University of Cyprus
Department of Economics

Working Paper 06-2020

Representative Democracy with or without Elections: An Economic Analysis

Miltos Makris, Theodore Palivos and Marios Zachariadis

Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
Tel.: +357-22893700, Fax: +357-22895028, Web site: http://www.ucy.ac.cy/econ/en

Representative Democracy with or without Elections: An Economic Analysis*

Miltos Makris (U of Kent) Theodore Palivos (AUEB) Marios Zachariadis (UCY)

May 31st 2020

Abstract

We compare the quality of policies set by a representative in the face of complexity/uncertainty under an electoral system where the representative is chosen out of the set of politicians versus a system where the policy-setter is randomly selected from the "demos", that is, the subset of citizens willing to serve as representatives if selected by lot. We do so by recognizing that the differences between the two systems affect the incentives of citizens to participate in the selection process in place. We find that for high enough returns from being the representative, drawing the decision-maker from the demos dominates elections because higher returns attract more able-for-the-job citizens while the probability of winning elections is decreasing in the number of politicians/candidates. Importantly, we also find that an increase in the complexity of policy issues makes it more likely that drawing the representative from the demos dominates elections since it reduces the probability of being selected under elections and thereby the incentives of more able-for-the-job citizens to become politicians. Calibrating our model we show that selection by lot is more likely to dominate elections in terms of the quality of policy decisions in countries with high income inequality as compared to those with low inequality.

Keywords: Elections; sortition; political selection; policy decisions; imperfect information.

JEL codes: D72, D82, H11

*For useful comments and suggestions we thank Francesco Giovannoni, Christian Shlater, Mathan Satchi, and seminar participants at Kent. Makris and Palivos also thank University of Cyprus for its hospitality during the 2019-2020 academic year. The usual disclaimer applies. Email addresses: mmakris.econ@gmail.com (Makris); tpalivos@aueb.gr(Palivos); zachariadis@ucy.ac.cy (Zachariadis).

1 Introduction

Electoral outcomes across the globe and proposed economic, environmental, and health policies over the recent past, have highlighted the potential importance of information for voters and policy makers alike. In this paper, we will be concerned with the quality of policy decision-making under idiosyncratic imperfect information.

More precisely, we compare the quality of policies under a system where the representative is chosen out of the set of politicians/candidates versus a system where the representative is randomly selected from the "demos", i.e., the subset of citizens willing to serve as representatives if selected.[1] We will refer to the latter system as "*selection by lot*", "*sortition*" or "*demarchy*".[2] In our set up, the representative chooses a policy under uncertainty given idiosyncratic imperfect information ("signal"). This information becomes available after selection takes place and hence the selection process is not influenced by it. The quality of the chosen policy depends on the human capital of the representative, which directly affects potential earnings in the private sector as well as the quality of her signal. The differences between the two democratic systems we study affect which citizens become available for selection and thereby the expected quality of the policy decision.

In our model, being the representative carries certain returns such as salary and ego-rents. Citizens also face a cost to establish networks of influence, to exchange favors or promises with interest-groups, and to build a "know-how" on relaxing constraints imposed by existing institutions. These networks, exchanges and know-how ultimately enable the appropriation of a fraction of the polity's total income, with the size of the appropriated fraction depending on the strength of the country's institutions and the extent of corruption. Compared to members of the demos, however, politicians also need to build support to rise through the ranks onto the public platform in order to have a chance to be elected. On the other hand, being a politician typically bears higher perks than being one-among-many demos-members. If the difference in the perks (net-of-costs) from being involved in the selection process under elections and under sortition is large enough, then there is no rent-seeking under sortition. As a result, participation in the selection process under sortition is driven only by the returns from being the representative.

[1] The subset of citizens who are not willing to serve, if selected, forms the group of "idiotes." In ancient Greek the word *idiotes* meant "private persons" as opposed to officials and, according to some scholars, those who were not interested in civic life. From this comes the English word "idiot" which originally meant "ignorant person" and later acquired its modern derogatory meaning. Interestingly, in our model, the idiotes, if any, will be those with the highest level of human capital.

[2] The latter was coined by Hayek (1973) "for the ideal for which democracy was originally adopted" (Volume 3, Chapter 13, page 40). Burnheim (1985) used it to describe sortition.

The differences between the two systems ultimately affect the incentives of citizens to participate in the selection process in place, and hence the associated quality of policy decisions. We find that for high enough returns from being the representative, selection by lot dominates elections because higher returns attract citizens with higher human capital while the probability of winning elections is decreasing in the number of politicians.

Increasing complexity of policy issues is particularly relevant for modern societies faced with economic and science-related issues that require a certain degree of technical competence or expertise. In our context, increasing complexity of policy issues is modeled as a drop in mean expertise affecting the ability of the representative to implement the right policies. Interestingly, we find that an increase in the complexity of policy issues makes it more likely that demarchy dominates elections. The reason is that an increase in complexity reduces the probability of winning elections and thereby the incentives of citizens with high human capital to become politicians.

Calibrating our model, we show that it is more likely for sortition to dominate elections in terms of the quality of policy decisions in countries with higher income inequality - as measured by the Gini coefficient - such as Brazil, Botswana, or even the US as compared to countries such us Sweden, The Netherlands, or Germany. These results are merely indicative as our baseline theoretical setup currently abstracts from many aspects relevant to complex decision making by state representatives. This would involve, for example, extending our baseline to account for decisions being made collectively by a group of individuals rather than treating the decision-making body as a single representative under both sortition or elections. As it is, our baseline model is more applicable to instances where the acquisition of information about which decision is the appropriate one is exogenous, abstracting from the reality that representatives need to use advisors with technocratic expertise and that, therefore, selecting which experts to get advice from could be crucial for the quality of policy decisions in modern democracies where high complexity is often an issue.

Related Literature A historical precedent of selection by lot which foreshadowed the election of representatives in later democratic systems is reported in Aristotle's Athinaion Politeia and goes back to Cleisthenes who around 507 BC created the Council of Five Hundred composed of representatives chosen by lot from small territorial entities in rough proportion to their population, where each citizen served

for one year and was restricted to a maximum of two terms during his lifetime.[3] Acemoglou and Robinson (2016) cite this as an example of inclusive institutions. Modern day examples of "selection by lot" include of course jury duty in the United States as well as citizens' assemblies.[4] The merits of sortition have been discussed by political scientists and philosophers to this day.[5] However, an economic analysis comparing such a system to elections is currently missing. Our paper aims to contribute towards filling this gap by focusing on the implications of each system on the quality of decision-making.

Our paper also contributes to the literature on political selection (see, e.g., Besley 2005), which examines how political systems determine the characteristics of politicians. Besley (2004) studies how representatives' pay affects the quality of decision-making when the pool of politicians is endogenous and finds that higher wages improve the quality of politicians, while Mattozzi and Merlo (2008) studies the effects of monetary incentives on the quality of politicians and show that in their model "an increase in the salary a politician receives while in office decreases the average quality of individuals who become politicians." In our context, an increase in the representative's payoff raises the incentive to be the representative under both systems. However, an increase in the representative's income raises the returns from public life under elections by less than under sortition, since the original increase in income is mitigated by a decrease in the probability of election as more citizens become politicians. Thus, we find that for a high-enough payoff for the representative, drawing a representative from the demos attracts citizens with higher human capital and, thereby, dominates elections in terms of the quality of policy decisions.

The organization of our paper is as follows. We describe our theoretical model next. Following that, in Section 3, we compare elections to sortition in terms of the

[3]For Aristotle, selection by lot was thought to be democratic while election was thought of as oligarchic as he stated in Politics and quoted in Manin (1997), p.43. Selection by lot was also taken seriously by political thinkers such as Montesquieu and Rousseau (Besley, 2006, p.180), which nevertheless favored elections as the means of selecting a natural aristocracy.

[4]Ireland's Citizens' Assembly comprised of 99 random individuals on the issue of abortion in 2016 has been widely seen as a game changer that brought the country to the brink of radical change of its abortion laws, an issue politicians dared not touch. The issue of sortition versus elections has resurfaced in the public debate in recent years, with Hennig (2017) bringing forceful arguments in favor of sortition.

[5]Becker, Szet and Ritter (1976) proposed that half the U.S. House should be selected by lot. Besley (2006, p.179) outlines the merits of selection by lot among the citizens as guaranteeing the widest possible access to public office thus maintaining a unity of purpose in the community, and that representatives have everyday life experience. Barber (1984), Burnheim (1985), Callenbach and Phillips (1985), Stone (2000) and Bouricius (2013) discuss the benefits of selection by lot more extensively. These include what Callenbach and Phillips (1985) describe as "descriptive representation", the notion that if a group comprises a certain fraction of the population then that same fraction of the legislature representing the population should belong to that group.

quality of decisions. The fourth section presents our calibration exercise, and the last section briefly concludes.

2 The Model

A polity of N citizens needs to make a collective decision $a \in \{0, 1\}$ under uncertainty about the state of the world $x \in \{0, 1\}$, with $\Pr(x = 1) = \rho$. Citizens have identical preferences $u(a, x)$, which are such that the polity would like to "match the state".[6] Let

$$b_x = u(x, x) - u(y, x) > 0 \text{ for all } y \neq x$$

One example of such a decision problem is whether the citizens of a small country should match their lifestyle to the probabilistic signal they receive regarding climate change.

The polity selects a representative to choose the policy that will be followed. The representative acquires a signal about the unknown state of the world $\sigma \in \{0, 1\}$ after they are selected. The precision of the signal is given by

$$t(\theta) = \Pr(x = \sigma | \sigma, \theta),$$

that is, by the probability of the signal revealing the truth. The precision of the signal a citizen receives can depend on their education, their innate ability to analyze information, their experience, etc., which are all captured here by the variable θ. We will refer to θ as the human capital of an agent, which, from now on, is used as a proxy for the "expertise" in discovering the true state of the world. Given this interpretation, we assume that $t(.)$ is non-decreasing. Human capital (and hence the precision of the signal) is identically and independently distributed across citizens according to the cdf F (with density f) over a subset $\Theta = (\underline{\theta}, \overline{\theta})$ of the real line. We assume that $t(\theta) \geq \max\{\rho, 1 - \rho\}$.[7] Given the binary nature of the state of the world, we can assume without loss of generality that $\rho \geq 1/2$. The "complexity" of the policy issue for (or the inverse of the ability as an expert of) a citizen with precision $t(\theta)$ is thus captured here by the difference $t(\theta) - \rho$.

[6] A possible extension would be to have some citizens with "extreme preferences" in that they prefer a certain action regardless of the state. This could capture the presence of ideological voters on either end of the policy spectrum (some always preferring $a = 1$ and others always preferring $a = 0$), alongside the rest of the voters who would like to match the state. A conjecture is that if the society is not very polarized (and so the majority cares relatively more about policy quality rather than the ideology of the representative), then the results below would still hold qualitatively.

[7] An application of Bayes rule shows this inequality requires that, under expertise θ, signal σ is more likely to occur when $x = \sigma$ than when $x \neq \sigma$, for all σ and θ. The proof is in Appendix A.3.

In this society, the ideal policy for each citizen depends on their human capital and their signal. We assume that $t(\theta) \geq \max\{\frac{b_1}{b_1+b_0}, \frac{b_0}{b_1+b_0}\}$ for all θ. Under this assumption, the ideal policy for a citizen with human capital θ is to "match their signal," i.e., set $a = \sigma$.[8] The higher the human capital θ of a citizen, the higher the quality of the decision taken by this citizen who acts upon their signal. Thus, agents with higher human capital in this model would make better decisions, if they were responsible for choosing the policy on behalf of the polity. We assume that θ is citizens' private information.[9]

Citizens may also differ with respect to their income in the private sector of the economy. In the spirit of the so-called AK model of Frankel (1962), we assume that citizens' income y is proportional to their human capital, that is, $y = A\theta$.[10] Through normalization of units, we henceforth assume that $A = 1$.

The selection of the representative will be according to one of the following two democratic systems. The first, which we will refer to as an "electoral" system, has the representative selected from the subset of citizens that are "politicians." More specifically, given that all citizens have the same preferences and human capital is private information and i.i.d., voters are indifferent over who will be the representative and so we assume, following Besley (2006, p.115), that they vote by randomizing over the set of politicians. Politicians are citizens themselves who have chosen to invest in a political career instead of entering the private sector and becoming entrepreneurs, in which case they would have earned θ. The second democratic system we consider, which we will refer to interchangeably as "sortition", "demarchy" or "selection by lot", has the representative randomly selected out of the subset of citizens that comprise the demos. Demos stands here for the subset of citizens who are willing to act as representatives if selected. If not selected, then they continue being entrepreneurs, i.e., they participate in the private sector and earn θ. Under the latter system, there will be no politicians in equilibrium as becoming one carries a cost. We emphasize that both the set of politicians and the demos are endogenously determined in our model. We return to this shortly.

[8] The proof can be found in Appendix A.3.

[9] Human capital as used here consists of observable as well as non-observable characteristics. Examples of the former are number of years and quality of education, and examples of the latter are effort, innate ability, habits and personal traits.

[10] We note that even if the product $A\theta$ is observable, the individual components are not.

2.1 Elections

Becoming a politician carries a cost $c(\theta) > 0$.[11] This cost captures the investment that is necessary for regular citizens to build support to rise through the ranks onto the public platform, establish networks of influence, build a "know-how" on relaxing constraints imposed by existing institutions, and exchange favors or promises with interest-groups as part of the process of becoming politicians with the hope that they will eventually come to be the representative. These networks, know-how, and exchanges, ultimately allow the appropriation of a fraction μ of polity's realized total income of entrepreneurs, which is a random variable. Denote with $\overline{y}$ the expected total income of entrepreneurs. The appropriated resources are taken away from the polity's raised tax revenues. The remaining tax revenues are used to provide a public good (or finance a per capita lump sum), which is suppressed hereafter from the description to avoid notational cluttering. The residual fraction $1 - \mu$ captures in a stylized manner how difficult it is for the representative to appropriate the economy's income. This degree of difficulty will naturally depend on the strength of institutions and the extent of corruption.[12] The cost of being a politician is also net of any tax benefits from not being an entrepreneur. This would tend to make the cost function decreasing in human capital. On the other hand, if individuals choose in an ex ante stage between investing resources in political capital necessary for becoming a politician versus accumulating expertise via investments in human capital, then resources invested in political capital are diverted away from the accumulation of expertise. While, we do not consider human capital (expertise) versus political capital accumulation explicitly in our model, we do consider the case where the cost of becoming a politician is higher for high expertise types, consistent with substitutability between human capital and political capital accumulation. Hence, this channel would push the cost of becoming a politician to be increasing in human capital. The case where the cost of entering politics is constant, and in particular independent of the politician's human capital, is essentially identical

[11] The assumption here is that without paying the investment cost to become a politician, a candidate will lose an election with certainty; thus, if there is even a small cost of running for election, a non-politician would never become a candidate.

[12] The availability of rents or the magnitude of how costly it is to appropriate these should generally relate to the present degree of executive constraints, captured by Polity IV measures such as xconst. Besley and Reynal-Querol (2017) and Besley and Mueller (2018) argue that only when the latter measure is at its maximum (xconst=7), e.g. in the case of Sweden, can we be confident that such executive constraints bind. In all other cases, these will not bind and the presence of rents for politicians will realistically be the likely outcome in most representative democracies under elections.

to the increasing-cost case analyzed below.[13] We assume in what follows that the cost-function is non-decreasing. However, we also consider the case where the cost of becoming a politician is decreasing in human capital and show in Appendix A.1 that under reasonable assumptions (that ensure uniqueness of equilibrium) our main results remain intact.[14] It should be noted that we also view the cost $c(\theta)$ as being net of any constant income, perks, or ego-rents enjoyed by citizens who belong to the group of politicians. In other words, the cost $c(\theta)$ could include a constant term that captures these pecuniary or non-pecuniary rewards.

We now turn to the occupational choice of citizens. Under sortition, citizens choose whether they will be part of the demos, i.e., be willing to act as a representative if selected to be one, or whether they will be idiotes, i.e., entrepreneurs but not part of the demos. Here, being a member of the demos requires no private cost.[15]

Under the electoral system, citizens choose whether they will become politicians or entrepreneurs. The payoff of an entrepreneur of type θ net of the (expected) utility from the policy chosen by the representative is

$$U^e(\theta) = \theta - \mu\overline{y}\mathbf{1},$$

where $\mathbf{1}$ is an index function that takes the value one if selection takes place by means of an election, and zero if selection takes place by means of sortition.[16] The burden of

[13]The case where the cost of becoming a politician differs among citizens in a manner that is unrelated to expertise, ability or entrepreneurship is quite plausible. To begin with, anecdotal evidence exists for nearly every country in the world whereby certain family names have repeated entries in the list of public officials over the past several decades. Empirical evidence from the US, e.g., Dal Bo et al. (2009) shows that "holding legislative power for more than one term doubles the probability that a politician will have a relative entering Congress in the future," for reasons unrelated to original differences in ability across families. This suggests that the entry cost is lower for potential politicians with close links to other politicians, rather than for those with higher human capital or ability. Merlo et. al (2008) investigate skill endowment, including observable characteristics, e.g., education and characteristics that are hard to measure directly such as ability, and find that there was "a deterioration of the quality of elected legislators, both with respect to their level of education as well as their ability score for the period 1994-2006 relative to 1981-1994" in the case of Italy. Thus, whereas characteristics such as education and ability are becoming increasingly important for entrepreneurship as the technology frontier shifts out over time, these became less important for elected politicians in Italy which again supports the notion that the cost of becoming a politician is not decreasing in one's human capital.

[14]In particular, assuming the cost of becoming a politician is sufficiently flat, namely, $c'(\theta) > -1$ for all θ, we show that proposition 1 below holds also for the case of decreasing cost.

[15]A public investment may be needed instead for those in the demos (or, equivalently, for those randomly selected to be representatives), but we assume that the cost of this is equally shared by every citizen and so does not enter into the private decision of citizens to enter the demos. In Appendix A.1, we allow for a private cost paid by demos members that ensures ability to appropriate resources if selected, and where politicians pay an extra cost as compared to that paid by demos members since, under sortition, no cost for building support to rise through the ranks onto the public platform is necessary to have the chance of being selected.

[16]We relax this latter assumption in the last section of Appendix A.1 where we allow demos

the (expected) income appropriated by the elected politician $\mu N\overline{y}$ is equally shared by all N citizens; hence, each one bears (in expectation) the average cost $\mu\overline{y}$. The payoff of a politician of type θ who is not elected to be the representative net of the expected utility from the policy chosen by the representative is

$$U^p(\theta) = -c(\theta) - \mu\overline{y}.$$

Finally, the payoff of a politician of type θ who is elected to be the representative net of the expected utility from the policy they chose is

$$U^w(\theta) = R - c(\theta) + \mu(N-1)\overline{y}$$

where $R > 0$ denotes all the pecuniary (after-tax) and non-pecuniary benefits awarded to the winner of the election, i.e. to the elected representative.

Let ω^p denote the probability of a politician becoming the representative. Let also $\Theta^p(\omega^p)$ be the set of types who are politicians, and hence not entrepreneurs, when ω^p is the probability of a politician becoming the representative. That is,[17]

$$\Theta^p(\omega^p) = \{\theta | \theta \leq \omega^p(R + \mu N\overline{y}) - c(\theta)\}. \tag{1}$$

Crucially, note that the size of this set is endogenous and depends on the cost of becoming a politician, the probability of being elected, and the size of the rents that an elected representative can embezzle. Notice, thus, that under election the expected utility from the policy chosen by the representative is[18]

$$W^p = \rho u(0,1) + (1-\rho)u(1,0) +$$

members to obtain rents, and show that our main results remain. We note, however, that one could conjecture that in a dynamic version of our theory setup the strength of institutions converges to its maximum level in a long-run equilibrium under sortition so that no rents are available for appropriation. That is, given that sortition resets the political environment each time in that, e.g., randomly selected leaders would be unlikely to be drawn a second time in any realistic setting, there would be less inherent strife with institutions under sortition as compared to elections and this would result in the strengthening of institutions over time.

[17] To understand the definition of the set Θ^p in the main text note that under election, the expected payoff of being a politician is

$$(1-\omega^p)[-c(\theta) - \mu\overline{y}] + \omega^p[R - c(\theta) + \mu(N-1)\overline{y}] + W^p = -c(\theta) - \mu\overline{y} + \omega^p(R + \mu N\overline{y}) + W^p,$$

while the expected payoff of an entrepreneur is

$$\theta - \mu\overline{y} + W^p,$$

where W^p is the expected utility from the policy chosen by the representative under election.

[18] The proof can be found in Appendix A.3.

$$\rho b_1 E[\Pr(\sigma = 1|x = 1, \theta)|\theta \in \Theta^p(\omega^p)] + (1-\rho) b_0 E[\Pr(\sigma = 0|x = 0, \theta)|\theta \in \Theta^p(\omega^p)].$$

In the expression above, $\rho u(0,1) + (1-\rho)u(1,0)$ is the expected utility from the policy chosen when the representative makes a mistake, and $b_x E[\Pr(\sigma = x|x, \theta)|\theta \in \Theta^p(\omega^p)]$ is the expected gain when the true state is x and the representative uses their human capital (knowledge or expertise) to choose this policy under election.

Let now ω^{p*} be the equilibrium probability of a politician becoming the representative, as seen from a single politician's point of view. To understand the formula below that gives ω^{p*}, observe first that, from an ex ante point of view, the probability that the typical citizen will be a politician is equal to the total probability that their type θ is in the set $\Theta^p(\omega^{p*})$. Moreover, from the point of view of someone who does not know the type of a given citizen, the distribution characterizing the probability of that citizen becoming a politician is the Bernoulli with "success probability" $\int_{\theta \in \Theta^P(\omega^{p*})} f(\theta) d\theta$. Recall also that types are private information. Given that types are i.i.d., we then have that from the point of view of a politician, the distribution of the n *other* citizens becoming rival politicians is the Binomial with size parameter $N-1$ and "success probability" $\int_{\theta \in \Theta^p(\omega^{p*})} f(\theta) d\theta$. Finally, given the assumed random selection among the politicians for the role of the representative, we have that if there are also n *other* rival politicians, then the probability that a given politician is elected is equal to $1/(n+1)$. Putting all these together, we then have that the probability assigned by a politician on the event that they will become a representative, ω^{p*}, is given implicitly by

$$\omega^{p*} = \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\theta \in \Theta^p(\omega^{p*})} f(\theta) d\theta)}{n+1}, \tag{2}$$

where the notation $b(n|N,p)$ stands for the Binomial probability mass function with size parameter N and probability of success p. Note also, for future use, that, since the Binomial distribution $b(n|N,p)$ increases stochastically with p and the function $1/(n+1)$ decreases in n, it follows that $\omega^{p*} \in [1/N, 1]$. Moreover, if $\Theta^p = [\underline{\theta}, \theta]$, then $\lim_{\theta \to \underline{\theta}} \omega^{p*} = 1$ and $\lim_{\theta \to \overline{\theta}} \omega^{p*} = 1/N$. The intuition for these is clear. In the first case, as θ approaches $\underline{\theta}$, the probability that someone, randomly selected, will become a politician is zero, and so is the probability of $n = 1, 2, \ldots N-1$ other citizens; hence, the probability that $n = 0$ other citizens will be politicians is one. It follows, then, that the probability as viewed by a given politician that she will be elected as the representative is one, given that she will be the only politician. Similarly, in the second case above, as θ approaches $\overline{\theta}$, the probability that someone,

again randomly selected, will be a politician is one, as is the probability for all $N-1$ other individuals. The probability then that a given politician will be selected as representative is $1/N$, as she will be one among N politicians.

Assume that

$$\underline{\theta} < \frac{1}{N}(R + \mu N \overline{y}) - c(\underline{\theta}) \tag{3}$$

and

$$\bar{\theta} > R + \mu N \overline{y} - c(\overline{\theta}) \tag{4}$$

The first inequality (3) states that the politician's investment cost is not too high, so that the expected (net) payoff for someone who is of the lowest human capital type will be above the lowest income in the private sector $\underline{\theta}$; this precludes the case where no type becomes a politician. The second inequality (4) ensures that this same investment cost is sufficiently high, so that the expected (net) payoff for politicians who are of the highest human capital type is below the highest income in the private sector $\overline{\theta}$; this rules out the case where all types become politicians. These two inequalities are *sufficient* for the existence of an equilibrium type θ. Consider

Proposition 1: *There exists a unique threshold θ^p such that a citizen of type θ is a politician if and only if $\theta \in \Theta^p = [\underline{\theta}, \theta^p]$.*

Proof: See Appendix A.1. In Appendix A.1, we also prove that the same result holds if we allow for decreasing cost of becoming a politician insofar as it is sufficiently flat, namely, $c'(\theta) > -1$ for all θ.

Observe that θ^p is increasing in R, μ, N, and $\hat{\theta}$. The reason is simply that an increase in any of these model parameters raises the income of the elected politician and hence gives more incentives to citizens with higher human capital to become politicians. More on this below.

2.2 Sortition

Becoming a politician requires an investment on the part of the citizen and being elected as a representative entails being a politician. Sortition, on the other hand, simply selects a representative from the set of citizens who are willing to be part of the demos and, as we have already mentioned, being in the demos requires no private investment.[19] We assume that all the benefits, R, offered to an elected politician are also available to the selected decision-maker under sortition. Thus, the only exogenous differences between election and sortition are that: a) politicians emerge only under the selection mechanism of election (due to the cost of becoming a

[19]We relax this assumption in Appendix A.1, where we allow for a private cost paid by demos members that ensures ability to appropriate resources if selected.

politician); and b) elected politicians have access to (additional) net rents $\mu(N-1)\overline{y}$. These differences will have an impact on the set of types who become available for selection and thereby on the quality of the policy decision.

Let now Θ^d be the set of types who become members of the demos.[20] That is,

$$\Theta^d = \{\theta | \theta \leq R/(1-\tau)\}. \tag{5}$$

where τ is the exogenously given income tax rate. Note that Θ^d is independent of the probability of a demos-member becoming the representative (and thereby of the equilibrium size of the demos) because the non-selected members of the demos remain in the private sector.

Assume that

$$R > \underline{\theta}/(1-\tau). \tag{6}$$

Inequality (6) ensures that the reward for the selected representative under sortition is greater than the lowest income in the private sector $\underline{\theta}$; this precludes the case where no type wants to be in the demos.

The following proposition is immediate.

Proposition 2: *There exists a unique threshold θ^d such that a citizen of type θ belongs to the demos if and only if $\theta \in \Theta^d = [\underline{\theta}, \theta^d]$.*

Obviously, if $R \leq \overline{\theta}/(1-\tau)$, then $\theta^d = R/(1-\tau)$, and θ^d is an interior point of Θ. Moreover, an increase in R trivially raises the threshold θ^d. If, on the other hand, $R > \overline{\theta}/(1-\tau)$, then $\Theta^d = \Theta$, that is, the entire polity belongs to the demos. Unlike the case encountered in the previous subsection and ruled out by condition (4), in which all types become politicians, the present case where the demos consists of the entire polity, i.e., all types belong to the demos, is interesting and we analyze it in the next section.[21]

Under sortition, the expected utility from the policy chosen by the representative

[20] To understand the definitions of the set Θ^d in the main text note first that under sortition the expected payoff of being a member of the demos is

$$\omega^d R + (1-\omega^d)\theta(1-\tau) + W^d,$$

where ω^d is the probability of a demos-member becoming the representative and W^d is the expected utility from the policy chosen by the representative under sortition. Note also that the expected payoff of an entrepreneur who is not a member of the demos, i.e., a member of the group of idiotes, is

$$\theta(1-\tau) + W^d = \omega^d\theta(1-\tau) + (1-\omega^d)\theta(1-\tau) + W^d.$$

[21] We thus impose only one boundary condition for sortition while we had to impose two in the case of elections.

is

$$W^d = \rho u(0,1) + (1-\rho)u(1,0) +$$

$$\rho b_1 E[\Pr(\sigma = 1|x = 1, \theta)|\theta \in \Theta^d] + (1-\rho) b_0 E[\Pr(\sigma = 0|x = 0, \theta)|\theta \in \Theta^d].$$

In the expression above, $\rho u(0,1) + (1-\rho)u(1,0)$ is, once again, the expected utility from the policy chosen when the representative makes a mistake. In addition, $b_x E[\Pr(\sigma = x|x, \theta)|\theta \in \Theta^d]$ is the expected gain when the true state is x and the representative uses their knowledge to choose this policy under sortition.

3 Comparing the two Selection Mechanisms

Given the above results and assuming that $\Pr(\sigma = x|, \theta)$ is non-decreasing in θ to capture the definition of θ as human capital,[22] we have that sortition is better than election in terms of the quality of policy decisions (and also in terms of expected welfare "behind a veil of ignorance") if

$$\rho b_1 E[\Pr(\sigma = 1|x = 1, \theta)|\theta \in \Theta^d] + (1-\rho) b_0 E[\Pr(\sigma = 0|x = 0, \theta)|\theta \in \Theta^d] >$$

$$\rho b_1 E[\Pr(\sigma = 1|x = 1, \theta)|\theta \in \Theta^p(\omega^*)] + (1-\rho) b_0 E[\Pr(\sigma = 0|x = 0, \theta)|\theta \in \Theta^p(\omega^*)]. \tag{7}$$

A sufficient condition for this is that the conditional distribution $F(\theta|\theta \in \Theta^d)$ dominates the conditional distribution $F(\theta|\theta \in \Theta^p(\omega^*))$ according to First-Order Stochastic Dominance. In light of the results derived above regarding the structure of the sets Θ^p and Θ^d, we have directly that the regime with the best policy choice is the one with the highest threshold. We distinguish then between two cases and we start with the simplest one.

Case 1: $R \geq \overline{\theta}/(1-\tau)$. In this case $\theta^d = \bar{\theta}$ and hence $\Theta^d = [\underline{\theta}, \overline{\theta}]$; all citizens

[22] Note, using Bayes rule, that

$$t(\theta) \equiv \Pr(x = 1|\sigma = 1, \theta) = \frac{\rho \Pr(\sigma = 1|x = 1, \theta)}{\rho \Pr(\sigma = 1|x = 1, \theta) + (1-\rho)[1 - \Pr(\sigma = 0|x = 0, \theta)]}$$

and

$$t(\theta) \equiv \Pr(x = 0|\sigma = 0, \theta) = \frac{(1-\rho) \Pr(\sigma = 0|x = 0, \theta)}{(1-\rho) \Pr(\sigma = 0|x = 0, \theta) + \rho[1 - \Pr(\sigma = 1|x = 1, \theta)]}$$

Therefore, $t(\theta)$ is increasing if and only if $\frac{1-\Pr(\sigma=0|x=0,\theta)}{\Pr(\sigma=1|x=1,\theta)}$ and $\frac{1-\Pr(\sigma=1|x=1,\theta)}{\Pr(\sigma=0|x=0,\theta)}$ are decreasing in θ. Moreover, equating the right-hand sides of the above equalities we have that it must be that:

$$\rho^2 \Pr(\sigma = 1|x = 1, \theta)[1 - \Pr(\sigma = 1|x = 1, \theta)] = (1-\rho)^2 \Pr(\sigma = 0|x = 0, \theta)[1 - \Pr(\sigma = 0|x = 0, \theta)]$$

belong to the demos. Clearly, sortition is superior to election. Indeed, if the payoff of the selected representative under sortition, R, is equal to or greater than the maximum reward in the private sector, $\overline{\theta}$, divided by $(1-\tau)$ then even the citizens of the highest human capital will be part of the demos. This maximizes the expected expertise of the selected representative. Any attempt to do the same under election, by raising the payoff of the elected politician, e.g., setting $R \geq \overline{\theta} - \mu N \overline{y} + c(\overline{\theta})$, will result in the unworkable case where every citizen is a politician and thus there is no production (see 4).

From the case above, it also follows that if participation in the demos is mandatory, as with jury service, then sortition is superior to election. Once again, this is so because sortition in this case will result in a representative from the entire domain of the distribution of human capital, whereas the elected politician will be selected from a right-truncated distribution. Nevertheless, we dismiss this case because one cannot force someone to do something they do not want to do.[23]

The following case is more pragmatic and hence more interesting.

Case 2: $R < \overline{\theta}/(1-\tau)$. In this case, we have that $\theta^d > \theta^p$ and hence sortition dominates election in terms of policy quality if and only if

$$R/(1-\tau) > \left[\sum_{n=0}^{N-1} \frac{b(n|N-1, F(\theta^p))}{n+1}\right](R + \mu N \overline{y}) - c(\theta^p). \tag{8}$$

Recall that the higher the human capital θ, the higher the quality of the decision taken by a citizen with human capital θ who acts upon their signal. Hence, between the two mechanisms dominates the one that attracts the highest human capital. Inequality (8), then, requires that the payoff of the representative selected under sortition is higher than the payoff of the elected politician. Recall that the former is constant, and, in particular, independent of the probability of being selected, whereas the latter depends on the probability of election.

Institutions

If $\lim_{\theta \to \underline{\theta}} c(\theta) > \mu N \overline{y}$, i.e., the entry cost is always higher than the additional funds that the elected politician peculates, then, as expected, sortition is superior to election. A special case of this result is when institutions are so strong that $\mu = 0$. More generally, the lower μ, due to stronger "institutions", the lower the RHS of the above inequality becomes. Additionally, the LHS of the above inequality is independent of μ. Therefore, the stronger the institutions, the more likely it is that selection by lot dominates election. The reason is that a lower μ leads to lower "extortion rents" and this would diminish the incentive of citizens with high human

[23] A search on the internet using the phrase "how to get out of jury" yields over 110,000 results.

capital to enter public life under election and would thereby lead to policies of lower quality.

Polity Size

The lower the size of the polity, N, the lower the RHS of the above inequality becomes. Additionally, the LHS of the above inequality is independent of N. Thus, the lower the size of the polity, the more likely it is that sortition dominates election. The reason is that a lower N leads to lower "extortion rents" and thus diminishes the incentive of citizens with high human capital to enter public life under election, leading to policies of lower quality. The implication is that sortition is more likely to dominate elections for smaller entities. This might explain the case of ancient Athens. It might also serve as food for thought for the democratic processes most appropriate for small democratic entities such us municipalities.

Representatives pay

Besley (2004) investigates how paying representatives affects the quality of decision-making when the pool of politicians is endogenous and finds that higher wages improve the quality of politicians in the pool of candidates available for public office and raise the fraction of congruent politicians who put themselves forward for office.[24] Matttozzi and Merlo (2008) investigate the effects of monetary incentives on the quality of politicians and show that in their model "an increase in the salary a politician receives while in office decreases the average quality of individuals who become politicians."[25] In our context, an increase in the representative's pay, R, raises the incentives to be the representative under both mechanisms and hence increases both thresholds θ^d and θ^p. Nevertheless, the higher R is the higher the LHS of inequality (8) becomes relative to the RHS. Therefore, the higher the "representative's income" R, the more likely it is that sortition dominates election. The reason is that an increase in R raises the returns from public life under sortition by more than under elections, and therefore sortition attracts citizens with higher expertise. Intuitively, the original increase in θ^p is mitigated by a decrease in the probability of being elected in the case of an electoral system. On the contrary, under sortition, the threshold θ^d is independent of the probability of being selected.

[24]He also shows that an increase in the wage of politicians increases the probability of congruence between voter preferences and policy outcomes, and that increasing the value of holding office increases voters' welfare.

[25]In their context, "an increase in the salary in the political sector makes politics a relatively more attractive option for all levels of political skills, thus lowering the quality of the worst politician" and also makes better incumbent politicians willing to remain in politics as "wages in politics are now better relative to the market wages" so that "the overall impact on the average quality of career politicians depends on which of the two effects dominates".

Complexity of policy issues

Finally, let us study the impact of an increase in mean expertise $\hat{\theta}$. Such an increase captures the scenario where the complexity of the policy issue for citizens is reduced on average; in the case we focus on here, this means that the polity becomes more productive on average and the average political investment cost is higher. To see the effects, note first that an increase in mean expertise is represented here by an increase in the area above $F(.)$ (up to 1) while maintaining $F(\underline{\theta}) = 0$ and $F(\overline{\theta}) = 1$.[26] Such a change has the following consequence in the dominance of sortition. Note that when $F(\theta)$ goes down for all $\theta \in (\underline{\theta}, \overline{\theta})$, then $\sum_{n=0}^{N-1} \frac{b(n|N-1,F(\theta^p))}{n+1}$ goes up. This increases the value of θ^p. In addition the size of rents that the elected politician appropriates for any given μ and N goes up and this raises the value of θ^p even more. On the other hand, the threshold θ^d does not depend on the distribution of human capital. We then have that a reduction in the complexity of the policy issue makes it less likely that sortition dominates election. The reason is that for the given quality-threshold under sortition, switching to election would attract higher quality politicians due to the greater probability of election and the higher prize for the elected. Conversely, an increase in the complexity of the policy issue (or a reduction in mean expertise $\hat{\theta}$) for citizens makes it more likely that sortition dominates election. As increasing complexity is an important aspect of modern day policy-making spanning economic and science-related policy issues pertaining to, e.g., health and the envirorment, we view this as an empirically relevant case.

4 Calibration

In principle, we could find some data (of varying quality) for some countries on politicians' salaries. However, it would be very difficult to estimate perks, rent-seeking costs and the "strength of institutions" for a large number of countries. This is a problem in general in the applied political economy literature. Our model could provide an index of "attractiveness" of being a politician by using the equilibrium condition for θ^p and the assumption of a linear cost function $c(\theta) = \gamma\theta$, $\gamma > -1$, if we had θ^p and F. In fact, for given θ^p and F, and under linear cost in a neighbourhood around θ^p, the equilibrium condition that gives implicitly θ^p can be rewritten as

$$\frac{R + \mu\overline{y}}{1+\gamma} = \frac{\theta^p}{\sum_{n=0}^{N-1} \frac{b(n|N-1,F(\theta^p))}{n+1}}$$

[26] Equivalently, this amounts to changing the cdf F in a First-Order Stochastic Dominance sense, i.e. $F(\theta)$ going down for all $\theta \in (\underline{\theta}, \overline{\theta})$.

giving us the index in question $\frac{R+\mu\overline{y}}{1+\gamma}$.

It turns out that we can calibrate θ^p and F from existing data on income in the following way, regardless of the functional form of $c(\theta)$. First, using Borel's law of large numbers we have for large N that $f(\theta)$ approximates well the proportion of citizens with human capital, and that $\frac{\overline{y}}{N} \simeq \hat{\theta}$. Second, we assume here, with some abuse of notation, that human capital is given by $\underline{\theta} + \theta$ where θ is distributed according to the pdf $f(\theta)$, given by the Pareto-Lognormal which is characterized by three parameters (M, σ, λ). The parameter λ governs the Pareto right tail of the distribution, whereas the parameters (M, σ^2) govern the Lognormal left tail of the distribution. Third, we normalize M so that $\hat{\theta} = 1$. Fourth, we assume that the observed income distribution for a country is, in effect, described in terms of our model by the conditional distribution $F(\theta|\theta > \theta^p)$. Finally, we calibrate the parameters $(\theta^p, \sigma, \lambda)$ – which characterize $F(\theta|\theta > \theta^p)$ completely under the above assumptions – to match World Development Indicators data on the Gini index, Y_3 and Y_1/Y_5, where Y_i, $i = 1, ..., 5$, is the share of income held by the i^{th} poorest 20% of the population. The parameters $(\theta^p, \sigma, \lambda)$ are calibrated by minimizing the sum of the squared deviations of the simulated Gini index, Y_3 and Y_5/Y_1 from their corresponding values in the data. All squared deviations of target moments in percentage terms are between 3% and 9.2%, with most of them below 3% and only two being around 9%.

We note that in the above calibration exercise we can associate the mean income from the data, shown in the first column of Table 1, with the productivity parameter A, and that, given our definition of human capital in this section, $\underline{\theta}+\theta$, the minimum income in a country is $A\underline{\theta}$. Thus, the actual threshold level of entrepreneurs' income that defines the marginal politician as a percentage of mean income is $\underline{\theta} + \theta^p$ rather than θ^p shown in the third column of Table 1.

Given that sortition dominates elections for all benefits that accrue to the representative R such that $R(1-\tau) > \theta^p$, we can associate a low θ^p with an economy where it is more likely that sortition dominates elections (in terms of the quality of decisions), and vice versa. Here, θ^p should be interpreted as a deviation from the minimum salary in mean income percentage terms. With this interpretation, our calibration results in Table 1 indicate that it is more likely for sortition to dominate elections when income inequality as measured by the Gini coefficient is high. More specifically, as shown in Table 1, sortition is more likely to dominate elections in Brazil, Botswana or even the US as compared to countries such as Germany, the Netherlands, or Sweden. Given the strong institutions of Sweden, our finding in the calibration exercise that sortition is less likely to be optimal there, is not necessarily

Country/Variable	Mean Income	Gini	θ^p
Sweden	43642.48047	27.41538429	0.11118574
Netherlands	45862.97656	28.63333321	0.110135584
Germany	38515.55469	30.31111145	0.121350222
Korea, Rep.	29419.08008	31.89999962	0.110205351
Canada	36082.82813	32.83333206	0.056821027
United Kingdom	37246.88672	34.10833359	0.08018896
Italy	36306.76172	34.40000153	0.067531279
Lithuania	21862.01367	35.3538475	0.067743228
United States	46687.61328	40.45555496	0.03759439
Uruguay	16945.08203	42.46666718	0.037835149
Argentina	15627.44141	46.70384598	0.024593653
Dominican Republic	9592.858398	48.75	0.021188991
Brazil	12342.79199	56.04166794	0.007721084
Botswana	11643.85352	59.82499695	0.010708413

Table 1: Calibrated θ^p interpreted as a deviation from the minimum salary in mean income percentage terms. Mean income is PPP-adjusted GDP per capita in constant 2011 international US dollars from the WDI averaged over the period from 1990 to 2018. Sampled countries are ranked in terms of their Gini index averaged over 1990-2018. These countries score high in indices of democratic institutions. All of them score a value of 1 in both the Democracy indices from Boix et al. (2013) and Cheibub et al. (2010), with the exception of Botswana for the second index, and with Argentina, Korea, Botswana and the Dominican Republic ranked relatively lower than the other countries in terms of the Electoral democracy index from Teorell et al. (2019) and the Institutionalized democracy index from POLITY IV (Marshall, Gurr and Jaggers 2018). We describe these variables in detail in Appendix A.2.

incompatible with data due to the fact that its Gini index is so low. In other words, election can be better under strong institutions if income inequality is sufficiently low.[27]

5 Conclusion

Electoral outcomes across the globe and proposed economic, environmental, and health policies over the recent past have highlighted the potential importance of information regarding the appropriate policies, on behalf of voters and policy makers alike. We have considered the quality of policy decision-making in representative democracies under imperfect information about what constitutes a good decision, and compared it to that in a democratic system where representatives are selected by lot from the demos.

We have found that an increase in the complexity of the policy issue makes it more likely that the latter democratic system dominates elections in terms of the quality of decision-making. Furthermore, we have shown that if the reward from public office is high enough, drawing a representative from the demos dominates elections in terms of the quality of policy decisions as it attracts citizens with higher human capital and thereby better-quality information. Finally, we have calibrated our model and shown that random selection from the demos is more likely to dominate elections in terms of the quality of policy decisions in countries with high income inequality as compared to those with low inequality. Our paper contributes to the current debate on elections versus sortition by offering an economics perspective as to why sometimes elections could be better than sortition and sometimes not.

To isolate the effects on political participation and thereby the quality of decision-making of the differences between the two regimes under comparison, we have made a number of simplifying assumptions. For instance, we treated the decision-making

[27]If F First-Order Stochastic Dominates (FOSD) G then Gini(F) $<$ Gini(G). Proof: F FOSD G implies F Second-Order Stochastic Dominates G, which implies that the Lorenz curve of F $>$ Lorenz curve G (Atkinson 1970) which implies Gini(F) $<$ Gini(G). In our case then, we have that if F(.) goes down almost everywhere, making sortition less likely to be optimal, then it is the case that F's Gini index goes down. This explains the case of Sweden in a weak sense: It has extremely strong institutions and so for sortition to be less likely to be optimal it must be that $F(\theta^p)$ is low, which implies an extremely low Gini index. This is confirmed by the data. The other direction is not always true. That is, we cannot say that Sweden has a low Gini Index and this pushes for sortition to be less likely to be optimal. Our findings for Sweden are consistent with Dal Bo et. al (2017) who show that "Swedish politics attracts competent people who are not restricted to the scions of elite families". They find that for Sweden, politicians are on average significantly smarter than the population they represent even when conditioning on family background, suggesting that "individual competence is key for selection" in this case. They thus conclude that "it is possible for democracy to generate competent and socially representative leadership."

body as a single representative. Extending our model to a setup where decisions are made collectively by a body comprised of several representatives, with the number of representatives possibly increasing with country size, is a very interesting avenue for future research.[28] Related, yet another important extension of our work would be to allow for the co-existence of upper and lower decision-making bodies and thereby the possibility of the co-existence of elected and randomly selected bodies each deciding on different policy dimensions.

In our model, the acquisition of information about which decision is the appropriate one is exogenous. In reality, however, representatives use advisors, i.e., citizens with technocratic expertise. In fact, selecting which experts to get advice from is a key issue for the quality of policy decisions in modern democracies where high complexity is often an issue.[29] A very interesting question then is: how do the differences between sortition and elections affect the selection of advisors and hence information acquisition? A conjecture – to be examined in future work – is that as politicians often owe favors to some individuals and typically value certain groups more than others due to ideology or because their future political success depends on them, an elected politician might be more likely to choose from a narrow pool of advisors and less likely to depend on objective technical expertise from the best available specialists as compared to the pool of advisors a randomly selected group of citizens would draw from. The latter could have an advantage over politicians in terms of the quality of decisions for complex issues faced by modern societies in that they would be more likely to draw from a larger pool of technocrats with higher expertise. This link between quality of experts and the way representatives are selected has not been included in our baseline model to create a level playing field for comparing election versus sortition, as it would only strengthen results in favor of the quality of experts and policy decisions under sortition.[30]

To better understand and evaluate the comparison between sortition and elec-

[28]This would alleviate the potential problem of extreme outcomes under sortition where drawing a single extremist poses a threat to the quality of decisions. The presence of large rents under elections and the resulting pursuit of political power favors single-leader outcomes so that the more relevant question is perhaps whether a group of individuals that previously stayed away from politics going about their own business would do worse than a professional politician once in office. In this case, the relevant comparison is between a collective decision-making body under sortition and a single decision-maker under elections.

[29]Issues faced by modern times policy-makers are so complex that no single person could tackle appropriately without relying on experts, as choosing the right policy nowadays involves knowledge no individual solely possesses.

[30]This would serve to counteract the potential disadvantage of sortition that arises from the possibility that the demos may not produce the best decision outcomes in complex societies as it asks citizens to make decisions on complicated and possibly multi-dimensional issues, while politicians may have built some expertise as part of their investment.

tions one would also need to account for the dynamic interaction between each of these systems and the endogenous quality of institutions that limit the size of economic rents.[31] Extending our theoretical model to a dynamic setup accounting for endogenous human capital versus political capital accumulation would further enhance our understanding of how the different democratic systems we consider here compare in terms of the expected quality of policy decisions they induce. Human capital levels could very well differ under the two systems as the incentive to accumulate human capital rather than political capital would likely be higher in a society where individuals cannot be elected. This is particularly important at an economic stage where human capital has been on a trend to becoming the main factor of production in economies across the globe.

Finally, an important extension of our baseline model would involve the addition of pandering and populist policies whose likelihood is inversely related to the level of human capital under elections. A key consideration in this case would be the interaction of human capital and outcomes under sortition versus elections, where lower levels of human capital or expertise would suggest lower access to and capacity to analyze information and would coexist with a greater degree of populism under elections.

[31] In such a dynamic setting, we could address why sortition has not emerged as a way modern societies govern themselves. While politicians and powerful organized groups have strong incentives to maintain the elections system rendering it a stable political equilibrium, such incentives might be weaker among those that govern under sortition as (unless they internalize the potential loss to society) they would have less to lose if sortition was abandoned for elections. The incentive to organise in political groups with the goal of abandoning sortition to run for elections would also undermine the stability of the sortition equilibrium. In a dynamic setting, however, the extent to which institutions become stronger over time under sortition would enhance the stability of the sortition equilibrium by making it less desireable to revert back to elections as politicians would face initially strong institutions limiting their ability to earn rents.

References

[1] Acemoglou, Daron and James, Robinson (2016) "Paths to Inclusive Political Institutions", The Pearson Institute Discussion Paper No. 39.

[2] Barber, Benjamin (1984) "Strong Democracy" Berkeley: University of California Press.

[3] Becker, Ted, Paul Szep, and Dwight Ritter (1976) "Un-Vote for a New America" Boston: Allyn and Bacon.

[4] Besley, Timothy (2004) "Paying Politicians: Theory and Evidence", Journal of the European Economic Association, 2(2-3) 193–215.

[5] Besley, Timothy (2005) "Political Selection", Journal of Economic Perspectives, 19(3) 43-60.

[6] Besley, Timothy (2006) "Principled Agents? The Political Economy of Good Government", Oxford University Press.

[7] Besley, Timothy and Marta Reynal-Querol (2017) "The Logic of Hereditary Rule: Theory and Evidence."

[8] Besley, Timothy and Mueller, Hannes (2018) "Cohesive institutions and the distribution of political rents: theory and evidence" In: Basu, Kaushik and Cordella, Tito, (eds.) Institutions, Governance and the Control of Corruption. International Economic Association. Palgrave Macmillan, Cham, Switzerland, 165-208.

[9] Boix, Carles, Michael Miller, Sebastian Rosato (2013) "A Complete Data Set of Political Regimes, 1800–2007".

[10] Bouricius, Terrill G. (2013) "Democracy Through Multi-Body Sortition: Athenian Lessons for the Modern Day," Journal of Public Deliberation. 9(1) article 11: 1-19.

[11] Burnheim, John (1985) "Is Democracy Possible? The Alternative to Electoral Politics" University of California Press.

[12] Callenbach, Ernest Callenbach and Michael Phillips (1985) "A Citizen Legislature", Banyan Tree Books and Clear Glass Publishing.

[13] Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland (2010) "Democracy and dictatorship revisited" Public Choice 143(1/2) 67-101.

[14] Ernesto Dal Bo, Pedro Dal Bo and Jason Snyder (2009) "Political Dynasties" Review of Economic Studies 76 115-142.

[15] Ernesto Dal Bo, Frederico Finan, Olle Folke, Torsten Persson nd Johanna Rickne (2017) "Who Becomes a Politician?" The Quarterly Journal of Economics, 1877-1914

[16] Frankel, Marvin (1962) "The Production Function in Allocation and Growth: A Synthesis." The American Economic Review 52 995-1022.

[17] Hayek, Friedrich (1973) "Law, Legislation and Liberty."

[18] Hennig, Brett (2017) "The End of Politicians: Time for a Real Democracy." Unbound.

[19] Manin, Bernard (1997) "The Principles of Representative Government" Cambridge University Press.

[20] Marshall, Monty G., Ted Robert, Gurr and Keith, Jaggers (2018) Polity IV Project "Political Regime Characteristics and Transitions, 1800-2017"

[21] Mattozzi Andrea and Antonio Merlo (2008) "Political careers or career politicians" Journal of Public Economics 92 597-608.

[22] Merlo, Antonio, Vincenzo Galasso, Massimiliano Landi and Andrea Mattozzi (2008) "The Labor Market of Italian Politicians, Second Version," Penn Institute for Economic Research Working Paper 09-024.

[23] Stone, Peter (2000) "The Luck of the Draw: Revisiting the Lot as a Democratic Institution" Ph.D. Dissertation, University of Rochester.

[24] Teorell, Jan, Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Valeriya Mechkova, Johannes von Römer, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-tingWang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem Codebook v9" Varieties of Democracy (V-Dem) Project.

A Appendix

A.1

Proof of Proposition 1

The proof is completed in two steps. First, given any probability of election ω^p, we show that the set $\Theta(\omega^p)$ is an interval of the real line, that is, $\Theta^p = [\underline{\theta}, \theta(\omega^p)]$. Given that, we then show that there exists an equilibrium probability ω^{p*} satisfying equation (2).

Step 1. Fix a probability of election ω^p. Observe that, given the Inada conditions (3) and (4), (1) defines implicitly a unique threshold $\theta^*(\omega^p)$ such that all types below or equal to this threshold type become politicians, with $\theta^*(\omega^p)$ being interior and increasing in ω^p. This threshold is given implicitly by the equation

$$\theta = \omega^p[R + \mu N A\widehat{\theta}] - c(\theta)$$

To see this, note that the left-hand side of the above equation is increasing in θ, while the right-hand side is decreasing in θ. The Inada conditions and the fact that the right-hand side is also increasing in ω^p implies that for any ω^p

$$\underline{\theta} < \omega^p[R + \mu N\widehat{\theta}] - c(\underline{\theta})$$

and

$$\overline{\theta} > \omega^p[R + \mu N\widehat{\theta}] - c(\overline{\theta}).$$

The intermediate-value theorem implies then the above preliminary result. Thus, a citizen of type θ is a politician if and only if $\theta \in [\underline{\theta}, \theta^*(\omega^p)]$ for the given ω^p.

Step 2. It follows that (2) becomes

$$\omega^p = \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\underline{\theta}}^{\theta^*(\omega^p)} f(\theta)d\theta)}{n+1}$$

Obviously, the left-hand side of the above equality is increasing in ω^p. Given that $\theta^*(\omega^p)$ is increasing in ω^p, we have that $\int_{\underline{\theta}}^{\theta^*(\omega^p)} f(\theta)d\theta$ is also increasing in ω^p with $\int_{\underline{\theta}}^{\theta^*(\omega^p)} f(\theta)d\theta \in (0,1)$ due $\theta^*(\omega^p)$ being interior. Therefore, from the first order stochastic dominance properties of the Binomial distribution and the fact that $\frac{1}{n+1}$ is decreasing in n, we have that the right-hand side of the above equality is decreasing

in ω^p. Note now that because $\theta^*(\frac{1}{N}) < \overline{\theta}$ we have

$$lim_{\omega^p \to \frac{1}{N}^-} \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\underline{\theta}}^{\theta^*(\omega^p)} f(\theta)d\theta)}{n+1} > \sum_{n=0}^{N-1} \frac{b(n|N-1,1)}{n+1} = 1/N$$

Similarly, because $\theta^*(1) > \underline{\theta}$ we have

$$lim_{\omega^p \to 1^+} \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\underline{\theta}}^{\theta^*(\omega^p)} f(\theta)d\theta)}{n+1} < \sum_{n=0}^{N-1} \frac{b(n|N-1,0)}{n+1} = 1$$

The intermediate-value theorem implies then directly that there is a unique probability of election ω^{p*} such that

$$\omega^{p*} = \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\underline{\theta}}^{\theta^*(\omega^{p*})} f(\theta)d\theta)}{n+1}$$

Combining the above two findings we have that the unique threshold type in the statement of the proposition is given by

$$\theta^p \equiv \theta^*(\omega^{p*})$$

which is such that

$$\theta^p = \sum_{n=0}^{N-1} \frac{b(n|N-1, \int_{\underline{\theta}}^{\theta^p} f(\theta)d\theta)}{n+1} [R + \mu N \widehat{\theta}] - c(\theta^p).$$

Decreasing Cost

First note that when $c(.)$ is decreasing $\omega^p(R + \mu N\widehat{\theta}) - c(\theta)$ is increasing in θ with slope $-c'(\theta) > 0$. For the equation

$$\theta = \omega^p(R + \mu N\widehat{\theta}) - c(\theta)$$

to accept, for the given ω^p, a unique solution, given the Inada conditions, we need that for all θ

$$-c'(\theta) < 1.$$

If the above inequality is satisfied then, given the Inada conditions, the above equation defines implicitly a unique threshold $\theta^*(\omega^p)$ such that all types below or equal to this threshold type become politicians, with $\theta^*(\omega^p)$ being interior and increasing

in ω^p. It follows that the second part of the proof above is still valid and hence the desired result follows. Moreover, the rest of the results in the main text hold with the modification that when $R < \overline{\theta}$, a *sufficient* condition for inequality (8) to hold and hence for sortition to dominate election is now $\lim_{\theta \to \bar{\theta}} c(\theta) > \mu N \hat{\theta}$, which means, once again, that the entry cost is always higher than the additional funds that the elected politician peculates.

No Rent-Seeking under Sortition

Assume that, under any of the two democratic regimes we study here, citizens choose between being an entrepreneur or a candidate for public office without having invested in rent-seeking or a candidate for public office having invested in rent-seeking. Under any regime, a citizen who decides to become an entrepreneur earns $\theta(1-\tau)$, where τ is the exogenously given income tax rate.

Start with elections. Let π denote the perks earned by a politician under elections and R the legal benefits/perks from being elected in addition to perks π. Let $\kappa(\theta)$ be the cost of building support to rise through the ranks onto the public platform, necessary to have a chance of getting elected, and $c^p(\theta)$ the cost of ensuring the politician has the ability to appropriate resources if elected. In terms of the notation in the main text, we have that $c(\theta) = \kappa(\theta) + c^p(\theta) - \pi$.

A citizen who invests only in building support to rise through the ranks onto the public platform and not towards ensuring their ability to appropriate resources if elected, earns payoff $\omega^p R + \pi - \kappa(\theta)$. A citizen who invests both in building support to rise through the ranks onto the public platform and towards ensuring their ability to appropriate resources if elected, earns payoff $\omega^p[R + \mu\overline{y}] + \pi - \kappa(\theta) - c^p(\theta)$.

Assume that

$$\frac{1}{N}\mu N\underline{\theta} = \mu\underline{\theta} > \max_{\theta \in \Theta} c^p(\theta)$$

This implies that for any $\omega^p \in [1/N, 1]$ and any realization of per capita income $\overline{y}/N \in [\underline{\theta}, \overline{\theta}]$, the payoff of any citizen from investing in rent-seeking is higher than that of simply becoming a politician with no access to resource appropriation. Therefore, no citizen will ever choose to become a politician without investing in rent-seeking. Using that $c(\theta) = \kappa(\theta) + c^p(\theta) - \pi$, we thus have that Proposition 1 stays the same.

Turning to sortition, let $\pi - \Delta$ denote the perks earned by a demos member, where $0 < \Delta \leq \pi$,[32] and R denote the legal benefits/perks from being selected in

[32] $\Delta > 0$ can be justified under the scenario that demos-members just put their names down (in a lot) and thus do not get the same public exposure with politicians. In fact, under anonymous declaration of membership they get zero perks, i.e. $\Delta = \pi$.

addition to perks $\pi - \Delta$. A citizen who decides to be part of the demos without investing towards ensuring their ability to appropriate resources if selected, earns payoff $\omega^d R + (1 - \omega^d)\theta + \pi - \Delta$. A citizen who decides to be part of the demos and invests towards ensuring their ability to appropriate resources if selected, earns payoff $\omega^p[R + \mu\overline{y}] + \pi - \Delta - c^p(\theta)$.[33]

Assume that

$$\underline{\theta}(1-\tau) > R + \mu N\overline{\theta} + \pi - \Delta - \min_{\theta\in\Theta} c^p(\theta)$$

This implies that for any $\omega^p \in [1/N, 1]$ and any realization of per capita income $\overline{y}/N \in [\underline{\theta}, \overline{\theta}]$, the payoff to any citizen from investing in rent-seeking is lower than that of simply becoming an entrepreneur. Therefore, no demos-member will ever choose to invest in rent-seeking.

Two observations are in order here. First the Inada conditions under elections are

$$\underline{\theta}(1-\tau) < \frac{[R + \mu N\underline{\theta}]}{N} + \pi - \min_{\theta\in\Theta}\{c^p(\theta) + \kappa(\theta)\}$$

and

$$\overline{\theta}(1-\tau) > R + \mu N\overline{\theta} + \pi - \min_{\theta\in\Theta}\{c^p(\theta) + \kappa(\theta)\}$$

Therefore, the above assumption is consistent with the first Inada condition if

$$\frac{[R + \mu N\overline{\theta}][N - \frac{\theta}{\overline{\theta}}]}{N} < \Delta - [\min_{\theta\in\Theta}\{c^p(\theta) + \kappa(\theta)\} - \min_{\theta\in\Theta} c^p(\theta)]$$

Second, the equilibrium condition under sortition becomes

$$\Theta^d = \{\theta | \theta \leq \frac{R + [\pi - \Delta]/\omega^d}{1-\tau}\}$$

If $\Delta = \pi$ then Proposition 2 stays the same, while if $\Delta < \pi$ then Proposition 2 needs to be adjusted along the lines of Proposition 1. However, qualitatively, the main result that there is a sufficiently high level of R that makes sortition better does not change.

[33]While we have assumed that the strengh of institutions and thus the fraction of resources that can be appropriated, μ, are identical under elections and under sortition, one could argue that in a dynamic setting institutions would tend to become stronger under sortition and thus the fraction of resources that may be appropriated would be lower. That is, given that sortition resets the political environment each time, there would be less inherent strife with institutions under sortition as compared to elections and this would result in the strengthening of institutions over time. This would only make the inequality below more likely to hold.

A.2 (Not Intended for Publication) Data

All of the political variables below have been obtained from the Varieties of Democracy (V-Dem) Project, v9. We thank Antonis Ellinas for bringing these data to our attention.

Electoral democracy index (D) (v2x_polyarchy) from V-Dem Codebook section 2.1.1. Question: To what extent is the ideal of electoral democracy in its fullest sense achieved? Clarification: The electoral principle of democracy seeks to embody the core value of making rulers responsive to citizens, achieved through electoral competition for the electorate's approval under circumstances when suffrage is extensive; political and civil society organizations can operate freely; elections are clean and not marred by fraud or systematic irregularities; and elections affect the composition of the chief executive of the country. In between elections, there is freedom of expression and an independent media capable of presenting alternative views on matters of political relevance. In the V-Dem conceptual scheme, electoral democracy is understood as an essential element of any other conception of representative democracy — liberal, participatory, deliberative, egalitarian, or some other. Scale: Interval, from low to high (0-1). Source(s): v2x_freexp_altinf v2x_frassoc_thick v2x_suffr v2xel_frefair v2x_elecoff Data release: 1-9. Release 1-5 used a different, preliminary aggregation formula. Aggregation: The index is formed by taking the average of, on the one hand, the weighted average of the indices measuring freedom of association thick (v2x_frassoc_thick), clean elections (v2xel_frefair), freedom of expression (v2x_freexp_altinf), elected officials (v2x_elecoff), and suffrage (v2x_suffr) and, on the other, the five-way multiplicative interaction between those indices. This is half way between a straight average and strict multiplication, meaning the average of the two. It is thus a compromise between the two most well known aggregation formulas in the literature, both allowing partial "compensation" in one sub-component for lack of polyarchy in the others, but also punishing countries not strong in one subcomponent according to the "weakest link" argument. The aggregation is done at the level of Dahl's subcomponents with the one exception of the non-electoral component. The index is aggregated using this formula: v2x_polyarchy = .5 MPI + .5 API = .5 (v2x_elecoff v2xel_frefair v2x_frassoc_thick v2x_suffr v2x_freexp_altinf) + .5 ((1/8) v2x_elecoff + (1/4) v2xel_frefair + (1/4) v2x_frassoc_thick + (1/8) v2x_suffr + (1/4) v2x_freexp_altinf) Citation: Teorell et al. (2019); V-Dem Codebook. Years: 1789-2018.

Institutionalized democracy (E) (e_democ) from V-Dem Codebook section 7.8.2. Question: Is the polity an institutionalized democracy? Clarification: Democracy is conceived as three essential, interdependent elements. One is the presence of in-

stitutions and procedures through which citizens can express effective preferences about alternative policies and leaders. Second is the existence of institutionalized constraints 319 on the exercise of power by the executive. Third is the guarantee of civil liberties to all citizens in their daily lives and in acts of political participation. The Democracy indicator is an additive eleven-point scale (0-10). The operational indicator of democracy is derived from codings of the competitiveness of political participation (PARCOMP), the openness and competitiveness of executive recruitment (XROPEN and XRCOMP), and constraints on the chief executive (XCONST). Responses: Numeric. Source(s): Polity IV (Marshall and Jaggers 2016). Data release: 5-9. Citation: Polity IV (Marshall and Jaggers 2016). Years: 1800-2017.

Democracy (BMR) (E) (e_boix_regime) from V-Dem Codebook section 7.2.1. Question: Is a country democratic? Clarification: Dichotomous democracy measure based on contestation and participation. Countries coded democratic have (1) political leaders that are chosen through free and fair elections and (2) a minimal level of suffrage. Responses: 0: No 1: Yes Source(s): Boix et al. (2013). Data release: 5-9. Citation: Boix et al. (2013). Years: 1800-2010.

Democracy (E) (e_chga_demo) from V-Dem Codebook section 7.9.2. Clarification: A regime is considered a democracy if the executive and the legislature is directly or indirectly elected by popular vote, multiple parties are allowed, there is de facto existence of multiple parties outside of regime front, there are multiple parties within the legislature, and there has been no consolidation of incumbent advantage (e.g. unconstitutional closing of the lower house or extension of incumbent's term by postponing of subsequent elections). Transition years are coded as the regime that emerges in that year. Responses: 1: Democracy. 0: Otherwise. Scale: Dichotomous. Source(s): Cheibub et al. (2010). Data release: 5-9. Citation: Cheibub et al. (2010). Years: 1946-2008.

A.3 (Not Intended for Publication)

Proof of $Pr(\sigma \mid x = \sigma) \geq Pr(\sigma \mid x \neq \sigma), \forall\theta \Rightarrow t(\theta) \geq \max\{\rho, 1-\rho\}$

Applying Bayes' rule we have that $\forall\theta$

$$Pr(\sigma \mid x = \sigma) \geq Pr(\sigma \mid x \neq \sigma) \Leftrightarrow \frac{Pr(x = \sigma \mid \sigma)}{Pr(x = \sigma)} \geq \frac{Pr(x \neq \sigma \mid \sigma)}{Pr(x \neq \sigma)}.$$

If we let $\sigma = 1$ and use the definitions of $t(\theta)$ and ρ, then we have

$$\frac{t(\theta)}{\rho} \geq \frac{1 - t(\theta)}{1 - \rho} \Leftrightarrow t(\theta) \geq \rho \tag{A}$$

Similarly, if we let $\sigma = 0$, then

$$\frac{t(\theta)}{1-\rho} \geq \frac{1-t(\theta)}{\rho} \Leftrightarrow t(\theta) \geq \rho \tag{B}$$

If we combine inequalities (A) and (B), then we have that $t(\theta) \geq max\{\rho, 1-\rho\}$, $\forall \sigma \in \{0,1\}, \forall \theta$.

Proof of: If $t(\theta) \geq \max\{\frac{b_1}{b_1+b_0}, \frac{b_0}{b_1+b_0}\}$ $\forall \theta$, then a citizen should set $a = \sigma$.

A citizen with human capital θ and signal $\sigma = 1$ should set $a = \sigma = 1$, if

$$Pr(x = 1 \mid \sigma = 1)u(1,1) + Pr(x = 0 \mid \sigma = 1)u(1,0) \geq Pr(x = 1 \mid \sigma = 1)u(0,1) + Pr(x = 0 \mid \sigma = 1)u(0,0)$$

or

$$Pr(x = 1 \mid \sigma = 1)u(1,1) + [1 - Pr(x = 1 \mid \sigma = 1)]u(1,0) \geq Pr(x = 1 \mid \sigma = 1)u(0,1) + [1 - Pr(x = 1 \mid \sigma = 1)]u(0,0)$$

or, using the definition of $t(\theta)$,

$$t(\theta)\{[u(1,1) - u(0,1)] + [u(0,0) - u(1,0)]\} \geq u(0,0) - u(1,0)$$

or, using the definition of b_x,

$$t(\theta) \geq \frac{b_0}{b_1 + b_0}.$$

Following very similar steps, we can show that a citizen with human capital θ and signal $\sigma = 0$ should set $a = \sigma = 0$, if

$$t(\theta) \geq \frac{b_1}{b_1 + b_0}.$$

The last two inequalities then establish the result.

Proof of the expression regarding W^P

Under the electoral system, the expected utility from the policy chosen by the representative, i..e, for $\theta \in \Theta^p(\omega^p)$

$$
\begin{aligned}
W^P &= Pr(\sigma = 1 \cap x = 1)u(1,1) + Pr(\sigma = 1 \cap x = 0)u(1,0) + Pr(\sigma = 0 \cap x = 0)u(0,0) \\
&+Pr(\sigma = 0 \cap x = 1)u(0,1) = Pr(\sigma = 1 \mid x = 1)Pr(x = 1)u(1,1) \\
&+Pr(\sigma = 1 \mid x = 0)Pr(x = 0)u(1,0) + Pr(\sigma = 0 \mid x = 0)Pr(x = 0)u(0,0) \\
&+Pr(\sigma = 0 \mid x = 1)Pr(x = 1)u(0,1) = Pr(\sigma = 1 \mid x = 1)\rho u(1,1) \\
&+Pr(\sigma = 1 \mid x = 0)(1-\rho)u(1,0) + Pr(\sigma = 0 \mid x = 0)(1-\rho)u(0,0) \\
&+Pr(\sigma = 0 \mid x = 1)\rho u(0,1) = Pr(\sigma = 1 \mid x = 1)\rho u(1,1) \\
&+[1 - Pr(\sigma = 0 \mid x = 0)](1-\rho)u(1,0) + Pr(\sigma = 0 \mid x = 0)(1-\rho)u(0,0) \\
&+[1 - Pr(\sigma = 1 \mid x = 1)]\rho u(0,1) = (1-\rho)u(1,0) + \rho u(0,1) \\
&+Pr(\sigma = 1 \mid x = 1)\rho[u(1,1) - u(0,1)] + Pr(\sigma = 0 \mid x = 0)(1-\rho)[u(0,0) - u(1,0)] \\
&=(1-\rho)u(1,0) + \rho u(0,1) + \rho b_1 Pr(\sigma = 1 \mid x = 1) + (1-\rho)b_0 Pr(\sigma = 0 \mid x = 0).
\end{aligned}
$$