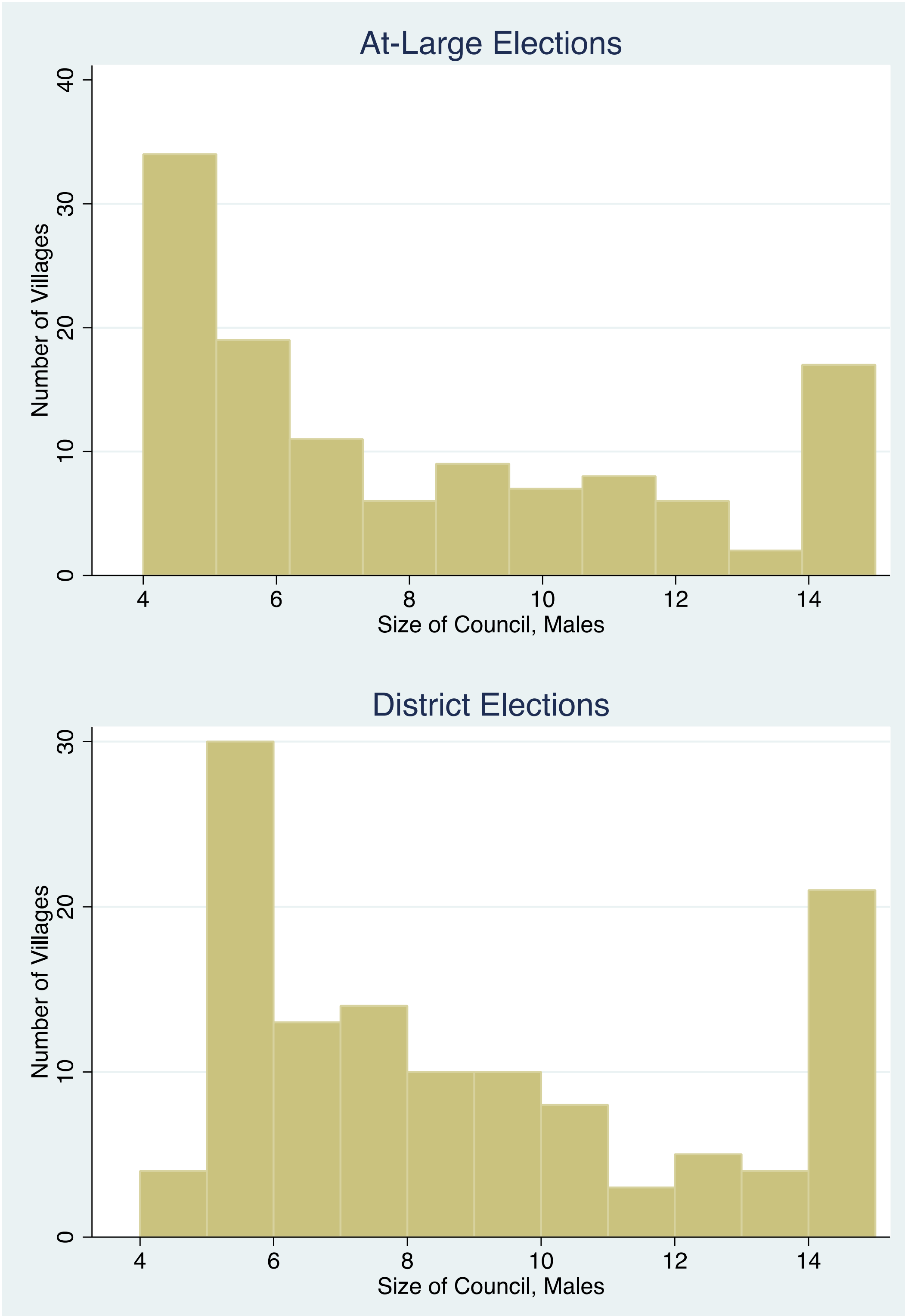

Figure C2. Cumulative Distribution Function for The Days Between Elections And Project's Start by Type of Elections.

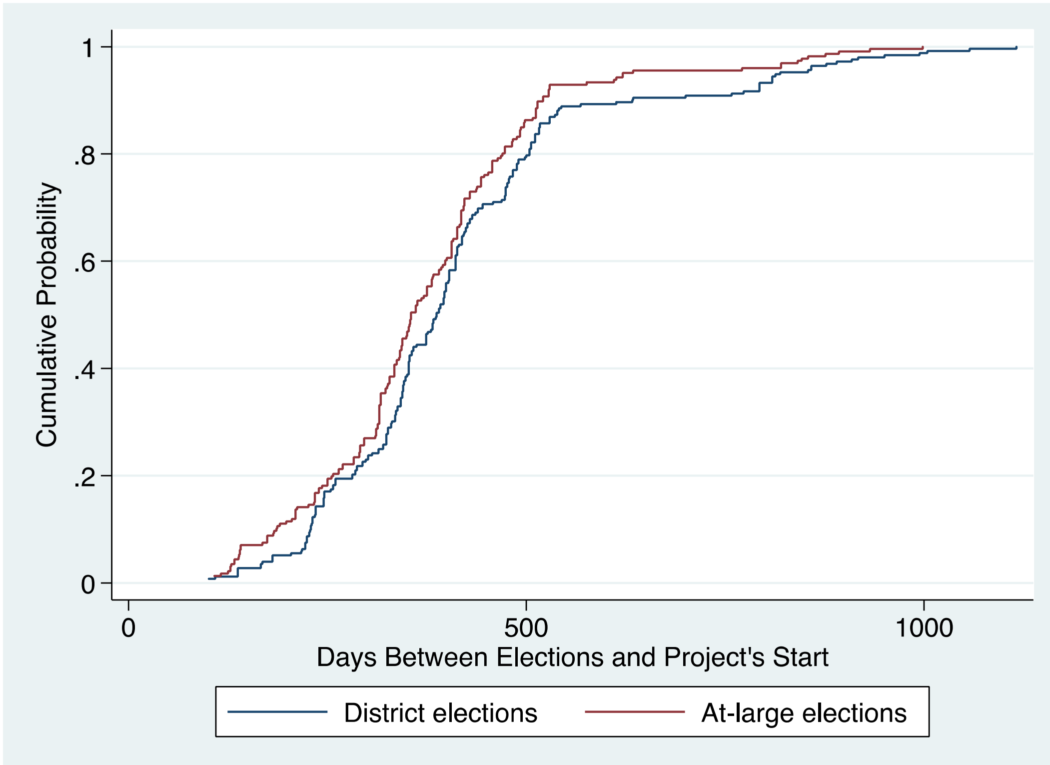

Figure C3. Cumulative Distribution Function for The Days Between Elections And Project's Completion by Type of Elections.

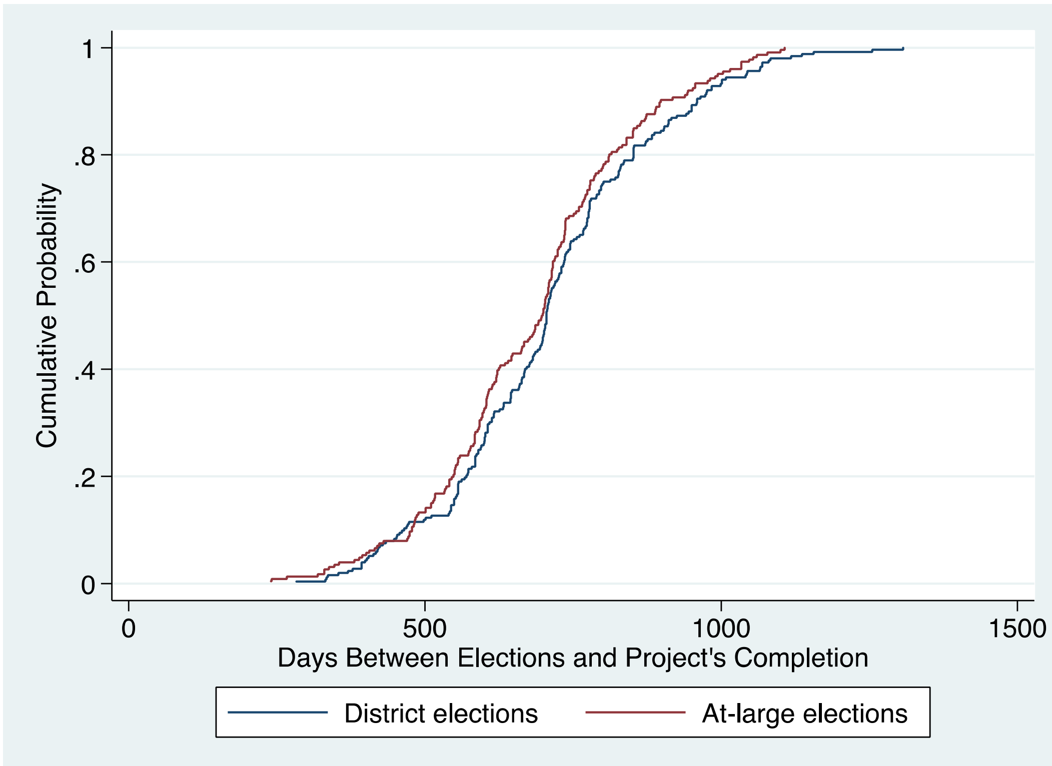

Figure C4. Average Education of Villagers by Quartiles of Distance to the Center of a Village.