PROGRESS AND PERSPECTIVES IN THE STUDY OF POLITICAL SELECTION

Ernesto Dal Bó
Frederico Finan

Progress and Perspectives in the Study of Political Selection
Ernesto Dal Bó and Frederico Finan
NBER Working Paper No. 24783
June 2018, Revised August 2018
JEL No. D72

ABSTRACT

We provide a model of self-selection by candidates in a probabilistic voting environment to shed light on the forces shaping the quality of politicians from both the supply and demand sides of politics. The model highlights that the patterns of selection and the comparative statics of politician quality depend critically on how the costs of running for office vary for candidates with different qualities. The model offers predictions on how the quality of the political class will vary with key parameters pertaining to both the supply and demand for candidates. We use the model to frame a review of the empirical literature on political selection that has emerged in the last two decades. We contrast areas where significant progress has been made with others where important theoretical predictions remain untested or existing evidence does not allow a consensus, highlighting areas for future research.

Ernesto Dal Bó
University of California, Berkeley
Haas School of Business
545 Student Services Building #1900
Berkeley, CA 94720-1900
and NBER
dalbo@haas.berkeley.edu

Frederico Finan
Department of Economics
University of California
508-1 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
ffinan@econ.berkeley.edu

1   Introduction

The last two decades have witnessed the development of a vibrant new literature in economics and political science focused on political selection. This literature has investigated how various features of the sociopolitical environment shape the quality of the political class. This has not been an idle pursuit given two important facts. First, there is evidence that political leaders matter for policy outcomes at all levels, including the national executive (Jones and Olken 2005; Besley, Montalvo, and Reynal-Querol 2011), national legislators (Lee, Moretti, and Butler 2004; Washington 2008; Clots-Figueras 2012), and local politicians (Pande 2003; Chattopadhyay and Duflo 2004; Meyersson 2014). Second, societies appear to vary in their ability to select politicians with desirable traits like education.[1] These two facts prompt the question of what forces shape the quality of the political class. In this paper, we review what we have learned about those forces after almost two decades of research.

"Quality" is of course a loaded term. Theoretical work in political economy often distinguishes between ideological differences among politicians, and differences in "valence," understood as features that voters value positively regardless of ideology, such as charisma, competence, or integrity. We will refer to quality and valence interchangeably to indicate performance-relevant traits like competence or integrity.

A brief review of the theoretical work involving candidates with valence seems to indicate an open gap. In particular, we are missing a workhorse model of the determinants of the quality of politicians that uses a standard probabilistic voting setting to capture partisan competition (as in the textbook treatment by Persson and Tabellini (2000)), while considering candidates of varying valence making entry decisions. In this paper, we propose a simple model of political selection and then use it to guide a review of the empirical evidence on what factors tend to produce politicians of higher quality.

We consider a probabilistic voting world with candidates from two rival social groups who must decide whether to pay a cost to run for office. Candidates care about winning per se and may differ in their valence. Since valence is observable and valuable to voters, higher valence gives candidates an electoral edge. Candidates learn their valence before running, and decide whether to run under uncertainty about the realized valence of the opponent. While valence does not affect the valuation of holding office, we distinguish between cases where the costs of running do not vary with valence (arguably the case if valence is about honesty) and when they do (arguably the case if valence is a form of competence that yields a higher market income that must be forgone when running). The model distinguishes between parameters belonging to the supply side of politics, such as rewards from office and the available range of qualities among would-be candidates, and parameters belonging to the demand side, such as the size and partisanship of both groups, the distance between their preferences (an inverse of political competition), and the intensity with which voters care about valence as opposed to group allegiance.

We find that the cost structure is crucial. When the cost of running does not vary with valence, it is the relatively high quality types who run, encouraged by their better electoral

---

prospects. When the cost of running rises fast enough with valence, the pattern reverses and it is the relatively low quality types who run, as they face lower costs. As a result, some of the comparative statics on candidate quality switch signs depending on which cost structure prevails. In addition, we find that while supply-side parameters tend to affect candidate quality in the same way across both parties, changes in demand-side parameters will affect opposing parties differently. We also use the model to elucidate how the study of the drivers of quality of elected politicians differs from that of candidates. Finally, we point out ways in which the model could be extended to analyze issues of accountability and party organization.

We take the main assumptions and comparative static predictions from our theory and examine what the empirical literature has to say about them. In particular, we begin by examining two key assumptions in the model: 1) that the rewards from office are positive; and 2) that voters care about valence.

The empirical literature supports both assumptions. In addition to any official salary a politician may receive, the benefits to political office can come in many other forms. There is considerable empirical evidence that the pecuniary returns to holding office can include such things as private-sector opportunities, both during and after leaving office, as well as unofficial earnings, both legal and otherwise. In contrast, the evidence on any non-pecuniary benefits is practically non-existent. Given the difficulties in measuring non-financial motives, we are largely left to infer the existence of these benefits through model-based approaches. Also, voters do seem to care about valence, whether characterized as honesty, competence, or charisma. The literature has not yet been able to isolate the effects of any one individual trait.

We also examine three main comparative-static predictions from the model under different cost structures: 1) the impact of the benefits to office; 2) the impact of political competition; and 3) the impact of group size and cohesion.

If costs are increasing with valence and parties are not too asymmetric, the model predicts that higher expected benefits of holding office will induce higher quality individuals to run for office, all else equal. Conversely, if costs are constant, an increase in benefits will lower the quality of the candidate pool. Interestingly, the empirical evidence can go in either direction depending on the nature of the benefit. When the benefit involves an increase in official salaries, the evidence suggests that the quality of the applicant pool increases, which is consistent with candidates facing steeply increasing costs of running. But if instead the benefit involves an increase in the opportunity to engage in more rent seeking, then studies tend to show a decrease in the quality of the applicant pool, which is consistent with a constant cost structure. The theory can help to rationalize these two disparate findings if the nature of the benefits is creating selection along different dimensions of valence. On the one hand, higher wages attract more competent individuals, who are also likely to have better outside options. On the other hand, an increase in the opportunity to be corrupt attracts less honest types, and running costs are unlikely to vary with a person's level of honesty.

The second prediction of the model involves political competition. Whether and how political competition helps candidate quality again depends on the cost structure. If costs are constant competition helps quality in the strong party while hurting quality in the weak party, but if costs are increasing candidates in the weak party get better without those in

the strong party being affected. Here, the majority of the evidence shows a strong positive association between the competitiveness of a district and the quality of the elected official.

It is important to note, however, that the prediction is on quality of the candidate pool, and the only empirical evidence we have in the literature is for elected politicians. Despite its simplicity, the model has less to say about the drivers of quality for the elected politicians. When it comes to the comparative statics on the elected, changes in the model's parameters will set into action various countervailing forces that render the predictions generally ambiguous. Under some restrictions, certain predictions are possible: if costs are increasing and parties are relatively even in their electoral appeal, an increase in rewards from office will increase the quality of elected politicians. This is what the empirical literature tends to find.

The final prediction involves the effect of group size and cohesion on the quality of the candidate pool. Unfortunately, the empirical evidence has little to say on this matter. We are unaware of any studies that have examined the relationship between group size and cohesion on the quality of the candidate pool. There are, however, a few recent papers that examine the effects on quality of the elected. More research is needed before a clear pattern can emerge.

In summary, the empirical literature is in its infancy, and still learning to speak. Data and identification challenges mean we actually know very little about our politicians. We know some general characteristics of those who get elected, such as their gender, age, education, and previous occupation. But for those who are not elected, we typically only know little more than their names, and this makes it difficult to study the supply and demand factors that determine political selection. In addition, valence can mean many things, and as we have highlighted, both the theory and empirics depend on how we define it. To achieve a deeper understanding of political selection, we need to go beyond the standard proxies for quality, such as education. Access to administrative data and primary data collection will be key steps in this direction.

Identification remains a significant challenge. Nobody (who we know of) gets randomly assigned to become a politician, nor are the reasons why people become politicians typically determined at random. As a result, endogeneity and omitted variable bias are ever present concerns. However, in much the same way that randomized controlled trials (RCTs) have revolutionized empirical work in other fields, the regression discontinuity design has transformed political economy. Since Lee (2008), there have been many studies using close elections to determine important questions such as whether female leaders matter, whether financial returns to political office exist, or whether policy converges across parties. As we will document below, the empirical literature has spoken loudest on the questions that lend themselves to a regression discontinuity design approach, whether it be in the use of close elections or in the evaluation of a policy rule that is applied discontinuously but based on a continuous eligibility criterion.

However, there are well known limitations to the RDD approach. As has been noted before, this design produces treatment effects that are localized, which makes comparative analysis more difficult even among studies in the same setting. Moreover, not all questions lend themselves to such convenient variation, and many of them deserve more empirical investigation, even if it falls short of quasi-experimental standards. In these circumstances,

theory–the ultimate form of pre-analysis–acquires even higher value as a guide to interpretation.

The rest of the paper is organized as follows. The next section provides a brief review of the models with valence. This will help situate our model in the literature and highlight its contributions. We present our model in Section 3, followed by a discussion of the relevant empirical literature in Section 4. We conclude in Section 5 with a discussion of some of the key open empirical and theoretical questions on this exciting topic within political economy.

2 A brief review of models where candidates differ in valence

There are a number of papers featuring models where candidates differ in their valence. We group these into roughly five strands of research.

Downsian politics and policy platforms The earliest strand involves Downsian politics; as such, these models aimed to understand whether existing differences in valence push platforms toward or away from the median voter. Some of the earliest contributions to this literature are those by Londregan and Romer (1993), Adams (1999), Ansolabehere and Snyder (2000), Groseclose (2001), Aragones and Palfrey (2002), and Schofield (2003). More recently, Ashworth and Bueno de Mesquita (2009), Bernhardt, Câmara, and Squintani (2011), Hummel (2013), and Kamada and Sugaya (2014) have added to that literature. This strand of research tends to abstract from candidate entry.[2]

Agency and screening A second strand of the literature features models in which politicians have private information about their underlying type and can use this information to extract rents while in office. Voters use elections as a way to both select and discipline politicians. But to the extent that reelection incentives force politicians to act according to voters' preferences while in office, it also makes it harder for voters to tell apart different types of politicians just by observing their record. Papers in this line include Banks and Sundaram (1993), Fearon (1999), Besley (2004) and Smart and Sturm (2013), and although they provide insights into the effects electoral accountability may have on political selection, these models tell us less about how entry decisions may be affected differently for different types of candidates.

Agency and self-selection A third strand of research focuses more on candidate self-selection decisions and derives propositions for the quality of politicians. The emphasis is usually on how the quality of politicians responds to various elements of the sociopolitical system, such as compensation in office, opportunity costs from private life, party governance, the prevalence of the rule of law and more generally the threats and opportunities, legal or illegal, of a public life. Papers in this strand include Dal Bó and Di Tella (2003), Caselli and Morelli (2004), Messner and Polborn (2004), Dal Bó, Dal Bó, and Di Tella (2006), Poutvaara and Takalo (2007), Mattozzi and Merlo (2008), Brollo, Nannicini, Perotti, and Tabellini

---

[2]Carter and Patty (2015) present a model in the Downsian tradition where candidates may abstain from participating. However, their paper is not concerned with evaluating how candidate incentives, as shaped by valence, affect self-selection decisions and ultimately the quality of politicians.

(2013) and Dal Bó, Finan, Folke, Persson, and Rickne (2017). While concerned with valence, these papers, like those in the previous strand, tend to de-emphasize the partisan elements of electoral competition. One resulting advantage is that these models help us to think about several aspects that may be orthogonal to partisanship. The disadvantage is that they remain limited in their ability to connect political selection to the more distinct aspects of electoral competition, such as group conflict and partisan differences.

Citizen-candidates Some of the papers in the third strand adopt the citizen-candidate framework from Osborne and Slivinski (1996) and Besley and Coate (1997) but tend to abstract from the ideological elements present in those papers. Conversely, the original citizen-candidate work focused on ideology more than on valence: Osborne and Slivinski (1996) abstracted from valence considerations altogether and Besley and Coate (1997) focused on the efficiency of politician-policy combinations (which could involve valence) rather than on how fundamentals shape the quality of politicians. More recently, Di Lonardo (2017) produced a citizen-candidate model where candidates differ in valence. The main bulk of his analysis abstracts from self-selection, as candidates do not know their own valence at the time they decide to run. An extension considers privately informed candidates and examines conditions for contested and uncontested election equilibria, and the link between candidate ideology and valence.

Social groups and parties A fifth strand of the literature, like the first, distinctly incorporates the electoral dimension without the same concern for Downsian policy determination. In some of this work, candidate quality is totally random rather than shaped by self-selection decisions, as in Banerjee and Pande (2009) who study ethnicity-driven politics; other work tends to place parties instead of candidates as the key strategic actor. Models along these lines include those by Carrillo and Mariotti (2001), Galasso and Nannicini (2011), Mattozzi and Merlo (2015), Galasso and Nannicini (2017), and Besley, Folke, Persson, and Rickne (2017). Placing the political party at center stage allows for important gains in understanding the calculus of a fundamental actor of electoral democracies, but it usually comes at a cost, namely detracting attention from candidate self-selection decisions.

A review of these strands of work leaves us with a hole in the literature that seems worth filling. Much of the work on agency models incorporates self-selection by candidates with varying valence but abstracts from group and partisan dimensions. Conversely, much of the work outside of agency theory looks at partisan differences, but eschews self-selection on valence. In the next section we introduce a model that combines both aspects through a simple, textbook-style probabilistic voting model à la Persson and Tabellini (2000).

3 Valence differentials and candidate self-selection in a world with probabilistic voting: a simple model

3.1 Social fundamentals

There are two groups of voters $j = r, l$, (right and left) with respective measures $\rho \in (0, 1)$ and $1 - \rho$. Voters elect a single leader. There are two groups of politicians (i.e., parties)

representing groups r and l, each with measure 1.[3] A politician from each party is randomly recognized to be able to run, and that person must then decide whether to run or not.[4] Politicians vary not only in their group affiliation, but also in terms of valence. A politician has a valence v_j drawn from a uniform distribution in the range $[0, \bar{v}_j]$; $j = r, l$, so we allow for the two groups to have different valence ranges. A politician's type is a pair (v_j, j), $j = r, l$, and is observable to voters.

3.1.1 Voter preferences

Politicians are office motivated. Winning office confers a benefit $b \geq 1$, whereas a politician outside of office receives zero. Importantly, this model does not distinguish between legal and illegal rewards from office.

A voter i obtains a benefit from electing a leader of type (v_j, j) given by

$$\varpi v_j + I(j) \delta_i,$$

where $\varpi > 0$ tracks how much voters care about valence relative to partisan alignment, and $I(j) = 1$ when $j = r$ and is zero otherwise.

The random shock δ_i captures an idiosyncratic preference of voter i for candidate r, and it is distributed uniformly in $\left[\phi - \frac{1}{2R}, \phi + \frac{1}{2R}\right]$ among r voters and in $\left[-\phi - \frac{1}{2L}, -\phi + \frac{1}{2L}\right]$ among l voters. The parameter ϕ captures the preference distance between groups. As such, it proxies for the degree of societal division. We assume $\phi > 0$, and also $\phi - \frac{1}{2R} < 0$ and $-\phi + \frac{1}{2L} > 0$, respectively guaranteeing that some r voters could vote for an l candidate and that some l voters may support the r candidate. This implies that there will be some swing voters who could be swayed even by small valence differentials. However, since $\phi > 0$, in the absence of valence differences l voters will tend to side with the l candidate, and r voters with the r candidate. The parameters R, L track the importance of group affiliation for an individual's political preference, and can be seen to describe group cohesion, or partisanship.

In this model, group membership is a candidate characteristic analogous to ideological positioning. As in citizen-candidate models, this feature of the candidate is an ascribed characteristic that cannot be altered. Once in office, the candidate will act like a member of her group, and utility consequences for voters will follow. Differently from citizen-candidate models, we focus on purely office-motivated candidates in order to isolate the self-selection implications of valence.

3.1.2 Timing

Two candidates, one from each group, are recognized to run. Their valences are realized and each candidate learns her own privately. Then each candidate decides whether to run. If one

---

[3]The measure of politicians could be made smaller than that of voters for realism, but this is inessential, as is the fact that politicians do not vote.

[4]The assumption that a party member is randomly recognized to run reflects our abstracting from party governance, a maintained assumption in this paper, and a natural point from which to develop extensions in future work.

person runs, that person wins automatically. If both run, an election is held (with the utility shocks δ_i being drawn and valences becoming observable to voters just prior to the election). If nobody runs, parties draw candidates again. A candidate who has declined to run cannot be drawn again and collects a payoff of zero. Once there is a winner she is appointed and the payoffs are collected.

3.1.3 Decisions to run

A candidate with type (v_j, j) will want to run if and only if the expected private benefit of doing so is non-negative,

$$P(j \text{ wins}) b - c(v_j) \geq 0. \tag{1}$$

3.1.4 Two formulations for costs

We will consider a model where $c(v_j) = c$, and one where $c(v_j) = c \cdot v_j$. In both cases the parameter c only matters relative to b, so from now on we will adopt the normalization $c = 1$. For ease of exposition, we will develop the model and establish the results for the case of constant costs, and later explain how the results change under the alternative cost structure. Proofs are in the appendix and contain results for both cost structures.

3.2 Solving the model

A voter supports the r candidate if and only if

$$\delta_i \geq \varpi(v_l - v_r),$$

which implies that the vote shares of the r and l candidates are,[5]

$$s_r = \frac{1}{2} + [\rho R + (1 - \rho) L] \varpi(v_r - v_l) + [\rho R - (1 - \rho) L] \phi \tag{2}$$

$$s_l = \frac{1}{2} - [\rho R + (1 - \rho) L] \varpi(v_r - v_l) - [\rho R - (1 - \rho) L] \phi. \tag{3}$$

These shares are intuitive: the vote share of the r candidate increases with the distance between groups ϕ whenever the R group has an electoral advantage (i.e., whenever $\rho R > (1 - \rho) L$). This becomes more likely when a larger share ρ of the population is r and when these voters are more partisan due to having more concentrated idiosyncratic shocks (high R). Given $\phi > 0$, a higher R and L make group membership a stronger determinant of the

---

[5]The vote shares of the r party among the two groups, r and l, are respectively, $1 - \frac{\varpi(v_l - v_r) - \phi + \frac{1}{2R}}{\frac{1}{R}}$ and $1 - \frac{\varpi(v_l - v_r) + \phi + \frac{1}{2L}}{\frac{1}{L}}$, making the total vote share of the r candidate

$$s_r = \rho \cdot \left[\frac{1}{2} + (\varpi(v_r - v_l) + \phi) R \right] + (1 - \rho) \cdot \left[\frac{1}{2} + (\varpi(v_r - v_l) - \phi) L \right].$$

vote vis-a-vis valence.[6] When $\rho = \frac{1}{2}$ and $R = L$, the two groups are electorally symmetric and absent valence differentials both parties will obtain half of the votes. Some restrictions are necessary in order to guarantee that vote shares are interior and involve votes from both groups going to both parties.[7]

As stated above, candidate j's decision to run depends on the expression $P(j \text{ wins}) \geq \frac{c(v_j)}{b}$. The probability that candidate j wins depends on whether the other candidate $-j$ runs, and conditional on $-j$ running, on how high $-j$'s valence is:

$$P(j \text{ wins}) = P(j \text{ wins} \mid -j \text{ runs}) P(-j \text{ runs}) + 1 - P(-j \text{ runs}). \quad (4)$$

This is explained by the fact that if candidate $-j$ does not run, candidate j wins automatically, but if candidate $-j$ runs, then j needs to have a high enough valence relative to that of candidate $-j$, such that j can obtain the majority of the vote. Combining expressions (1) and (4) shows that a candidate j with type v_j will run if and only if,

$$\frac{b - c(v_j)}{b} \geq P(-j \text{ runs}) P(-j \text{ wins} \mid -j \text{ runs}). \quad (5)$$

The analysis of decisions to run depends on how each side in the last inequality behaves, which will partly depend on the cost structure.

In the case $c(v_j) = 1$, the inequality in (5) governing candidate j's decision to run becomes,

$$\frac{b-1}{b} \geq P(-j \text{ runs}) P\left(s_j < \frac{1}{2} \mid v_j\right). \quad (6)$$

This expression says that, in order to run, the *benefit-cost margin* on the left hand side has to compensate for the *risk of defeat* on the right hand side. Candidate j's valence is private information, so the probability that the opponent runs ($P(-j \text{ runs})$) does not depend on candidate j's valence; also, the probability of losing a contested election ($P\left(s_j < \frac{1}{2} \mid v_j\right)$) is decreasing in j's own valence type. Hence, candidate j's risk of defeat in the right hand side of (6) is decreasing in that candidate's own valence v_j.

In standard probabilistic voting models the uncertainty about the electoral outcome stems from shocks to the distribution of idiosyncratic preferences. Here, uncertainty in the eyes of

---

[6]Note that both ϕ, a measure of preference separation, and (R, L), which capture group cohesion, can be seen to affect competition in that they alter the margin of swing voters. That said, in our view ϕ better captures the potential for competition because, as made clear in (2)-(3), with constant costs (and holding valences fixed), an increase in ϕ will always raise the vote-margin gap in favor of the dominant party (reducing competition thus measured), while the effects of changes in (R, L) depend on which is the dominant party, and which candidate has higher valence.

[7]A sufficient condition is:

$$\bar{v}_r < \frac{1}{\varpi} min\left(\frac{1}{2L} + \phi, \frac{1}{2R} - \phi\right)$$

$$\bar{v}_l < \frac{1}{\varpi} min\left(\frac{1}{2R} + \phi, \frac{1}{2L} - \phi\right)$$

In what follows, we assume these conditions to hold.

one candidate (herself endowed with private information on her own type) stems from not knowing the realized type of her potential rival.

3.3 Equilibrium quality of candidates

The pattern of selection for candidate j will depend on the number of times the left hand side of (6) intersects the right hand side as v_j varies, and whether in the case of a unique intersection, the benefit-cost margin intersects the defeat risk function from above or below. The assumptions of uniform density for idiosyncratic shocks and valence types ensure that the risk of defeat is linearly decreasing in a candidate's own type. Therefore, if an intersection exists it is unique. With constant costs, the left hand side is cut by the decreasing risk of defeat from above, so it is the candidates with relatively high valences who prefer to run. This is detailed in the following proposition:

Proposition 1 *Suppose $c(v_j) = 1$. In equilibrium, it is the relatively high valence politicians who run. Specifically, politicians run if and only if their type is above a cutoff given by*

$$v_l^* = \max\left\{0, \min\left\{\bar{v}_r\frac{1}{b} + x, \bar{v}_l\right\}\right\}$$

$$v_r^* = \max\left\{0, \min\left\{\bar{v}_l\frac{1}{b} - x, \bar{v}_r\right\}\right\},$$

where $x \equiv \frac{\rho R - (1-\rho)L}{\rho R + (1-\rho)L}\frac{\phi}{\varpi}$.

Proof: See Appendix.

This proposition goes against the usual expectation in the literature that the less able have a comparative advantage in entering politics (cf. Caselli and Morelli (2004); as we will see later, the opposite result will hold when costs are increasing fast enough.) The reason is simple: higher valence candidates are more confident of victory and hence more willing to sink the cost of running. Inspection reveals that (interior) equilibrium cutoffs are given by two terms neatly separating "supply side" parameters $\bar{v}_j\frac{1}{b}$ and "demand side" parameters grouped in the electoral asymmetry expression x. As is clear from the cutoffs, the strategies for candidates on each side are identical if the electoral asymmetry term x goes to zero. As one groups grows stronger, a wedge appears between the quality of candidates running on each side. In what follows we assume without loss of generality that $\rho R - (1-\rho)L \geq 0$ so $x \geq 0$ and we can refer to r as the electorally advantaged party. Inspection of the equilibrium cutoffs shows that,

Remark 1 *Suppose $\bar{v}_r = \bar{v}_l = \bar{v}$, so that any asymmetries stem from electoral advantage. If costs are constant ($c(v_j) = 1$), then $v_l^* > 0$ and $v_l^* \geq v_r^*$, and the disadvantaged (l) candidates have higher average quality.*

Sources of electoral strength such as a large group size (high ρ), highly partisan voters (high R), or a lack of electoral contestability (high ϕ) will make candidates in the dominant

r party more confident that they can win and incentivize lower quality candidates to run. As a result, the electorally strong party will feature worse candidates on average and increases in electoral strength will amplify that valence gap.

3.4 Comparative statics on the quality of candidates

The relevant parameters in this model are: $b, \rho, R, L, \phi, \varpi, \bar{v}_r, \bar{v}_l$. While the model can be studied keeping $\bar{v}_r$ and $\bar{v}_l$ different, we will for simplicity study the model under the symmetric restriction $\bar{v}_r = \bar{v}_l = \bar{v}$. Under both cost structures, if valences are drawn from the same distribution ($\bar{v}_r = \bar{v}_l = \bar{v}$), only the demand side term x drives a wedge between the cutoffs governing running decisions in each party.

Before formally stating the comparative static effects, note that when costs are constant the expected quality of candidates for party j is $\frac{v_j^* + \bar{v}}{2}$. Thus, differentiation of cutoffs directly yields the sign of changes in expected quality of candidates in any party j, as detailed in the following proposition (the proof if standard and hence omitted).

Proposition 2 *If costs are constant ($c\left(v_j\right) = 1$),*
    (a) The expected quality of candidates from both parties decreases with the rewards from office b, and it increases with the maximum possible type $\bar{v}$.
    (b) The expected quality of candidates from the electorally strong (weak) party decreases (increases) in the electoral asymmetry term x.

One take-away from this proposition is that the "supply side" parameters $(b, \bar{v})$ affect both parties in the same way. In contrast, changes in demand side parameters have opposing effects across parties. This is because they directly shift the wedge x that drives apart the strategies of candidates on each side.

The last proposition also provides some specific lessons. If we unbundle the electoral asymmetry term x, it is easy to see that,

Corollary 1 *If costs are constant ($c\left(v_j\right) = 1$),*
    (i) The expected quality of candidates from a party j decreases with the share of voters belonging to group j, and with the importance of own group affiliation. That is, the quality of r candidates decreases with ρ and R, and it increases in $1 - \rho$ and L, while the opposite holds for l candidates.
    (ii) The expected quality of candidates from the electorally strong (weak) party increases (decreases) with the degree of electoral contestability $\frac{1}{\phi}$, and with voters' concern for valence ϖ.

All of these results are intuitive. When costs are *constant*, it is the relatively high valence types who run, and more confidence in victory translates into lower valence types deciding to take the campaign gamble. Therefore, any increase in a party's electoral strength – whether from group size or own-group partisanship – translates into worse candidates, while the rival, weakening party selects better candidates. The parameters for voter preference distance ϕ and voter concerns with valence ϖ operate together, in the sense that only the ratio $\frac{\phi}{\varpi}$

matters, and its role is to amplify the effects of electoral asymmetry. The reason is intuitive. If ϕ is high, voters from each group are more distinct in their preferences, there are fewer swing voters, and the effects of electoral asymmetry (as measured by $\frac{\rho R - (1-\rho)L}{\rho R + (1-\rho)L}$) make more of a difference. Conversely, if concerns for valence ϖ increase, the pure partisan aspect of preferences matters less.

A conclusion from the analysis in this section is that the empirical study of some parameters – especially those on the demand side in the term x – may have to pay special attention to heterogeneous effects across parties. Empirical designs that do not accommodate for possibly heterogeneous effects may find attenuated or even zero effects.

3.5 A different cost structure

An important goal of our theoretical analysis is to contrast two different cost structures. We developed the model and main results relying on a constant cost of running. This might make sense if a higher valence captures something like integrity, which may be unrelated to the cost of running. But if valence captures something like ability and market opportunities that are forgone when running, then the costs of running could be increasing. In this subsection we detail how the previous results change under increasing costs.

3.5.1 Equilibrium selection patterns

If costs are increasing, the benefit-cost margin on the left hand side of (6) becomes decreasing in v_j at rate $\frac{1}{b}$. In addition, if that rate of decrease is faster than that of the right hand side, we could expect a pattern of selection that reverses the one under constant costs: high valence individuals have a small risk of defeat but also small benefit-cost margins due to their high costs, and therefore might prefer to stay out. Indeed, the alternative version of Proposition 1 states that if $\bar{v}_l, \bar{v}_r > b$, and $0 \leq x \leq b(\bar{v}_r - b)/\bar{v}_r$, then when $c(v_j) = v_j$ there exists an equilibrium in which candidates run if and only if their type is below a cutoff level given by

$$v_r^* \equiv b$$

$$v_l^* \equiv b\left(1 - \frac{x}{\bar{v}_r - b}\right),$$

where again $x \equiv \frac{\rho R - (1-\rho)L}{\rho R + (1-\rho)L}\frac{\phi}{\varpi}$.

As a result, if costs rise fast enough, the selection pattern flips relative to the constant cost case: it is the lower valence types who run.[8] Moreover, when costs are increasing, $v_r^* > v_l^*$ and the electorally advantaged (r) candidates have higher average quality. Why does it take steeply increasing costs to invert the selection pattern? The reason is that higher valence candidates face higher costs, but they are also more confident in victory; to invert the selection pattern, the discouragement from higher costs must dominate the encouragement from increased confidence in victory. In addition, since attracting higher quality candidates

---

[8]By fast enough we mean the slope $-\frac{1}{b}$ of the benefit-cost ratio in the left hand side of (5) is smaller than that of the defeat risk expression on the right hand side. The latter is $-\frac{1}{\bar{v}_{-j}}$.

11

requires additional confidence in victory, it is the electorally dominant party that fields the better candidates.

3.5.2 Comparative statics of candidate quality

When costs are increasing, the expected quality of candidates in party j is $\frac{v_j^*}{2}$ and differentiation yields the comparative statics effects. The alternative version of Proposition 2 (the proof of which is again standard and hence omitted) states that (i) if $c(v_j) = v_j$ and the electoral asymmetry term x is not too large, the expected quality of candidates from both parties increases with the rewards from office b, and it increases among the weak party with the maximum possible type $\bar{v}_r$. In addition, (ii) an increase in the electoral asymmetry term x leaves the strong party unaffected and decreases the quality of candidates from the weak party.

As with constant costs, the supply-side parameter b affects both parties in the same way (if asymmetry is not large), rather than making the candidates of one party better and the candidates from the other worse. But the sign of the comparative static effect for changes in rewards b changes: while in the constant cost case increases in rewards decreased candidate quality, now an improvement in the rewards from office *improves* candidate quality. The intuition is tightly linked to the drivers of the selection pattern, indicating that attracting higher quality candidates would require compensating them for the costs of running - stronger confidence in victory is not enough, and higher rewards are necessary.

When we consider the results for both cost structures together, two broad implications follow. First, the two cost structures imply opposite selection patterns and often different comparative statics results. Second, the empirical study of effects of changing rewards b on candidate quality can help pin down the prevailing cost structure independent of party affiliation: b is a parameter that affects all parties equally, and does so in opposite ways strictly based on the cost structure.

3.6 Comparative statics on the quality of the elected

It is generally easier to collect information about elected officials than about the entire candidate pool. Moreover, we tend to care more about those who are elected, since they set policy. Unfortunately, the study of what drives the quality of elected officials is theoretically much more complicated. Our model helps explain the complications.

Changes in the parameters of the model set off countervailing forces that tend to create ambiguous effects over the expected quality of the elected. There are three types of competitive scenarios: One where the r candidate runs alone, one where the l candidate runs alone, and one where both candidates run. In the last case, there is an actual electoral contest. The overall expected quality of the elected person reflects the expected quality of the winner in

each competitive scenario, weighted by the probability that each scenario occurs,[9]

$$E_{elected}(v) = E(v|r \text{ alone}) P(r \text{ alone}) + E(v|l \text{ alone}) P(l \text{ alone}) + E(v|\text{both}) P(\text{both}).$$

Therefore, given a change in a parameter θ, the change in expected quality is $\frac{dE_{elected}(v)}{d\theta} = A + B$ where,

$$A \equiv E(v|r \text{ alone}) \frac{dP(r \text{ alone})}{d\theta} + E(v|l \text{ alone}) \frac{dP(l \text{ alone})}{d\theta} + E(v|\text{both}) \frac{dP(\text{both})}{d\theta} \quad (7)$$

$$B \equiv \frac{dE(v|r \text{ alone})}{d\theta} P(r \text{ alone}) + \frac{dE(v|l \text{ alone})}{d\theta} P(l \text{ alone}) + \frac{dE(v|\text{both})}{d\theta} P(\text{both}).$$

One obvious source of ambiguity is that demand side parameters can have opposing effects on the candidates from the two parties. But the complexities do not end there. Even if we focus on the supply side parameter b, which affects the quality of candidates from both parties in the same fashion, countervailing forces are still present.

The expression for $\frac{dE_{elected}(v)}{d\theta}$ helps show that when a parameter changes there are three types of forces at play. First, there is the direct effect on the quality of candidates conditional on each competitive scenario (reflected in the terms labeled with B). Thus, a parametric change that is expected to, say, increase cutoffs v_r^* and v_l^* will introduce a force toward higher quality elected politicians because the candidates who run are better. We call this the *"candidate quality"* effect. But secondly, when a parameter changes, the probability of each competitive scenario changes as well (as reflected in the terms labeled with A). The terms in A form a convex combination of three expected qualities, and a change in cutoffs that makes more types run will reduce the chances of the uncontested electoral scenarios and increase the chance of an electoral contest (formally, if a change in a parameter θ increased cutoffs, it is easy to show that we would have $\frac{dP(\text{both})}{d\theta} > 0$). This change in the likelihood of the different competitive scenarios introduces two additional effects. First, since a contested election creates an order statistic effect over the expected quality of the winner, the contested election tends to produce winners of higher quality than the uncontested elections. This is the *"competitive"* effect. But when parties are electorally asymmetric, one of them will field candidates of higher quality than the other. Thus, an uncontested election featuring the party with better candidates will typically produce winners of high quality; the contested election gives chances to both parties, including one that may on occasion win despite having candidates of worse quality. This is a *"composition"* effect. The "competitive" and the "composition" forces go in opposite directions and render the overall impact of contestability itself ambiguous. That said, as parties get close to electoral symmetry, the composition effect vanishes, and increases in contestability produce only the competitive effect which in turn

---

[9]The probability $P(j \text{ alone})$ that party j will have its candidate running unopposed is $\frac{P(j)(1-P(-j))}{1-(1-P(j))(1-P(-j))}$ where $P(j)$ is the probability that a candidate from party j who has been recognized decides to run. Recall that if neither candidate runs, parties draw new candidates –a process that could go on for arbitrarily many iterations. The term $P(j \text{ alone})$ can be seen as the probability that party j will field the lone candidate, starting at any iteration. Thus, $P(j \text{ alone}) = P(j)(1 - P(-j)) + (1 - P(j))(1 - P(-j)) P(j \text{ alone})$, which yields the expression above. Analogous reasoning yields $P(\text{both}) = \frac{P(j)P(-j)}{1-(1-P(j))(1-P(-j))}$.

enhances the quality of the elected. When parties are close enough to symmetric, parameter changes that increase the measure of types who run make the overall effect from the terms in A unambiguously positive. The question then is whether such positive effect compounds or cancels the candidate quality effect in line B.

These considerations underlie the following proposition.

Proposition 3 *Suppose parties are close to, but not perfectly, symmetric. (i) If costs are constant ($c(v_j) = 1$), an increase in b produces countervailing forces: the candidate quality effect is negative and the competitive effect is positive. As x approaches zero the former effect dominates the latter, decreasing the quality of elected politicians. (ii) If costs are increasing ($c(v_j) = v_j$), an increase in b produces a positive candidate effect and a positive competitive effect, increasing the quality of elected politicians.*

Proof: see Appendix.

This proposition tells us that, due to potential ambiguities that only disappear under symmetry, the study of the quality of *elected* politicians cannot in general be used to determine the effects of rewards, nor what cost structure is in place. It is more reliable to analyze the quality of *candidates*.

The analysis of the comparative statics of the quality of the elected under conditions of asymmetry is difficult and left for future research. In situations of very stark asymmetry where one party is almost always running unopposed, the comparative statics of the elected will tend to track the comparative statics of the candidates from the dominant party. It is worth mentioning that the model presented here readily accounts for uncontested elections, since it is often the case that a candidate runs unopposed and it is feasible to study the comparative statics of the likelihood of such events. This is also left for future research since it is not something that we will use to guide our empirical review.

3.7   Summary and discussion

Our simple textbook-like model produces clear predictions for how fundamentals will affect the quality of candidates. A higher valence is an electoral advantage (due to higher chance of besting the opponent), but may also be a cost disadvantage (due to higher opportunity costs). When the electoral advantage force dominates, the best types run and electoral strength, lower political competition, lower voter preferences for valence, and the attractiveness of office all worsen the average quality of candidates. When the cost disadvantage force dominates, the opposite holds.

The study of the quality of the elected is more complicated. Changes in parameters can create countervailing effects through a direct impact on the quality of candidates, and through changes in the contestability of elections. Under conditions close enough to electoral symmetry, higher rewards from office are guaranteed to improve the quality of the elected when the costs of running increase with valence fast enough. Part of the proof of our Proposition 3 shows that in contested elections the effects of rewards on the elected track the effects over candidates. Thus, we should expect higher rewards from office to lower the quality of

the elected when costs are constant, and to improve their quality when costs are (rapidly) increasing.

Our simple model generates enough complexity to entertain the analyst, but it certainly leaves out important dimensions that ought to be developed in future research. We have not dealt with the issue of accountability. The model could be extended to incorporate different benefits from office depending on how candidate type affects performance and reelection probabilities. Similarly, we have not dealt exhaustively with parties. For example, parties could play a role by screening members and therefore shaping the distribution of candidate valences to differ from that of the general population.

4   A review of the empirical evidence

Our model's predictions are based on two assumptions that seem natural yet we have not justified: 1) the returns to office are positive; and 2) voters care about valence. Before we test the model's main predictions, we explore whether the empirical evidence supports these two assumptions.

4.1   What are the returns to holding office?

The empirical evidence on the returns to office can be usefully classified along two dimensions: financial versus non-financial, and present versus future returns.

Financial returns   Most elected politicians receive a salary, and for some positions the pay can be quite attractive.[10] But at the same time, we have also seen several instances in which individuals take pay cuts in order to hold office. For instance, a U.S. senator earns $174,000 annually. Prior to running for office, Georgia Senator David Perdue received an annual salary of close to $1 million as the CEO for Dollar General; Carly Fiorina's annual salary at Hewlett-Packard, before she was fired and began an (unsuccessful) Senate bid, ranged from $1 to $4 million per year. Based on these casual observations, it seems unlikely that the salary one receives in office, while a clear enticement for many, can fully explain why individuals run for office. Are there other financial returns to holding office? In this respect, the empirical literature has largely focused on trying to identify two additional sources: 1) future returns to holding office; and 2) illicit returns to holding office.

Estimating the future returns to holding office: Since leaving the White House in 1996, Bill Clinton has earned in excess of 106 million dollars in speaking fees. George W. Bush allegedly received a 7 million dollar advance for his post-presidency memoir, "Decision Points." Barack Obama's expected advance for his memoir has been estimated to be somewhere between 25-45 million dollars. Most people cite these events as clear evidence of the future benefits of holding office. After all, when individuals have been elected to a position of power and have

---

[10]For example, Peichl, Pestel, and Siegloch (2013) estimate a positive wage gap of 35%-65% between German members of parliament and citizens with similar observable traits occupying executive positions in the private sector, with the difference shrinking to zero for top positions.

experienced how policy decisions are made, they should be able to monetize this experience and knowledge once they leave office. And yet, there are relatively few studies that can provide systematic evidence on the future returns to holding office.

There are two primary reasons this strand of the literature remains sparse. First is a data constraint: it is very difficult to obtain information about a politician's income or net wealth after they have left office. We only know how much Bill Clinton received in speaking fees because the Clintons have remained in the public eye, and even then speaking fees are not his only source of income. To estimate the returns to office, one would need to find a systematic way to collect income or wealth data for a reasonably-sized sample of politicians. This has proven to be extremely difficult.

The second hurdle is in finding an appropriate counterfactual. Even if we do know that Bill Clinton earned 106 million dollars in speaking fees since leaving office, do we know how much he would have earned during this period had he not been president? The answer is certainly no. Fortunately, some recent advances have been made in our ability to properly estimate the counterfactual.

The studies that have made progress on this question have overcome these limitations in fairly creative ways. For instance, Eggers and Hainmueller (2009) created a database of every candidate who ran for the House of Commons between 1950 and 1970. For competitive candidates who had passed away (23% of the sample), they were able to link the majority to their probate values to measure the total value of the candidate's estate at death. To identify the returns to holding office, they employ a regression discontinuity design (RDD) comparing Members of Parliament (MPs) with parliamentary candidates who narrowly lost. They find that serving in office almost doubled the wealth of Conservative MPs, but had no discernible financial benefits for Labour MPs. The authors suspect that Conservative MPs profited from office largely through lucrative outside employment they acquired as a result of their political positions.

Querubin and Synder (2009) take a similar approach in their analysis of U.S. Congressmen. They use a regression discontinuity design to estimate the rents from a seat in the U.S. Congress between 1850 and 1880. They find evidence of significant returns for the first half of the 1860s, during the Civil War, but not for other periods. Those who won their first election by a narrow margin and served during the period 1861 - 1866 accumulated, on average, almost 40% more wealth between 1860 and 1870 (roughly $800,000 in present-day values) relative to those who ran but did not serve. Consistent with these findings, wealth accumulation was particularly large for congressmen who represented states most involved in military contracting and those who served during the Civil War on the committees that were most responsible for military appropriations.

We suspect that it is not a coincidence that these two studies approached the question from a historical perspective. After all, one needs microdata to identify politicians, and Census microdata in the U.S. (U.K.) only becomes public 72 (100) years after they are collected. In other settings, most notably the Scandinavian countries, researchers have been able to solve the data constraint with access to more current administrative data. For instance, Lundqvist (2013) estimates the financial returns to being elected to a local council using administrative data from Sweden. The author again uses a RDD approach, but adapted for a proportional

representation (PR) system. In this case, the comparison is between candidates who just barely won to those who barely lost a seat for the same party. For local politicians who had run for office during the period 1991-2006, the author found no evidence of financial returns in the short or long run. Whether these results are specific to the time period, to Sweden, or the type of elected position remain interesting open questions. Kotakorpi, Poutvaara, and Terviö (2016) also do not find large returns for being elected to parliament in Finland. They find that while becoming a member of parliament increases annual earnings by about $20,000 while in office, the effect on subsequent earnings is small. Another way that studies have dealt with the need for extensive data requirements is to approach the topic of returns more narrowly, from a within-office rather than career perspective; for instance, Lenz and Lim (2009) find that U.S. representatives do not accumulate wealth faster than a matched sample of non-representatives.

Overall, the results on the future returns to office seem mixed, or at least highly dependent on the context. Without additional studies, it is difficult to know whether the differences in findings are due to the nature of the position, the type of electoral system being considered, or any other difference in the political or economic setting. One takeaway, however, is clear: close elections have been a valuable source of variation for trying to identify these effects, even if they do introduce another layer of "localness" that make comparisons across studies more difficult.

Estimating the illicit returns to holding office: Politicians need not wait to leave office to capitalize on their ability to wield political influence. Rent extraction is another potential benefit of holding office.

For evidence that corruption may be an important motive for holding office, one only needs to pick up the newspaper. The concern, of course, is that newspaper accounts may paint an exaggerated picture or only report on a subset of cases. Some studies have been able to provide systematic evidence of the extent to which politicians engage in corruption. For instance, Reinikka and Svensson (2004) track the flow of public funds that the central government of Uganda used for primary education. Based on a survey of 250 primary schools that compared the amount of funds disbursed from the central government to the amount received by schools, the authors estimated a leakage rate of 87 percent. The bulk of the missing funds were captured by local officials and politicians. Ferraz and Finan (2011) use data from an anti-corruption program in Brazil that audited municipalities' use of federal funds. They estimate that approximately 80 percent of mayors engaged in some form of corruption, at a cost of US $550 million per year. McMillan and Zoido (2004) examine unusually detailed records kept by the then-secret-police chief of Peru in the 1990s documenting bribes paid in exchange for support of President Fujimori's regime. They note that politicians generally received between US $5,000 and US $20,000 per month, relative to a salary of US $4,500 per month. Niehaus and Sukhtankar (2013) examine a wage increase in the context of India's National Rural Employment Guarantee Scheme program, and find that almost the entire 21% increase was captured by local officials and politicians.

While it is has been hard to find direct evidence of political corruption, numerous papers measure corruption indirectly by estimating the value of political connections. A number of papers have documented that firms linked to politicians perform better. For instance,

Fisman (2001) shows that the stock prices of Indonesian firms connected to President Suharto fluctuated in response to news about his failing health. He estimates that 23% of the value of the most connected firms was due to corruption. Faccio (2006) later extended this analysis to include the effects of political connections to over 20,000 publicly traded firms across 47 countries. She finds that having a board member or large shareholder who is a politician is associated with a 2.3-4.3 percent increase in the company's market value, depending on the corruption level of the country that the politician represents. Szakonyi (2017) uses an RDD and data from Russia to show that politically-connected firms experience a 60% increase in revenue and a 15% increase in profits during the politician's term. He argues that the increase in firm performance can be explained by greater access to government contracts. We have also seen convincing evidence that political connections are useful for obtaining access to credit for firms in Pakistan (Khwaja and Mian (2005)) and better labor outcomes for family members in Italy (Gagliarducci and Manacorda (2016)), the Philippines (Fafchamps and Labonne (2017)), and Sweden (Folke, Persson, and Rickne (2016)). Although all of these studies focus on only one side of the exchange, it is not unreasonable to think that the politician also benefits from the relationship. The obvious limitation in interpreting the existence of corruption or valuable connections as implying substantial returns to holding office is that in most of this research the exact benefits accruing to the individual politicians themselves cannot be estimated.

The work by Fisman, Schulz, and Vig (2014) makes progress by exploiting the introduction of India's Right to Information Act that required, starting in 2003, all candidates standing for public office to disclose the value and composition of their assets. Based on this information, the authors calculate the asset growth of politicians using the disclosures of those competing in consecutive state assembly elections and compare the asset growth of election winners and runners-up. They find that the annual asset growth of winners is 3-5 percent higher than the runners-up in close elections. This sizeable increase in yearly assets cannot be explained by formal compensation, but are instead suggestive of rent-seeking activities. Consistent with this hypothesis, the growth rates are higher in more corrupt states and for individuals holding ministerial positions.

Non-financial returns   Politicians may also have non-financial motives for holding office. These incentives have been modeled in various ways, such as the self-esteem boost one gets from holding office ("ego rents"), the rewards from performing one's civic duty, and the desire to implement one's preferred policy. These incentives can motivate individuals to seek office even if financial rewards are diminished or nonexistent.

While the existence of non-financial incentives has gained acceptance in the theoretical literature, those incentives have been difficult to quantify. After all, how does one systematically measure the utility politicians receive from performing their civic duty, or implementing their policy of choice? Imposing structure through theory has allowed partial progress, by effectively "residualizing" this component of a politician's utility.

Diermeier, Keane, and Merlo (2005) provided the first estimate of the non-pecuniary returns to office using a structural model of the career decisions of U.S. national legislators. By observing politicians' career decisions and financial rewards over their life cycle, these

authors estimate that general non-pecuniary rewards amount to over $200,000 per year for a senator and about $20,000 to $35,000 per year for a representative, depending on whether he/she is a member of a major House committee. To put this number in perspective the average annual salary of a member of Congress in 1995 dollars over their sample period is $120,378.

The non-financial benefits of holding office might also accrue in the future in the form of legacy effects. Studies have shown that one of the benefits of holding office is that one's offspring are also more likely to hold office. For example, Dal Bó, Dal Bó, and Synder (2009) study political dynasties in the US Congress since the 1700s. They show political power is self-perpetuating: legislators who hold power for longer become more likely to have relatives entering Congress in the future. Querubin (2015) finds similar evidence for the Philippines. He finds that candidates who barely win their first election are around 5 times more likely to have a relative in office in the future than individuals who barely lose and do not serve.

4.2   Do voters care about valence?

Valence can mean different things to different researchers. When the term was originally introduced by Stokes (1963), valence (or "valence issues") referred to any issue or attribute the electorate cared about beyond a candidate's stated policy platform. This definition can of course include various attributes, such as a politician's charisma, integrity, or the ability to deliver goods and services. Valence can be multi-dimensional, and on some level context-specific. This imposes an obvious empirical challenge, especially given the paucity of information that typically exists about political candidates.

Measuring valence: As we will see in the studies below, most empirical studies that feature valence define it as an indicator of quality, and therefore tend to rely on common proxies, such as a person's education level, pre-office income levels, or type of occupation. And while an individual's socio-economic outcomes are undoubtedly a function of that individual's traits, these proxies are not only likely to miss important attributes but are also likely to reflect luck or social class.

This measurement concern was a primary reason why Dal Bó, Finan, Folke, Persson, and Rickne (2017) turned to administrative data from Sweden in order to assess whether politicians were positively selected from the population. In contrast to other (non-Scandinavian) settings, Sweden maintains a national registry of its population which can be accessed for research. Besides the common set of socio-economic characteristics (e.g., age, education, gender, income, etc.) these data also contain rich measures of individual ability, including evaluations of IQ and leadership potential for the 1951-1980 cohorts done by the military on all males. These data also link individuals across generations, which allows for the measurement of an individual's social background. Based on these data, the authors settled on four different measures of quality: 1) education, 2) leadership score, 3) cognitive score, and 4) a measure of an individual's earning capacity as estimated from a fully saturated Mincer regression developed by Besley, Folke, Persson, and Rickne (2017). These measures, while positively correlated, also capture different margins of variation and suggest the possibility that there might be different dimensions of ability. That multiplicity highlights the hazards

of relying on a single measure of quality.

The contrast between Figures 2a and 3 illustrates some of the pitfalls in the measurement of politician quality. The former shows a negative correlation across countries between GDP per capita and the percentage of legislators with a college degree or higher (a correlation that, as shown in Figure 2b, obtains also when considering the education of politicians relative to that of the general population they represent). If instead we were to plot this relationship using the percentage of legislators who had a master's degree or higher as our measure of quality, we would reach the exact opposite conclusion: there is a positive correlation across countries between GDP per capita and the quality of the political class.

Is there any evidence that voters care about valence? In general, there is evidence that voters react to news that might be informative about valence, although voters' reactions may not always improve selection. Characterizing valence as honesty, there is a substantial literature noting that voters are less likely to support a candidate who has behaved (or been strongly alleged to have behaved) corruptly. This behavior has been documented in elicitation of survey preferences (Winters and Weitz-Shapiro 2013; Banerjee, Green, McManus, and Pande 2014) and in actual voting (Ferraz and Finan 2008; Chong, De La O, Karlan, and Wantchekon 2014; Chang, Golden, and Hill 2010; Bobonis, Cámara Fuertes, and Schwabe 2016; Fisman, Schulz, and Vig 2016).[11]

There is also substantial evidence that corrupt behavior by politicians affects citizens' propensity to vote. Kostadinova (2009), Caillier (2010), de Figueiredo, Hidalgo, and Kasahara (2012), and Chong, De La O, Karlan, and Wantchekon (2014) all document evidence of lower turnout when voters are made aware of corruption, although Malesky, Schuler, and Tran (2012) find no effect. Overall, almost all of the findings indicate that evidence of corruption impacts citizen responses, either through lower vote shares for the corrupt, or through lower turnout, which may also suggest that voters disapprove of, and are put off by, corruption. The catch is, lower turnout does not necessarily mean tighter accountability. Throughout, it is worth noting that evidence of corruption does not only demonstrate a politician's (lack of) honesty; it may also signal something about his performance in office.

Voters also seem to prefer candidates who are more competent, although this competence has been defined empirically in a variety of ways, some more compelling than others. Besley, Folke, Persson, and Rickne (2017) cite a survey of Swedish voters in 2000 where 71 percent of respondents agreed that competence was an important attribute of politicians. Candidate quality is often defined as having previously held office, and there is considerable evidence that such candidates tend to garner higher vote shares, as in Lublin (1994).[12] In a more recent example, Hobolt and Hoyland (2011) find that voters reward parties who choose more experienced top candidates in elections to the European Parliament.

Another common way to define competence is performance in office. Studies show that voters reward top performers in a way that is consistent with positive selection (Padró i Miquel and Snyder 2006; Alt, de Mesquita, and Rose 2011; Banerjee, Kumar, Pande, and Su 2011; Butler and Powell 2014; Humphreys and Weinstein 2012).[13] Clearly, however,

---

[11]See Pande (2011) for summaries of several papers on this question.

[12]For a thorough review of this literature, see Squire (1995)

[13]Pande (2011) also summarizes some of this literature.

performance in office, while likely correlated with competence, is not a direct measure.

Two studies provide more direct evidence that competence matters to voters, both in the context of randomized experiments embedded within campaigns. Bidwell, Casey, and Glennerster (2015) show that voters respond positively to new information about candidates' professional qualifications; Kendall, Nannicini, and Trebbi (2015) show that voters respond positively to a campaign that featured the incumbent's competence as a manager. Kendall, Nannicini, and Trebbi (2015) use a random utility voting model to estimate the weights that voters place on valence versus ideological considerations. They find that while most voters place equal weights on valence and ideology, right-leaning voters tend to place much lower weight on valence. Although this paper does well to separately identify the effects of ideology from valence, one complication in the interpretation of the results is that their informational treatment about valence does not separate the incumbent's competence from effort. Thus, we might have estimated similar effects even in a pure moral hazard world: voters may have been rewarding candidates based on effort rather than screening on valence.

Voters also select on other characteristics, such as appearance (Atkinson, Enos, and Hill 2009) or ethnicity (Norris and Mattes 2003; Ferree 2006; Eifert, Miguel, and Posner 2010; Adida 2015). Ethnicity is likely a heuristic shortcut (Ferree 2006), and may even be a reasonable one, given that in some contexts ethnicity correlates with redistributive preferences (Pande 2003). Interestingly, this behavior is most pronounced in low information environments, and tends to decline as potential voters get more information about a candidate's competence (Casey 2015) or various other characteristics (Conroy-Krutz 2012). This suggests that such behaviors are not inconsistent with the existence of a preference by voters for high valence candidates.

4.3   Do higher rewards from office improve politician quality?

According to our model, it is important to distinguish between the pool of candidates and the elected politicians. Predictions about the quality of the elected combine predictions about the quality of those who run with predictions about changes in the type of competition scenarios that may occur. Theoretical predictions about the quality of the overall pool of willing candidates are more straightforward. If the costs of running are increasing (fast enough) with valence, then as the expected benefits of holding office increase, higher quality individuals will find it worthwhile to run for office, all else equal. If costs do not rise fast enough, an increase in benefits will lower valence instead. (In the case of the elected, the same predictions arise if one restricts attention to the case of increasing costs, or scenarios where contestability is held constant – something that the analyst cannot easily control for.) Given these opposite predictions, what does the empirical evidence say?

To address this question empirically, researchers have to overcome two serious but familiar challenges: data and identification. Not surprisingly, the limited empirical evidence we have on this topic has focused on the easier-to-measure financial rewards to office, and in particular the salaries politicians receive, rather than on non-financial rewards. It is also worth noting that the preferred empirical setting involves the entire candidate pool rather than just the elected. To wit, the data gods can be cruel: We often know very little when it comes to those

who choose to run, but are not elected.

Identification is, of course, the other concern. The rewards from office are almost never set exogenously. Instead, they are a function of the political and economic environment. This makes it difficult to cleanly identify the effects of, say, a politician's salary on the characteristics of the candidate pool, even assuming sufficient sacrifices to the data gods have been made.

Given these constraints, one shouldn't expect to see too many serious studies on this topic. To our knowledge, there have been very few studies that have examined the effects of politician salaries on the quality of the candidate pool. Ferraz and Finan (2008) examine whether higher salaries attract better candidates for local council in Brazil. The authors exploit the fact that salaries for local legislators are capped at different levels based on the population of the municipality. Based on this exogenous variation, they find that a 20 percent increase in wages leads to a modest increase of 0.2 years of schooling and 0.05 additional terms of experience. Gagliarducci and Nannicini (2013) examine the effects of higher wages on the pool of candidates who run for mayor in Italy. Using a similar research design, they also find that higher wages attract more educated candidates. Kotakorpi and Poutvaara (2011) take advantage of a reform in the year 2000 in Finland, when the salaries of members of parliament were increased by 35%. Using candidates in municipal elections as a control group, the authors find that a higher salary increased the fraction of candidates with higher education among female candidates, but not among male candidates. Lastly, Dal Bó, Finan, Folke, Persson, and Rickne (2017) document a positive correlation across Swedish municipalities between relative mayoral wages and several measures of quality among individuals running at the top of the various municipal party lists, who de facto constitute the set of mayoral candidates.

A few other studies have also looked at the effects of wages on political selection, but they have been limited to elected politicians. Fisman, Harmon, Kamenica, and Munk (2015) use data on Members of the European Parliament (MEPs) to examine the impact of salaries on the composition and behavior of legislators. They employ a differences-in-differences approach that exploits the introduction of a law that equalized the salaries of MEPs. They find that increasing salaries decreases the quality of elected MEPs, as measured by college quality. The intuition for this result is similar to why higher wages may decrease the quality of the candidate pool when opportunity costs do not vary with valence. The authors report that the wage increase leads those with lower quality degrees to seek reelection more, without affecting the reelection seeking behavior of those with top degrees. Braendle (2013) also studies the harmonization of the base pay for the Members of the European Parliament (MEPs), but in using alternative measures of human capital (e.g. education levels, political experience, and occupational background) he finds no significant impact of the reform on quality.

As we discussed above, the impact on entry decisions of financial and non-financial benefits beyond salary remains uncharted territory. A study by Brollo, Nannicini, Perotti, and Tabellini (2013) offers a glimpse into this area, however. These authors examine the effects of additional government revenues on political corruption and on the quality of politicians. If additional revenues yield benefits to politicians, variation in those revenues may lead to

changes in non-wage benefits. An obvious caveat to our interpretation of this study is that the link between central transfers and benefits accruing to the local politicians is likely complex and heavily dependent on context. The identification strategy of the authors is based on the fact that federal transfers to municipal governments change exogenously according to given population thresholds. Exploiting a regression discontinuity design, they show that larger transfers increase political corruption and reduce the quality of candidates for mayor, as measured by education. Pecenco and Sukhtankar (2018) document similar findings in the context of local elections in India. The authors exploit experimental variation in the introduction of a new payment technology used for two of India's largest welfare programs. This reform was estimated to have reduced leakage in both programs by 40% (Muralidharan, Niehaus, and Sukhtankar 2016). The authors show that the heads of the elected village councils in the treatment areas are less educated than those elected in the control areas. Relatedly, Gehlbach, Sonin, and Zhuravskaya (2010) document the high share of businessmen who run for public office in Russia. Using data from Russian gubernatorial elections, they show that "Businessman Candidates" emerge in regions with low media freedom and government transparency. They argue that when democratic institutions are comparatively weak, businessmen run for office to avoid the cost of lobbying elected officials. When democratic institutions are strong, the ability to influence public policy is limited. How one maps "Businessman Candidates" to valence is unclear, but the results do suggest that the opportunity to extract rents might attract candidates who are more venal.

In sum, there is fairly convincing evidence that the rewards from office matter for political selection, but how they matter appears to depend on the type of rewards in question. Increases to legal rewards tend to induce positive self-selection, whereas increases to illegal rewards can invite negative self-selection. One interpretation of this difference lies in our assumptions about opportunity costs and valence. If we view valence as being about honesty, then the costs of running should be constant in valence, and according to our model, and most of the evidence, increases in the attractiveness of office induce a negative selection effect. Still, these considerations suggest that it would be informative to enrich our model by allowing the nature of the benefits to vary.

4.4 Does political competition improve the quality of politicians?

A second main prediction of the model is that the quality of the candidate pool will depend on electoral contestability. If costs are constant, then we should observe higher quality candidates by the dominant, and hence likely more numerous, party in more competitive districts (with negative effects on the weak party); if costs are increasing we should expect no change in the strong party but higher quality among candidates in the weak party. Thus, higher quality may result on average in both cases. The empirical study of the effects of competition has met with the usual challenges of identification and data availability. With the exception of Dal Bó, Finan, Folke, Persson, and Rickne (2017) who document a positive relationship between political competition and the quality of the candidate pool, all the other empirical studies on this topic have only focused on elected politicians.

There are several strands of the literature that can claim empirical support for the pre-

diction that more competition should improve the quality of willing politicians. One line of study has looked at the relationship between political competition and performance. If we care about valence because politicians with higher valence perform better in office, then we should also expect to see better policy outcomes, broadly defined, in more competitive districts. Several studies suggest this is the case. For example, Besley, Persson, and Sturm (2010) show that among states in the U.S., higher political competition is associated with policy choices that are more pro-growth (i.e., low taxes, higher investment spending, etc.). Besley and Preston (2007) examine the performance of local government in the United Kingdom from 1980-1988. They show that policies tend to be more moderate in places with less electoral bias (and hence more competitive). Galasso and Nannicini (2011) provide evidence that members of the Italian parliament from 1994-2006 are less likely to be absent for roll-call votes if they come from more contested districts. Recently, Gavoille and Verschelde (2017) provide additional evidence of the effects of electoral competition on political selection, using data on members of the French National Assembly from 1958 to 2012. They show that a Herfindahl index of the vote shares during the first round of the legislative elections are strongly predictive of a deputy's legislative activity. The correlation between political competition and performance can even be seen at a cross-national level. Using an event-study design, Acemoglu, Naidu, Restrepo, and Robinson (2014) show that countries grow faster after becoming more democratic (which, one might expect, increases political competition).

Although the relationship between competition and performance is interesting for many reasons, there are two immediate concerns for using this line of attack in support of the theory. First, a positive relationship between competition and performance is also consistent with a model in which politicians are identical. In models that feature only moral hazard (e.g. Ferejohn (1986)), politicians have an incentive to behave "better" in places that are more competitive. Second, unless we are willing to take a normative stance about what good policies entail, it is not clear what better performance means, or even how valence should map onto these outcomes. After all, competent individuals are equally capable at engaging in corruption as they are at implementing growth-enhancing policies.

Another strand of the literature offers a more direct line of attack, by using a proxy for valence as the dependent variable. A nice example is again the paper by Galasso and Nannicini (2011). In addition to documenting that political competition leads to less absenteeism among Italian MPs, they also show that MPs who come from more competitive districts have more years of schooling, higher pre-election income, and more experience in local government. Similarly, De Paola and Scoppa (2011) use data from Italian local level governments for 1985-2008 to show that in municipalities that are more politically competitive, both mayors and councilors tend to have higher education levels and come from higher-skilled occupations.

Researchers have documented similar evidence in the U.S. In their study on political dynasties, Dal Bó, Dal Bó, and Synder (2009) find that dynastic legislators (who typically have less public office experience) are less common in more competitive places. Atkinson, Enos, and Hill (2009) find that more competitive congressional districts in the U.S. lead to contestation by candidates who also appear more competent. The authors asked college undergraduates to evaluate hundreds of pairs of photos of 972 Congressional candidates for the 2004 House elections and 1990-2006 Senate elections. For each pair, the students were

asked to pick which candidate appeared more competent. Based on these pairwise choices, the authors constructed competence scores for 972 Congressional candidates. They then regressed their measure of candidate facial competence score on district competitiveness as evaluated by the Cook Political Report, which classifies each campaign as Tossup, Lean, Likely, or Safe for each party. They find that moving from a race categorized as Safe to one that is considered a Tossup increases the facial competence of the challenger by almost one standard deviation. There is also evidence in support of the prediction at the cross-national level. Besley and Reyal-Querol (2011) use a dataset on 1,400 world leaders between 1848 and 2004 to examine whether democracies select more educated leaders than autocracies. Controlling for country and year fixed effects, they find that democracies are 20 percent more likely to select a highly educated leader.

In sum, there is broad empirical support for the second prediction. But even with the ball seemingly at the back of the net, there are at least three reasons why one should stop short of claiming victory. First, empirical work rarely analyzes heterogeneous effects by party, something that the theory strongly suggests should be done when studying the effect of demand-side parameters. If the detected effects operate across the whole political spectrum, then both versions of the model presented here would be falsified. Second, political competition as is often empirically measured (a vote-margin gap between parties) is itself an equilibrium outcome, determined not only by voters' preferences, but by candidates' best responses and the overall political environment. This raises important identification concerns. Places that are more politically competitive are also likely to be, among other things, more economically developed and populated by more educated citizens. Without accounting for these types of factors, we are likely to overstate the importance of political competition on political selection. Third, with empirical evidence so uniformly in favor of one version of the theory that makes competition look good, one cannot help but wonder how much publication bias is at play. While it is always difficult to induce sufficient excitement to publish a null result, it is even more so if the null result is based on a poor research design. Without a credible empirical strategy, such as an RDD to lean on, the literature has yet to establish convincing evidence on the relationship between political competition and candidate quality.

Although we are far from being able to causally estimate the extent to which political competition affects political selection, we have made some nice progress in identifying effects that bear some relation to competition. Beath, Christia, Egorov, and Enikolopov (2016) study the effects of electoral rules on political selection through a randomized experiment in villages in Afghanistan. As part of a community-driven development project, villages received block grants in order to fund local public goods. A precondition for receiving these block grants was that villages were required to create an elected community development council. The experiment randomized whether the councilors were elected based on an at-large system or on single-member-district elections. We interpret the latter as constraining political competition. Elected officials from the at-large elections were found to be more educated than those elected from single-member-district elections. Unfortunately, the authors do not present evidence regarding the effects on the entire candidate pool.

The literature on the impact of political reservations on political selection provides another nice example of studies relying on plausibly exogenous variation in the rules of political

representation. For instance, since the 1990s India has reserved political positions at the local level for women and for two disadvantaged minority groups – scheduled castes and scheduled tribes. If the effect of a political reservation is to increase (to 100%) the probability that a candidate from a certain group gets elected, one could construe reservations as lowering competition. Then a natural question to ask is whether political reservations induce lower quality candidates to run. An attractive feature of the setting is that the elections in which districts were subject to these reservations are determined exogenously.

Based on this variation, the evidence seems to indicate across various measures that political reservations have a negative effect on the quality of officials. In an important study by Chattopadhyay and Duflo (2004) on the effects of female reservations on public goods provision, they show that in the villages with reserved female pradhans, the leaders tended to be less educated, less experienced, have fewer household assets, and less likely to have received any formal training. Lamba and Spears (2013) provide indirect evidence of the effect of caste-based reservation on leadership quality. They show that reserved villages are less likely to receive a cash award for reducing open defecation.

However, it is not necessarily clear that the negative impacts of quotas extend outside of India or persist over time. Besley, Folke, Persson, and Rickne (2017) examine the effects of the introduction of a "zipper quota" in Sweden, which led to women and men being listed on ballots in alternating order. They found that this policy increased the earnings capacity of elected officials, driven by an increase in the quality of male elected politicians. Baltrunaite, Bello, Casarico, and Profeta (2014) find that levels of education and relevance of previous occupation among elected politicians also increase after the introduction of gender quotas in Italy. Bagues and Campa (2017) find that municipal-level quotas in Spain increase female representation without discernible effects on quality or the promotion of women to positions of higher influence.

Overall, the effect of reservations seems to indicate that the characteristics of the underlying population from which candidates emerge matter. Securing the representation of disadvantaged groups with serious shortages of human capital may mechanically reduce the quality of elected politicians in the short run. This is more likely to occur in the local context in India, where human capital asymmetries are large, and less likely to occur in Sweden where said asymmetries are smaller. These considerations suggest a potential tension between representation and competence: a society where no group has a monopoly on representation, but where instead all groups are represented, may have to make room to less educated politicians. Dal Bó, Finan, Folke, Persson, and Rickne (2017) take up the study of this tension in Sweden. Although it is true that members of the working classes have on average lower competence scores than members of economic elites, politicians from the working class are as much or even more positively selected, diluting the tension between broad representation and competence.

4.5 Does group size and cohesion improve the quality of politicians from the group?

Another prediction from our model relates to the impact of group size and cohesion on the quality of politicians. To our knowledge only two papers have explicitly studied the effect of group size and partisanship on candidate quality. Munshi and Rosenzweig (2016) examine the role of caste-based politics in India on the supply of local public goods. Using panel data from India across three elections, they show that when the population share of the largest eligible caste passes a threshold of 0.5, the schooling of the elected representative increases by almost 2 years. The authors interpret this finding in the context of a model in which each ethnic group runs a candidate who maximizes the collective benefit of the group, net of his effort and opportunity cost. Because more competent representatives deliver higher levels of public goods, it is in the group's interest to run a high quality type. To ensure that the high quality type is willing to run given that his opportunity cost is higher, the model further assumes that the group can commit to ex-post transfers. Thus, once the largest ethnic group becomes sufficiently numerous to become politically viable, we will see an increase in the quality of the elected politician.

This finding stands in contrast to the study by Banerjee and Pande (2009), also in India. They argue that in places with a dominant population group, the more citizens care about identity politics, the worse the quality of the political representation will be. The reason is that stronger concern for group identity makes the larger group's electoral advantage more unassailable. These authors derive their prediction from a model that captures one aspect of ours, namely a strictly statistical effect: suppose candidates are drawn at random from a distribution, and the probability of winning depends both on candidate valence and electoral strength of the party/group - then electorally stronger groups may win even when drawing the lower valence candidate. To test this claim, they assemble a dataset that combines electoral data from 1980 and 1996 with microdata on the corruption records of Indian politicians from the state of Uttar Pradesh. They find that when voters are more likely to vote along caste lines, winner quality is lower in jurisdictions where she or he is from the majority party relative to jurisdictions where her or his party represents the minority.

That these two intriguing studies took in place in India, in a setting where group identity is well defined, is unlikely to be a coincidence. In other contexts, the notion of group formation and group identity are likely to be much more endogenous, which makes identifying the effects of group size on political selection even more challenging.

5 Conclusion

The research on political selection has come a long way since Besley (2005) and Persson and Tabellini (2000) surveyed the literature almost two decades ago. We know a lot more about who becomes a politician and why. These insights have the potential to affect policy in ways that could improve the quality of politicians and governance more generally.

There is, however, still plenty to be learned. In particular, there are at least five major

gaps in the literature that will hopefully be addressed by the next time this literature is resurveyed, maybe two decades from now.

First, how does valence affect performance? We have a general sense that valence matters, or at least that it matters to voters. But we have a poor understanding of the ways in which it matters. Does it matter through the types of policies that are set and/or through how given policies are implemented? What are the dimensions of valence that affect performance, and how should we adjust the rewards from office to attract individuals who possess these traits? Such questions will undoubtedly require not only new sources of data, but also additional guidance from theory. If such theory is to take advantage of our simple model, it will be important to add some notion of accountability in order to better capture the incentives office holders face.

Second, political parties likely play a major role in who becomes a politician, and yet we have a very limited understanding of how political parties recruit and screen their candidates. The literature on primaries (e.g., Hirano and Snyder (2014)) is a promising step in this direction, but in the end it is only one of the many ways parties select their candidates. We need a much more comprehensive understanding of party objectives, governance structures, and strategies, and how these factors shape the quality of the political class. Models such as those by Carrillo and Mariotti (2001), Caillaud and Tirole (2002), Besley, Persson, and Sturm (2010), Galasso and Nannicini (2011), Mattozzi and Merlo (2015), Galasso and Nannicini (2017), and Besley, Folke, Persson, and Rickne (2017) provide a valuable start.

Third, while we have made progress in documenting some of the financial rewards of political office and how they affect political selection, we are still missing evidence on non-financial returns. Motives such as prestige or the desire to perform one's civic duty could play even larger roles in determining selection patterns. Different types of rewards may also hold differential interest to candidates with different types of valence. Higher wages seem effective in attracting more competent individuals who face higher opportunity costs. Attracting more honest types, however, whose opportunity costs may be unrelated to their type, may require improvements in other types of reward (or punishment) schemes. Progress – both theoretical and empirical – on this theme is needed.

Fourth, we need to better understand the relationship between political competition and valence for both the elected and non-elected. Thus far, the empirical evidence has consisted of observational studies mostly limited to the valence of the elected candidates. We still do not have clear evidence, causal or otherwise, on whether political competition improves the quality of the candidate pool as a whole.

Finally, despite some pioneering efforts, the empirical relationship between candidate quality and group size and cohesion is largely unexplored. How to deal with the endogeneity of group size or cohesion is a daunting empirical challenge that remains to be conquered. But given the importance of parochial politics and ethnic divisions in determining the way we vote and the way public resources are distributed, it is a relationship that should attract more attention in the years to come.

References

Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2014, March). Democracy Does Cause Growth. NBER Working Papers 20004, National Bureau of Economic Research, Inc.

Adams, J. (1999, July). Policy Divergence in Multicandidate Probabilistic Spatial Voting. *Public Choice 100*(1-2), 103–122.

Adida, C. (2015). Do african voters favor coethnics? evidence from a survey experiment in benin. *Journal of Experimental Political Science 2*, 1–11.

Alt, J., E. B. de Mesquita, and S. Rose (2011). Disentangling accountability and competence in elections: Evidence from u.s. term limits. *The Journal of Politics 73*(1), 171–186.

Ansolabehere, S. and J. Snyder, James M (2000, June). Valence Politics and Equilibrium in Spatial Election Models. *Public Choice 103*(3-4), 327–336.

Aragones, E. and T. R. Palfrey (2002, March). Mixed Equilibrium in a Downsian Model with a Favored Candidate. *Journal of Economic Theory 103*(1), 131–161.

Ashworth, S. and E. Bueno de Mesquita (2009, September). Elections with platform and valence competition. *Games and Economic Behavior 67*(1), 191–216.

Atkinson, M., R. Enos, and S. Hill (2009). Candidate faces and election outcomes: Is the face-vote correlation caused by candidate selection? *Quarterly Journal of Political Science 4*, 229–249.

Bagues, M. F. and P. Campa (2017). Can gender quotas in candidate lists empower women? evidence from a regression discontinuity design. IZA Discussion Papers 10888, Bonn.

Baltrunaite, A., P. Bello, A. Casarico, and P. Profeta (2014). Gender quotas and the quality of politicians. *Journal of Public Economics 118*, 62–74.

Banerjee, A., D. Green, J. McManus, and R. Pande (2014). Are poor leaders indifferent to whether elected leaders are criminal or corrupt? a vignette experiment in rural india. *Political Communication 31*(3), 391–407.

Banerjee, A., S. Kumar, R. Pande, and F. Su (2011). Do informed voters make better choices? experiment evidence from urban india. Mimeo.

Banerjee, A. and R. Pande (2009). Parochial politics: Ethnic preferences and politician corruption. Mimeo.

Banks, J. S. and R. K. Sundaram (1993). Adverse selection and moral hazard in a repeated elections model. *ch 12*, 295–311.

Beath, A., F. Christia, G. Egorov, and R. Enikolopov (2016). Electoral rules and political selection: Theory and evidence from a field experiment in afghanistan. *Review of Economic Studies 83*, 932–968.

Bernhardt, D., O. Câmara, and F. Squintani (2011). Competence and Ideology. *Review of Economic Studies 78*(2), 487–522.

Besley, T. (2004). Paying politicians. *Journal of the European Economic Association 2*, 193–215.

Besley, T. (2005). Political selection. *Journal of Economic Perspectives 19*(3), 43–60.

Besley, T. and S. Coate (1997). An economic model of representative democracy. *The Quarterly Journal of Economics 112*(1), 85–114.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from sweden. Forthcoming [American Economic Review].

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011). Do educated leaders matter?*. *The Economic Journal 121*(554), F205–227.

Besley, T., T. Persson, and D. M. Sturm (2010). Political Competition, Policy and Growth: Theory and Evidence from the US. *Review of Economic Studies 77*(4), 1329–1352.

Besley, T. and I. Preston (2007). Electoral bias and policy choice: Theory and evidence. *The Quarterly Journal of Economics 122*(4), 1473–1510.

Besley, T. and M. Reyal-Querol (2011). Do democracies select more educated leaders? *American Political Science Review 105*(3), 552–566.

Best, H. and M. Cotta (2000). *Parliamentary Representatives in Europe 1848-2000.* Oxford University Press.

Bidwell, K., K. Casey, and R. Glennerster (2015). Debates: The impact of voter knowledge initiatives in sierra leone. Mimeo.

Bobonis, G., L. Cámara Fuertes, and R. Schwabe (2016). Monitoring corruptible politicians. *American Economic Review 106*(8), 2371–2405.

Braendle, T. (2013). Does remuneration affect the discipline and the selection of politicians? evidence from pay harmonization in the european parliament. WWZ Discussion Paper, No. 2013/03.

Brollo, F., T. Nannicini, R. Perotti, and G. Tabellini (2013). The political resource curse. *American Economic Review 103*(5), 1759–1796.

Butler, D. and E. Powell (2014). Understanding the party brand: Exexperiment evidence on the role of valence. *The Journal of Politics 76*(2), 492–505.

Caillaud, B. and J. Tirole (2002). Parties as political intermediaries*. *The Quarterly Journal of Economics 117*(4), 1453–1489.

Caillier, J. (2010). Citizen trust, political corruption, and voting behavior: Connecting the dots. *Politics and Policy 38*(5), 1015–1035.

Carrillo, J. D. and T. Mariotti (2001). Electoral competition and politician turnover. *European Economic Review 45*(1), 1–25.

Carter, J. and J. W. Patty (2015). Valence and campaigns. *American Journal of Political Science 59*(4), 825–840.

Caselli, F. and M. Morelli (2004). Bad politicians. *Journal of Public Economics 88*, 759–782.

Casey, K. (2015). Crossing party lines: The effects of information on redistributive politics. Forthcoming, American Economic Review.

Chang, E., M. Golden, and S. Hill (2010). Legislative malfeasance and political accountability. *World Politics 62*(2), 177–220.

Chattopadhyay, R. and E. Duflo (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica 72*, 1409–1442.

Chong, A., A. De La O, D. Karlan, and L. Wantchekon (2014). Does corruption information inspire the fight or quash the hope? a field experiment in mexico on voter turnout, choice and party identification. *The Journal of Politics 77*(1), 55–71.

Clots-Figueras, I. (2012, January). Are female leaders good for education? evidence from india. *American Economic Journal: Applied Economics 4*(1), 212–44.

Conroy-Krutz, J. (2012). Information and ethnic politics in africa. *British Journal of Political Science 43*, 345–373.

Dal Bó, E., P. Dal Bó, and R. Di Tella (2006). "plata o plomo?": Bribe and punishment in a theory of political influence. *American Political Science Review 100*(1), 41–53.

Dal Bó, E., P. Dal Bó, and J. Synder (2009). Political dynasties. *The Review of Economic Studies 76*, 115–142.

Dal Bó, E. and R. Di Tella (2003, October). Capture by Threat. *Journal of Political Economy 111*(5), 1123–1152.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017). Who becomes a politician? Forthcoming QJE.

de Figueiredo, M., D. Hidalgo, and Y. Kasahara (2012). When do voters punish corrupt politicians? experiment evidence from brazil. Mimeo.

De Paola, M. and V. Scoppa (2011). Political competition and politician quality: evidence from italian municipalities. *Public Choice 148*, 547–559.

Di Lonardo, L. (2017). Valence uncertainty and the nature of the candidate pool in elections. *Journal of Theoretical Politics 29*(2), 327–350.

Diermeier, D., M. Keane, and A. Merlo (2005). A political economy model of congressional careers. *American Economic Review 95*(1), 347–373.

Eggers, A. and J. Hainmueller (2009). Mps for sale? rerturns to office in postwar british politics. *American Political Science Review 103*(4), 513–533.

Eifert, B., E. Miguel, and D. Posner (2010). Political competition and ethnic identification in africa. *American Journal of Political Science 54*(2), 494–510.

Faccio, M. (2006, March). Politically connected firms. *American Economic Review 96*(1), 369–386.

Fafchamps, M. and J. Labonne (2017). Do politicians' relatives get better jobs? evidence from municipal elections. *The Journal of Law, Economics & Organization 33*, 268–300.

Fearon, J. D. (1999). Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance. *Democracy, accountability, and representation 55*, 61.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1/3), 5–25.

Ferraz, C. and F. Finan (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 703–745.

Ferraz, C. and F. Finan (2011, June). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review 101*(4), 1274–1311.

Ferree, K. (2006). Explaining south africa's racial census. *68 68*(4), 803–815.

Fisman, R. (2001, September). Estimating the value of political connections. *American Economic Review 91*(4), 1095–1102.

Fisman, R., N. Harmon, E. Kamenica, and I. Munk (2015). Labor supply of politicians. *Journal of the European Economic Association 13*(5), 871–905.

Fisman, R., F. Schulz, and V. Vig (2014). The private returns to public office. *Journal of Political Economy 122*(4), 806–862.

Fisman, R., F. Schulz, and V. Vig (2016). Financial disclosure and political selection: Evidence from india. Mimeo.

Folke, O., T. Persson, and J. Rickne (2016). Dynastic political rents? economic benefits to relative of top politicians. Unpublished.

Gagliarducci, S. and M. Manacorda (2016). Politics in the family: Nepotism and the hiring decisions of italian firms. Technical report, IZA DP No. 9841.

Gagliarducci, S. and T. Nannicini (2013). Do better paid politicians perform better? disentangling incentives from selection. *Journal of the European Economic Association 11*(2), 369–398.

Galasso, V. and T. Nannicini (2011). Competing on good politicians. *The American Political Science Review 105*(1), 79–99.

Galasso, V. and T. Nannicini (2017, Jun). Political selection under alternative electoral rules. *Public Choice 171*(3), 257–281.

Gavoille, N. and M. Verschelde (2017). Electoral competition and political selection: an analysis of the activity of french deputies, 1958-2012. *European Economic Review 92*, 180–195.

Gehlbach, S., K. Sonin, and E. Zhuravskaya (2010). Businessman candidates. *American Journal of Political Science 54*(3), 718–736.

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science 45*(4), 862–886.

Hirano, S. and J. Snyder (2014). Primary elections and the quality of elected officials. *Quarterly Journal of Political Science 9*, 473–500.

Hobolt, S. and B. Hoyland (2011). Selection and sanctioning in european parliamentary elections. *British Journal of Political Science 41*(3), 477–498.

Hummel, P. (2013). Candidate strategies in primaries and general elections with candidates of heterogeneous quality. *Games and Economic Behavior 78*(C), 85–102.

Humphreys, M. and J. Weinstein (2012). Policing politicians: Citizen empowerment and political accountability in uganda. Mimeo.

Jones, B. F. and B. A. Olken (2005). Do leaders matter? national leadership and growth since world war ii. *The Quarterly Journal of Economics 120*(3), 835–864.

Kamada, Y. and T. Sugaya (2014). Valence candidates and ambiguous platforms in policy announcement games. Unpublished.

Kendall, C., T. Nannicini, and F. Trebbi (2015). How do voters respond to information? evidence from a randomized campaign. *American Economic Review*.

Khwaja, A. I. and A. Mian (2005). Do lenders favor politically connected firms? rent provision in an emerging financial market. *The Quarterly Journal of Economics 120*(4), 1371–1411.

Kostadinova, T. (2009). Abstain or rebel: Corruption perceptions and voting in east european elections. *Politics and Policy 37*(4), 691–714.

Kotakorpi, K. and P. Poutvaara (2011). Pay for politicians and candidate selection: An empirical analysis. *Journal of Public Economics 95*, 877–885.

Kotakorpi, K., P. Poutvaara, and M. Terviö (2016). Returns to office in national and local politics. Technical report, IZA DP No. 10003.

Lamba, S. and D. Spears (2013). Caste, cleanliness and cash: Effects of caste-based political reservations in rajasthan on a sanitation prize. *The Journal of Development Studies 49*(11), 1592–1606.

Lee, D. S. (2008, February). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics 142*(2), 675–697.

Lee, D. S., E. Moretti, and M. J. Butler (2004). Do voters affect or elect policies? evidence from the u. s. house*. *The Quarterly Journal of Economics 119*(3), 807–859.

Lenz, G. and K. Lim (2009). Getting rich(er) in office? corruption and wealth accumulation in congress. Unpublished.

Londregan, J. and T. Romer (1993). Polarization, incumbency, and the personal vote. In W. A. Barnett, M. Hinich, and N. Schofield (Eds.), *Political Economy: Institutions, Competition, and Representation*. New York: Cambridge University Press.

Lublin, D. (1994). Quality, not quantity: Strategic politicians in u.s. senate elections, 1951-1990. *The Journal of Politics 56*(1), 228–241.

Lundqvist, H. (2013). Is it worth it? on the returns to holding political office. Technical report, Barcelone Institute of Economics Working Paper.

Malesky, E., P. Schuler, and A. Tran (2012). The adverse effects of sunshine: A field experiment on legislative transparency in an authoritarian assembly. *American Political Science Review 106*(4), 762–786.

Mattozzi, A. and A. Merlo (2008). Political careers or career politicians? *Journal of Public Economics 92*, 597–608.

Mattozzi, A. and A. Merlo (2015). Mediocracy. *Journal of Public Economics 130*(C), 32–44.

McMillan, J. and P. Zoido (2004). How to subvert democracy: Montesinos in peru. *Journal of Economic Perspectives 18*(4), 69–92.

Messner, M. and M. Polborn (2004). Paying politicians. *Journal of Public Economics 88*, 2423–2445.

Meyersson, E. (2014). Islamic rule and the empowerment of the poor and pious. *Econometrica 82*(1), 229–269.

Munshi, K. and M. Rosenzweig (2016). Insiders and outsiders: Local ethnic politics and public good provision. Mimeo.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2016, October). Building state capacity: Evidence from biometric smartcards in india. *American Economic Review 106*(10), 2895–2929.

Niehaus, P. and S. Sukhtankar (2013). The marginal rate of corruption in public programs: Evidence from india. *Journal of Public Economics 104*, 52–64.

Norris, P. and R. Mattes (2003). Does ethethnic determine support for the governing party? the structural and attitudinal basis of partisan identification in 12 african nations. Mimeo.

Osborne, M. and A. Slivinski (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics 111*(1), 65–96.

Padró i Miquel, G. and J. Snyder (2006). Legislative effectiveness and legislative careers. *Legislative Studies Quarterly 31*(3), 347–381.

Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? theory and evidence from india. *The American Economic Review 93*(4), 1132–1151.

Pande, R. (2011). Can informed voters enforce better governance? experiments in low-income democracies. *Annual Review of Economics 3*, 215–237.

Pecenco, M. and S. Sukhtankar (2018, February). The impact of political rents on political selection: Evidence from india. *Working Paper*, 2895–2929.

Peichl, A., N. Pestel, and S. Siegloch (2013). The politicians wage gap: insights from german members of parliament. *Public Choice 156*(3), 653–676.

Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy.* Zeuthen lecture book series. MIT Press.

Poutvaara, P. and T. Takalo (2007). Candidate quality. *International Tax and Public Finance 14*(1), 7–27.

Querubin, P. (2015). Family and politics: Dynastic persistence in the philippines. Mimeo, New York University.

Querubin, P. and J. Synder (2009). Returns to u.s. congressional seats in the mid-19th century. In E. Aragones, C. Bevia, H. Llavador, and N. Schofield (Eds.), *Political Economy of Democracy.* Madrid: Fundación BBCA.

Reinikka, R. and J. Svensson (2004). Local capture: Evidence from a central government transfer program in uganda. *The Quarterly Journal of Economics 119*(2), 679–705.

Schofield, N. (2003, October). Valence Competition in the Spatial Stochastic Model. *Journal of Theoretical Politics 15*(4), 371–383.

Smart, M. and D. M. Sturm (2013). Term limits and electoral accountability. *Journal of Public Economics 107*(C), 93–102.

Squire, P. (1995). Candidates, money, and voters: Assessing the state of congressional elections research. *Political Research Quarterly 48*(4), 891–917.

Stokes, D. E. (1963). Spatial models of party competition. *The American Political Science Review 57*(2), 368–377.

Szakonyi, D. (2017). Businesspeople in elected office: Identifying private benefits from firm-level returns. *American Political Science Review*, 117.

Washington, E. L. (2008, March). Female socialization: How daughters affect their legislator fathers. *American Economic Review 98*(1), 311–32.

Winters, M. and R. Weitz-Shapiro (2013). Lacking information or condoning corruption: When do voters support corrupt politicians? *Comparative Politics*.
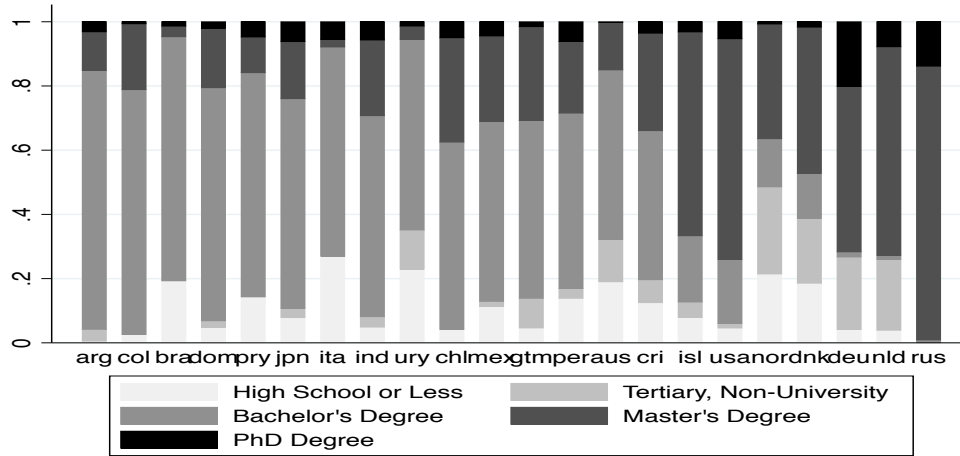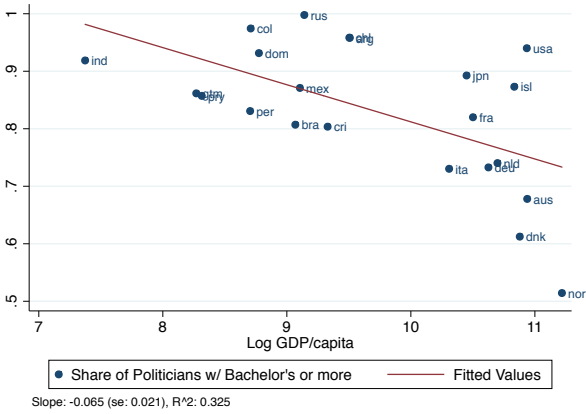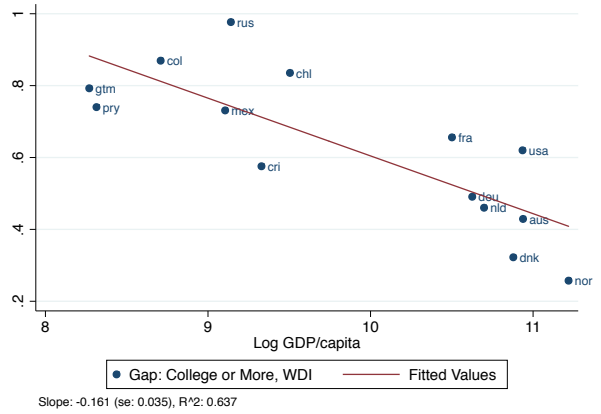
6 Figures

Figure 1: Cross-country variation in competence - education levels

Notes: This figure plots the proportion of national legislators by education level. Source: publicly available data obtained by the authors from congressional websites.

(a) Levels　　　　　　　　　　　　　　　　(b) Gap

Figure 2: Politician competence and GDP per capita

Notes: Panel A plots the relationship between GDP per capita and the percentage of legislators with a college degree or higher. Panel B plots the relationship between GDP per capita and the percentages of legislators with higher education relative to the respective national population. Source: publicly available data obtained by the authors from congressional websites.
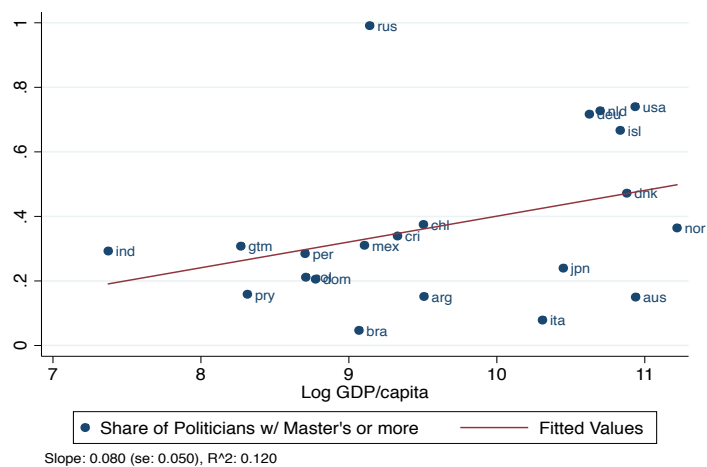
Figure 3: Politician competence and GDP per capita

Notes: This figure plots the relationship between GDP per capita and the percentage of legislators with a master's degree or higher. Source: publicly available data obtained by the authors from congressional websites.

7   Appendix

Proof of Proposition 1: We separate the analysis for the cases with $c(v_j) = 1$ and $c(v_j) = v_j, j = r, l$. We start with the first one.

Constant costs $(c(v_j) = 1)$

As stated in the main text, a candidate with type v_j runs iff

$$\frac{b-1}{b} \geq P(-j \text{ runs}) P\left(s_j < \frac{1}{2}|v_j\right). \tag{8}$$

In order to develop the expressions in the right hand side we postulate the existence of the equilibrium described in the proposition, namely one in which candidates run iff their type is above a cutoff v_j^*, where v_j^* is the equilibrium choice among possible cutoff choices $v_j' \in [0, \bar{v}_j]$. In this case, the type v_{-j} of candidates that j may confront is drawn from a uniform distribution in $[\underline{v}'_{-j}, \bar{v}_{-j}]$. Candidates with types v_l and v_r then want to run iff, respectively,

$$l: \frac{b-1}{b} \geq \frac{\bar{v}_r - v_r'}{\bar{v}_r} P(v_r \geq v_l - x|v_r \geq v_r')$$

$$r: \frac{b-1}{b} \geq \frac{\bar{v}_l - v_l'}{\bar{v}_l} P(v_l > v_r + x|v_l \geq v_l'),$$

where $x \equiv \frac{\phi[\rho R - (1-\rho)L]}{\varpi[\rho R + (1-\rho)L]}$ and where by assumption $b > 1$. Since $P(v_r \geq v_l - x|v_r \geq v_r') = \left(1 - \frac{v_l - x - v_r'}{\bar{v}_r - v_r'}\right)$ and $P(v_l > v_r + x|v_l \geq v_l') = \left(1 - \frac{v_r + x - v_l'}{\bar{v}_l - v_l'}\right)$ one immediately sees that the cutoff choices v_l', v_r' are strategically independent, and that the right hand side of both conditions is decreasing in the respective own types v_l, v_r. Therefore candidates run iff their types v_l, v_r satisfy the respective conditions,

$$v_l \geq \max\left\{0, \min\left\{\bar{v}_r \frac{1}{b} + x, \bar{v}_l\right\}\right\}$$

$$v_r \geq \max\left\{0, \min\left\{\bar{v}_l \frac{1}{b} - x, \bar{v}_r\right\}\right\}.$$

Assuming without loss of generality that r is the electorally strong party $(\rho R \geq (1-\rho)L)$, it is immediate that we have an interior equilibrium whenever $\bar{v}_r \frac{1}{b} + x < \bar{v}_l$ and $\bar{v}_l \frac{1}{b} > x$.

Increasing costs $(c(v_j) = v_j)$

The inequality in (5) governing candidate j's decision to run becomes,

$$c(v_j) = v_j: \frac{b - v_j}{b} \geq P(-j \text{ runs}) P\left(s_j < \frac{1}{2}|v_j\right).$$

We look for an equilibrium where l candidates enter iff $v_l \in [0, v_l^*]$ and the r candidates enter if $v_r \in [0, v_r^*]$ and where, since $x \geq 0$ we also have $v_r^* \geq v_l^*$. Given a putative cutoff v_j', and an equilibrium where candidates run as long as $v_j \leq v_j'$, the probability candidate j runs is

now $\frac{v'_j}{\bar{v}_j}$. Then, using the uniform distribution, candidates with types v_l, v_r run as long the respective following conditions hold,

$$\frac{b - v_l}{b} \geq \frac{v'_r}{\bar{v}_r} \min \left\{ \max \left\{ 0, 1 - \frac{v_l - x}{v'_r} \right\}, 1 \right\} \tag{9}$$

$$\frac{b - v_r}{b} \geq \frac{v'_l}{\bar{v}_l} \min \left\{ \max \left\{ 0, 1 - \frac{v_r + x}{v'_l} \right\}, 1 \right\}. \tag{10}$$

In an equilibrium where all types up to a given cutoff run, the left hand side in these inequalities must (start higher and) decrease faster in v_j than the right hand side, requiring $b < \bar{v}_{-j}$ for $j = r, l$, as assumed.

We first rule out $v_l^* > v_r^*$. Since $v_l^* \leq b$ (to ensure $P(l \text{ wins})b \geq c(v_l) = v_l$), it must be true that $v_r^* < b$. But consider a type who is not running $v_r = b - \varepsilon$ with $0 < \varepsilon < b - v_r^*$. For this type the left hand side of the running condition (10) reads $\frac{\varepsilon}{b}$ and the right hand side $\min \left\{ 0, \frac{v_l^* - b + \varepsilon - x}{\bar{v}_l} \right\} < \frac{\varepsilon}{b}$ implying this type strictly prefers to run, yielding a contradiction, so $v_l^* > v_r^*$ cannot hold in equilibrium.

If there is an equilibrium, it must be the case that $v_r^* \geq v_l^*$. Thus, a type $v_r = v_r^*$ wins for certain and therefore in equilibrium we cannot have $v_r^* < b$ (if this held, then there would be a type $\hat{v}_r > v_r^*$ for whom $P(r \text{ wins} \mid v_r = \hat{v}_r) = 1$ and $b > \hat{v}_r$ so running would be strictly profitable). Also, we cannot have $v_r^* > b$ (candidates with $v_r > b$ would make a loss as $c(v_r) = v_r > b$). Therefore, in equilibrium we can only have $v_r^* = b$. Given this, the l candidate would run iff:

$$\frac{b - v_l}{b} \geq \frac{b}{\bar{v}_r} min \left(1 - \frac{v_l - x}{b}, 1 \right) \Leftrightarrow v_l \leq b \left(1 - \frac{x}{\bar{v}_r - b} \right) \equiv v_l^*.$$

Types $v_l \in [0, x]$ stand no chance in a contested election. But types $v_l \in \left[0, \frac{b}{\bar{v}_r} (\bar{v}_r - b) \right]$ are willing to run just in the off-chance that r does not run. Types $v_l \in \left[\frac{b}{\bar{v}_r} (\bar{v}_r - b), v_l^* \right]$ are only willing to run given they stand a chance in a contest, and since $x \leq \frac{b}{\bar{v}_r} (\bar{v}_r - b)$, they do. It is easy to check that all r candidates with $v_r \leq b$ prefer to run, completing the characterization of the equilibrium.∎

Proof of Proposition 3:

We first tackle the case of a constant cost of running, and deal with the case of increasing costs right after.

(i) **Constant costs** $(c(v_j) = 1)$

We study the expression (7) under the restriction $\bar{v}_l = \bar{v}_r = \bar{v}$. We will also evaluate a number of expressions at the point of symmetry where $x = 0$ and thus $v_r^* = v_l^*$ and standard continuity considerations will extend the result to a neighborhood near the point of symmetry.

The relevant terms to be differentiated are,

$$E\left(v|r \text{ alone}\right) = \frac{\bar{v} + v_r^*}{2}$$

$$E\left(v|l \text{ alone}\right) = \frac{\bar{v} + v_l^*}{2}$$

$$E\left(v|\text{both}\right) = \frac{\int_{v_l^*}^{\bar{v}_l} \int_{v_l-x}^{\bar{v}_r} v_r dv_r dv_l + \int_{v_l^*}^{\bar{v}_l} \int_{v_r^*}^{v_l-x} v_l dv_r dv_l}{\left(\bar{v}_r - v_r^*\right)\left(\bar{v}_l - v_l^*\right)} \equiv \Xi$$

and,

$$P\left(r \text{ alone}\right) = \frac{\frac{v_l^*}{\bar{v}}\frac{\bar{v}-v_r^*}{\bar{v}}}{1 - \frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}} = \frac{v_l^*\left(\bar{v} - v_r^*\right)}{\bar{v}^2 - v_l^* v_r^*}$$

$$P\left(l \text{ alone}\right) = \frac{\frac{v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}}{1 - \frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}} = \frac{v_r^*\left(\bar{v} - v_l^*\right)}{\bar{v}^2 - v_l^* v_r^*}$$

$$P\left(\text{both}\right) = \frac{\frac{\bar{v}-v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}}{1 - \frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}} = \frac{\left(\bar{v} - v_r^*\right)\left(\bar{v} - v_l^*\right)}{\bar{v}^2 - v_l^* v_r^*}.$$

These follow because there are only two ways in which, say, r gets to run alone. As explained in footnote 9, in a given candidate draw iteration, with probability $\frac{v_l^*}{\bar{v}}\frac{\bar{v}-v_r^*}{\bar{v}}$ r runs and l does not. Alternatively, nobody runs, which happens with probability $\frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}$, in which case there is a re-draw and r faces the same chance as in the previous iteration of ending up running alone. Thus, $P\left(r \text{ alone}\right) = \frac{v_l^*}{\bar{v}}\frac{\bar{v}-v_r^*}{\bar{v}} + \frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}P\left(r \text{ alone}\right)$, implying $P\left(r \text{ alone}\right) = \frac{\frac{v_l^*}{\bar{v}_l}\frac{\bar{v}-v_r^*}{\bar{v}_r}}{1 - \frac{v_r^*}{\bar{v}_r}\frac{v_l^*}{\bar{v}_l}}$, which yields the expressions above.

We first analyze the "candidate quality" effect ($\frac{dE(v|r \text{ alone})}{db}P\left(r \text{ alone}\right) + \frac{dE(v|l \text{ alone})}{db}P\left(l \text{ alone}\right) + \frac{dE(v|\text{both})}{db}P\left(\text{both}\right)$) and show it to be negative. In the scenarios where $j = r,l$ run alone, this is immediate, as average quality in those cases is $\frac{\bar{v}+v_j^*}{2}$ and $\frac{dv_j^*}{db} < 0$ from proposition 2. That $E\left(v|\text{both}\right)$ (the expected quality of the winner in a contest) decreases in b is more tedious to show. A sketch of the steps is as follows:

$$\frac{dE\left(v|\text{both}\right)}{db} = \frac{d\Xi}{dv_r^*}\frac{dv_r^*}{db} + \frac{d\Xi}{dv_l^*}\frac{dv_l^*}{db}$$

$$= -\frac{\bar{v}}{b^2}\left(\frac{1}{\left(\bar{v}_r - v_r^*\right)}\left[\Xi - \left(\frac{\bar{v}_l + v_l^*}{2}\right)\right] + \frac{\Xi\left(\bar{v}_r - v_r^*\right) - \left(\begin{array}{c}\frac{\bar{v}_r^2}{2} + \frac{\bar{v}_l^2}{2} - \frac{\left(v_l^* - x\right)^2}{2} - \frac{v_l^{*2}}{2} \\ + \left(v_l^* - x - v_r^*\right)v_l^*\end{array}\right)}{\left(\bar{v}_r - v_r^*\right)\left(\bar{v}_l - v_l^*\right)}\right).$$

After some algebra, and taking $x \to 0$ so $v_l^* \to v_r^*$, one can show the above expression to be negative iff $2\Xi - \frac{\bar{v}+v^*}{2} > 0$, which holds. To see this, it helps to establish first that $\Xi > \frac{\bar{v}+v^*}{2}$

which says that the expected quality of the winner of a (symmetric) contest is higher than the expected quality of the winner of a solo election. The expression $\Xi > \frac{\bar{v}+v^*}{2}$ can be developed into $(\bar{v}-v^*)\frac{\bar{v}^2+v^{*2}}{2} > \int_{v^*}^{\bar{v}} v_l^2 dv_l$, which is true from the convexity of the quadratic on the right hand side.

We next show the competitive effect to be positive. Since as just established, under electoral symmetry $\Xi > \frac{\bar{v}+v^*}{2}$, the competitive effect is positive whenever an increase in b raises $P(\text{both})$ and decreases $P(j \text{ alone})$, $j = r, l$. Since $P(\text{both}) = 1 - P(l \text{ alone}) - P(r \text{ alone})$ it sufficient to show $\frac{dP(j \text{ alone})}{db} < 0$; $j = r, l$ which are the same under symmetry.

$$
\frac{dP(r \text{ alone})}{db} = \frac{\left(-\frac{dv_r^*}{db}v_l^* + (\bar{v}_r - v_r^*)\frac{dv_l^*}{db}\right)\left(\bar{v}_r\bar{v}_l - v_r^*v_l^*\right) + (\bar{v}_r - v_r^*)v_l^*\left(\frac{dv_r^*}{db}v_l^* + \frac{dv_l^*}{db}v_r^*\right)}{\left(\bar{v}_r\bar{v}_l - v_r^*v_l^*\right)^2},
$$

which is negative iff $(v_l^* - (\bar{v} - v_r^*))(\bar{v}^2 - v_r^*v_l^*) < (\bar{v} - v_r^*)v_l^*(v_l^* + v_r^*)$, which then using symmetry becomes $\bar{v}(\bar{v} - v^*) > (\bar{v} - v^*)v^*$ which is true. Thus, the preceding argument establishes the potential for ambiguity in a neighborhood around symmetry. Additional straightforward algebra shows that as $x \to 0$ the addition of the candidate quality and competition effects $A + B$ becomes negative - in the limit the candidate quality effect dominates the competition effect.

(ii) **Increasing costs** $(c(v_j) = v_j)$

We study the expression (7) under the restriction $\bar{v}_l = \bar{v}_r = \bar{v}$. The relevant terms to be differentiated are,

$$
E(v|r \text{ alone}) = \frac{v_r^*}{2}
$$
$$
E(v|l \text{ alone}) = \frac{v_l^*}{2}
$$
$$
E(v|\text{both}) = \frac{\int_0^x \int_0^{v_r^*} v_r dv_r dv_l}{v_r^* v_l^*} + \frac{\int_x^{v_l^*} \int_{v_l - x}^{v_r^*} v_r dv_r dv_l}{v_r^* v_l^*} + \frac{\int_0^{v_l^* - x} \int_{v_r + x}^{v_l^*} v_l dv_l dv_r}{v_r^* v_l^*} \equiv \Xi
$$

and,

$$
P(r \text{ alone}) = \frac{\frac{v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}}{1 - \frac{\bar{v}-v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}} = \frac{v_r^*(\bar{v} - v_l^*)}{\bar{v}(v_l^* + v_r^*) - v_l^* v_r^*}
$$

$$
P(l \text{ alone}) = \frac{\frac{v_l^*}{\bar{v}}\frac{\bar{v}-v_r^*}{\bar{v}}}{1 - \frac{\bar{v}-v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}} = \frac{v_l^*(\bar{v} - v_r^*)}{\bar{v}(v_l^* + v_r^*) - v_l^* v_r^*}
$$

$$
P(\text{both}) = \frac{\frac{v_r^*}{\bar{v}}\frac{v_l^*}{\bar{v}}}{1 - \frac{\bar{v}-v_r^*}{\bar{v}}\frac{\bar{v}-v_l^*}{\bar{v}}} = \frac{v_r^* v_l^*}{\bar{v}(v_l^* + v_r^*) - v_l^* v_r^*}.
$$

From proposition 2, an increase in b raises cutoffs v_l^*, v_r^* if x is small enough, and it is straightforward to show that this implies $\frac{dP(\text{both})}{db} > 0$ and $\frac{dP(r \text{ alone})}{db} < 0, \frac{dP(l \text{ alone})}{db} < 0$. So

if the competitive effect leads $E\left(v|\text{both}\right)$ to be larger than $E\left(v|r\text{ alone}\right)$ and $E\left(v|l\text{ alone}\right)$ (which must be true under symmetry) the sum of terms in A will be positive. Since higher cutoffs increase $E\left(v|r\text{ alone}\right), E\left(v|l\text{ alone}\right)$ and $E\left(v|\text{both}\right)$, the effect must be overall positive, proving the result. That $\frac{dE(v|r\text{ alone})}{db} > 0, \frac{dE(v|l\text{ alone})}{db} > 0$ is obvious. We now confirm that under electoral symmetry we find $\frac{dE(v|\text{both})}{db} > 0$. Note,

$$\frac{d\Xi}{db} = \frac{\left(\begin{array}{c} -\int_0^x v_r^* dv_l \frac{dv_r^*}{db} + \int_x^{v_l^*} v_r^* dv_l \frac{dv_r^*}{db} \\ -\int_{v_l^*-x}^{v_r^*} v_r dv_r \frac{dv_l^*}{db} + \int_0^{v_l^*-x} v_l^* dv_r \frac{dv_l^*}{db} \end{array}\right) v_r^* v_l^* - \Xi v_r^* v_l^* \left(\frac{dv_r^*}{db} v_l^* + v_r^* \frac{dv_l^*}{db}\right)}{\left(v_r^* v_l^*\right)^2}$$

$$\frac{d\Xi}{db} = \frac{\int_0^{v_l^*} v_r^* dv_l \frac{dv_r^*}{db} - \int_{v_l^*-x}^{v_r^*} v_r dv_r \frac{dv_l^*}{db} + \int_0^{v_l^*-x} v_l^* dv_r \frac{dv_l^*}{db} - \Xi \left(\frac{dv_r^*}{db} v_l^* + v_r^* \frac{dv_l^*}{db}\right)}{v_r^* v_l^*}.$$

In the case of symmetry, we have $x = 0$, $\frac{dv_r^*}{db} = \frac{dv_l^*}{db}$ and $v_r^* = v_l^* = v^*$, so we can write,

$$\frac{d\Xi}{db} = \frac{dv^*}{db} \frac{v^{*2} + v^{*2} - \Xi 2 v^*}{v^{*2}}$$
$$= \frac{dv^*}{db} 2 \frac{v^* - \Xi}{v^*} > 0$$

from the fact that $\Xi < v^*$, and $\frac{dv^*}{db} > 0$.∎