Delegation and Recruitment in Organizations: The Slippery Slope to "Bad" Leadership*

Selcen Çakır† Konstantinos Matakos‡ Janne Tukiainen§

November 26, 2024

Abstract

We construct a dynamic model of two-sided matching in labor markets with multi-dimensional agent and firm heterogeneity. We apply it to study optimal party structure and the decision of how (de)centralized candidate recruitment should be. Parties are non-unitary actors and compete at the *local* markets over recruitment of competent candidates. Local organizers possess an informational advantage over the distribution of politicians' skill, which is positively related to electoral rent generation. Party leadership has a dual objective: they want simultaneously to maximize a) the organization's rents and b) their retention probability. Thus, when deciding how centralized recruiting should be, leaders face a trade-off: delegating selection to local party organizations harnesses all available information and increases electoral returns, but also limits a leader's ability to stack the organization with loyalists who are more likely to retain her should a (stochastic) leadership challenge arrives. Moreover, ideology alters this trade-off in ways that generate welfare non-monotonicities. We characterize an equilibrium delegation rule with two key properties: a) some high-skilled politicians may select into lower performing parties due to ideological alignment, and b) more *moderate* and *competent* leaders rely excessively on market-based recruitment and, as a result, survive relatively shorter at the helm of the organization.

JEL codes: C78, D72, D82, J40, L20, M50

Keywords: delegation, intraparty organization, recruitment, information, two-sided matching

*For valuable comments and feedback, we thank Ben Marx, Vincent Pons, Jeff Frieden, Pia Raffler, David Yang, Sulin Sardoschau, Clement Minaudier, Carlo Prato, Ernesto Dal Bo, Charles Angelucci, Nico Voigtlander, Olle Folke, Harto Saarien, Orhan Aygün and participants in the following seminars and conferences: Harvard University, EUI, LSE, King's College London, University of Nottingham (NICEP 2024), EEA 2023, EPSA 2023, Madrid-PERG, and Bilkent University. This research is funded by the European Union (ERC, INTRAPOL, 101045239). Views and opinions expressed are only those of the authors, however, and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

†Boğaziçi University, Department of Economics, and University of Turku, Department of Economics. selcen.cakir@boun.edu.tr.

‡*Corresponding author:* King's College London, Department of Political Economy (Bush House NE), London WC2B 4BG, UK. konstantinos.matakos@kcl.ac.uk

§University of Turku, Department of Economics. janne.tukiainen@utu.fi.

1 Introduction

In a representative democracy, the nature and quality of political selection is of fundamental importance. The quality of politicians has attracted a great deal of scholarly attention (Besley (2005); Dal Bó and Finan (2018)), under the premise that better political selection improves the quality of government. Indeed, a large set of literature shows that decision-makers –from executives to individual members of legislatures– can influence the outcomes of policy-making.[1]

From a theoretical perspective, many authors have modeled political selection processes as a result of self-selection by candidates and screening by voters (Besley (2004); Caselli and Morelli (2004); Poutvaara and Takalo (2007); Mattozzi and Merlo (2008); Smart and Sturm (2013)). However, in a representative democracy, political parties also play a consequential role: they are at the core of citizens' representation because they manage the political selection process across various channels. Political parties have gate keeping powers over candidate selection, and parties can either directly – through mechanisms such as list ranks in a closed-list PR system– or indirectly –through, for instance, campaign contributions– influence the electoral prospects of those candidates. Thus, the quality of democratic output depends on how a party choose to structure and organize itself internally. Accordingly, there are also formal models that bring in political parties that are strategic in candidate selection (Carrillo and Mariotti (2001); Mattozzi and Merlo (2015); Galasso and Nannicini (2017)). One take-away from these approaches to political selection is that positive selection is far from obvious; for instance, less competent individuals might have a comparative advantage at entering politics due to weaker outside options or parties preferring less costly mediocre candidates Cakir (2019). However, to our knowledge, the formal theoretical literature has been silent about the role of internal organization and rules that parties have in political selection.

In this type of framework, political parties have a consequential role: they are

[1] A large empirical literature inspired by the citizen-candidate models of Osborne and Slivinski (1996) and Besley and Coate (1997) has demonstrated how politicians' characteristics matter for policy. Studies concern, for example, the causal effects of political partisanship (Lee et al. (2004)), female representation (Chattopadhyay and Duflo (2004); Clots-Figueras (2012)), minority representation (Pande (2003)), and politicians' occupational background (Hyytinen et al. (2018); Kirkland (2020)), and politicians competence (Meriläinen (2022)) on policy outcomes.

at the heart of citizens' representation because they manage the selection process. But, just like any other complex organization, parties are structured internally around some rules –in other words, they have an intraparty constitution. In turn, a party's internal structure will also determine the selection process. The decision of whom to select to represent a political organization can be made centrally or downstream. Hence, an important –yet unanswered question– remains: how are such rules chosen and why? Which entity within a party gets to decide the process that governs political selection? In the spirit of Barbera and Jackson (2004), we are interested in the positive characterization of intraparty constitutional arrangements.

We understand constitutions as a set of rules that govern the allocation of decision-making power among the members of an organization –such as a political party. To put it simply, the intraparty constitution determines the degree of power-sharing (see e.g., Cakir (2019); Dewan et al. (2015); Acemoglu and Robinson (2020)) between the members of the political organization. Thus, a more inclusive constitution will entail more power-sharing and a larger percentage of decisions delegated downstream (Invernizzi and Prato (2019)). This is what interest us: what is the optimal degree of intraparty delegation in decision-making processes? How much power would leaders delegate to downstream party members when it comes to political selection?

Our second question is normative in nature. That is, we are also interested in the welfare implications of delegation, leading us to question whether intraparty power-sharing is good or bad for political selection. In other words, do 'inclusive' parties with higher degrees of power-sharing (i.e., parties that delegate more) select higher quality politicians? What is the relationship between intraparty democracy and the quality of democratic representation? As numerous cases demonstrate, this also matters for the overall quality of a democratic polity. For instance, a more decentralized candidate selection process might make it easier for a party's legislators to replace an incompetent or authoritarian leader.

To address these questions we build a dynamic model of two-sided candidate recruitment and we endogenize the leader's choice of how (de)centralized the selection process should be. A leader cares both about maximizing her survival chances (her retention probability should an intraparty challenge be staged), as well as, about party electoral success because it reduces the probability of a leadership challenge occurring. These, in turn, depend on the type and characteristics, such as competence and loyalty, of the selected politicians/candidates. However, there is an apparent tension: selecting purely based on a candidate's competence increases the chances of electoral success but also relinquishes control over the identity/loyalty of the political personnel that is recruited.

We model the process of intraparty selection as follows. The *party leader* wants to maximize her chances of staying in power, but, to some degree, this also depends on making the right decisions in terms of selecting the party's political personnel that maximize the party's rents and probability of electoral success. Moreover, parties compete with each other over recruiting the most talented politicians. In turn, 'non-loyalist' politicians (that is, politicians who are not parachuted to the list by a leader) also care about maximizing their (and consequently their party's) re-election prospects, but they also put some weight in their ideological match with a given leader. Thus, they face a choice over which party to join. In other words, the recruitment of political talent is a two-sided matching process, which we explicitly model.

Our model underscores how, the key trade-off involved in deciding how decentralized selection should be –competence versus loyalty–, is modified by the presence of ideology and party competition. Since local party organizations hold an informational advantage over candidate recruitment, more delegation increases the chances of electoral success, which is positively related to the leader's survival. At the same time, however, increased delegation also increases the leader's probability of being replaced since there will be fewer loyalists among party ranks when a (stochastic) challenge arrives. In other words, delegating effective control, or real authority, over the party's selection and recruitment decisions is a double-edged sword: it helps the leader to reap the informational advantage at the cost of relinquishing party control. The latter echoes –albeit in a different setting– Aghion and Tirole (1997) who link the delegation of real authority in organizations with the structure of information. As a result, and due to its general nature, our model also speaks to similar recruitment and personnel sorting problems outside the realm of politics, as we discuss later in this paper.

Our main finding is that, in equilibrium, the degree of intraparty decentralization (delegation) increases with a leader's *competence* and ideological *moderation.* In other words, more extremist and less competent leaders delegate less out of *choice* not because of a character trait (although we find that incompetence is relatively a bigger drag). This result offers several insights into the relationship between intraparty constitutional arrangements and the quality of political representation, as well as into democratic politics in general. First, we provide a rational choice-based explanation of the link between authoritarian, non-inclusive leaders and ideological extremism. Extreme centralization of intraparty power, rather than mostly being a 'character' or psychological attribute of ideological extremist leaders, is predominantly a *strategic choice* made by them. Thus, our work offers a possible link to endogenize the relationship between ideological extremism and authoritarianism.

Second, we find that parties led by moderate and competent leaders dynamically expand over time because they allow local organizers to recruit more competent and moderate candidates/workers, as opposed to those led by extreme (and authoritarian) ones who instead fill the organization's ranks with loyalists. But while, in expectation, this increases the party's rent-generation ability and total assets it decreases their own survival probability. The latter observation also points to an interesting feature of party competition that we refer to as the 'slippery slope' of intraparty politics. This feature describes the idea that competent and moderate leaders are likely to be replaced more frequently because they are more likely to relinquish control over the party's selection process. That is, 'good' (competent and moderate) leaders fall victims to their tendency to delegate, leading them to delegate much more compared to extremist (and also sometimes competent) ones, in equilibrium. In other words, and without having ex ante differential risk preferences, moderate and competent leaders rationally assume more risk –they prefer higher the variance in expectation. This novel insight, that comes from opening the black box of intraparty politics and selection (Dal Bó and Finan (2018)), is in our view another manifestation of the general idea of the 'narrow corridor' of democratic politics (see Acemoglu and Robinson (2020)) applied within a political organization.

This relatively lower expected survival rate of moderate and competent leaders, in equilibrium, is also theoretically 'surprising' because it cannot be solely attributed to the –well-understood– trade-off between loyalty and competence; the latter follows from the fact that its terms-of-trade are relatively more beneficial for moderate and competent leaders.[2] Rather, it is due to leaders' preference to also match on ideology with their selected candidates. Since moderate politicians constitute the larger fraction of the labor market for skilled non-loyalists, moderate and competent leaders have an absolute advantage in making better matches when recruiting from the market: their type is more appealing and, hence, competent non-loyalists are more likely to select into a party led by them. As a result, moderate leaders endogenously have a higher valuation of the pool of candidates/resources (accessible via delegation).

Taken together with the –common for all leaders– asymmetric information problem, delegation presents moderate and competent leaders with the possibility to 'kill two birds with one stone'. Not only they resolve asymmetric information, but they can also fill the party ranks with high-skill, ideological lookalikes. That is, compared to simply packing the racks with loyalist, delegation presents moderate (and competent) politicians with a 'favorably loaded dice' which makes them willing to pay a higher price (that is, accept a relatively higher risk of replacement) in order to source the best political talent (also likely of similar ideological inclination). This, in turn, translates to a relatively shorter leadership tenure, and an equilibrium delegation rate that is higher (relative to a case where ideology was absent). Thus creating a dynamically expanding, more resourced, and ideologically more moderate (and more aligned) party comes at the cost of their own

[2]Given that loyalist candidates are drawn from a lower-mean distribution and always vote for the incumbent leader irrespective of her type, the marginal benefit of recruiting a loyalist (i.e. no delegation) is *constant* across *all leader types*. At the same time, because non-loyalist politicians –recruited from the local market– are simultaneously more skilled and more likely to be moderate, a 'good' leader's marginal benefit of selecting (via delegation) is *higher for any level* of delegation, compared to all other leader types. This is so, because good leaders always find a better match that, not only is more competent thus maximizing their assets, but is also more likely to retain them should a challenge arrives. Hence, given identical time/risk preferences (and no behavioral assumptions) across leaders, in equilibrium, any rational leader will delegate optimally up to the point where the marginal benefit of recruiting a non-loyalist equals that of parachuting a loyalist. In equilibrium 'good' leaders' will delegate more, yet, since this trade-off strictly benefits them in relative terms, it cannot explain their higher replacement rate unless the impact of ideology and the structure of political labor markets are taken into account. Or, in other words, there is an additional force (ideology) that pushes the equilibrium delegation level –but also their replacement rate– upwards.

career prospects. The latter also gives a normative aspect to the term of 'good' leadership: not only they are more competent but, in equilibrium, they also put party above themselves. Yet, this feature is not due to a miscalculation, or ex ante higher risk preference, but to purely instrumental and self-serving reasons.

Finally, in the last part of the paper, we perform a welfare analysis which points to the limits that the vertical dissemination of information has. That is, we show that aggregate societal welfare is, generically, *not strictly increasing* with more information acquisition/diffusion (which, in our case, implies full delegation). This is a new insight that contrasts the conventional wisdom –found in other market settings– of 'more information being always welfare-enhancing'. Thus, our welfare analysis uncovers an interesting contrast between political and 'ordinary' markets. The latter can potentially explain differences in the quality and the organization of political selection across polities (see e.g., Besley, 2005; Dal Bó et al., 2017)

The rest of the paper is organized as follows. In section 2, we present the paper's main contributions and its connection to the literature. Section 3 presents our model, while section 4 presents our main results, comparative statics and welfare analysis. Section 5 concludes.

2 Contribution and Relation to Literature

By providing a tractable dynamic model of two-sided matching, we highlight the presence of a key trade-off between ideological match (similar to product homogeneity) and competence –for a non-loyalist politician– on the one hand and for party (organizational) success on the other, which we model as maximizing the probability of collecting rents. The insight we gain by exposing this trade-off is that there exists a tension between organizational and individual (leader or politician) success in terms of rents/wages collected. This tension gets amplified by multi-dimensional agent heterogeneity and the stochastic, uncertain nature of the leader's survival at the helm of the party. Thus, our work contributes to the following strands of literature.

First, we contribute to the burgeoning literature on intraparty politics and the role that parties play in the process of candidate selection (see e.g., Dal Bó and

Finan (2018)).[3] Here, our contribution is twofold. First, we advance the literature by explicitly modelling how the choice of internal party rules and constitutions (selection/delegation rules) map into candidate selection and party success. To our knowledge, this is the first paper that endogenizes the choice of a political organization's personnel recruitment procedures. Second, we formally model both dimensions of competition (within and between parties) that political organizations engage with. To achieve this, we model the selection process as a two-sided matching market, and we derive analytical solutions. This adds additional complexity to the model, but in return, we gain useful insights that relate to the 'narrow corridor' of intraparty democracy and the strategic, non-idiosyncratic link between authoritarianism and ideological extremism. This is a new insight in the study of intraparty politics. What is more, our theoretical results are consistent with recent empirical studies of populism, which find that while populism has large economic costs, populist leaders can nonetheless be long-lasting (see e.g., Funke et al. (2020)). This is another way of understanding our model's equilibrium where more extreme and incompetent leaders survive longer.

Second, due to our choice to explicitly model the political selection process as a two-sided competitive market, we can isolate and focus on the influence that between-party competition exerts in political selection. We gain two additional insights from this. First, we can explicitly derive the influence that between-party (or firm) competition has on the optimal within-party distribution of skills. To our knowledge, we are the first to do so. Second, we show that the relatively more competent and moderate party leaders *double-down on decentralization* and relinquish party control to local party organizations in order to attract better political talent and thus improve the party's election prospect. The latter comes at the cost of party leaders being replaced more often, should a challenge arise, since these leaders lack loyal, hand-picked allies among party ranks.

Third, our work contributes to the literature on organizational theory in bureaucracies. We endogenize the choice of the optimal structure of the organization –in our case, how decentralized decision-making should be and the optimal degree

[3]There are a few recent advances in the literature exploring the role of intraparty dynamics (see e.g., Cakir (2019); Buisseret et al. (2022); Matakos et al. (2018)) on political selection but from a different angle. Importantly, they do not endogenize the choice of the selection rule.

of downstream delegation of decision-making. In this sense, our work extends recent advances on understanding state bureaucratic organizations by Dal Bó et al. (2021) in a more complex setting since, unlike state bureaucratic organizations, parties operate in an environment of both vertical (within) and horizontal (between) competition. In a different direction, it also extends work by Dewan et al. (2015) who explore how ideology and information aggregation technology affect optimal centralization of authority in executive decision-making and its impact on policy quality. Our work also has important implications for span-of-control literature (see e.g., Bandiera et al. (2021); Eeckhout and Kircher (2012); Akcigit et al. (2018); Aghion and Tirole (1997)) and the incentives that managers and CEOs face in allocating authority and control over key organizational decisions, such as personnel recruitment or monitoring.

3 Model

This section develops a model that explains how political party leaders decide whether to involve lower-level entities, such as local branches, in the process of candidate selection. Leaders aim to balance two key goals: enhancing the party's chances of electoral success and maintaining their own leadership position. While local branches often have better insight into which candidates are likely to win, these candidates may not remain loyal if a leadership challenge arises.

The model highlights the "fragility of democratic politics," showing that less competent and more radical leaders tend to stay in power longer by favoring loyalty over competence. This dynamic can negatively affect the party's overall quality and growth. In contrast, more capable and centrist leaders may delegate candidate selection to local branches, which tends to improve party quality and growth. However, by doing so, they risk losing control over their leadership.

In Section 3.1, we outline the basic structure of the model. We then explore the equilibrium of two-sided matching between parties and politicians recruited by local organizations, conditional on leaders' decentralization rules (Section 3.2), and introduce the dynamics of leadership contests in Section 3.3. Finally, in Section 3.4, we derive the optimal rate of decentralization for party leaders and characterize the equilibrium properties of the model.

3.1 Model setup

Our model focuses on a given election cycle in a continuous time setting, and abstracts from dynamics related to transitions between different election cycles. The political arena is fragmented into ideological subgroups such as the left wing and the right wing, and we focus on an arbitrarily chosen subgroup. There are K parties indexed by $k \in \{1, 2, ..., K\}$. Leadership within these parties are selected from a continuous distribution of leaders, who are heterogeneous in their ideology and competence. A leader's ideology, denoted j, can be either moderate, M, or extremist, E. In our context, extremism means views/policies with fewer supporters. The probability that a leader will have a moderate ideology is p^M and we assume that $p^M > 0.5$.[4] Leaders further differ in their competence, η, which shows how productively they lead the party resources. This productivity parameter is distributed uniformly with support on [0,1].

Politicians are a disjointed set from leaders. There are two types of politicians: loyalist and non-loyalist. A loyalist always supports the incumbent when a leadership challenge arrives. The loyalists have heterogeneous amounts of political assets, distributed $L^L(z)$, with density $\ell^L(z) > 0$ over $[0, z^L_{max}]$.

Non-loyalists differ in both ideology and the amount of political assets they have. The ideology of a non-loyalist affects his leader choice when a party leadership competition occurs. Non-loyalist politicians' ideology is assumed to have the same distribution as party leaders' ideology. The heterogeneous amounts of j-ideology non-loyalist politicians' infinitesimal assets, denoted by z, are distributed according to $L^j(z)$ with positive density $\ell^j(z) > 0$ on $[0, z^j_{max}]$, for $j \in \{M, E\}$. Moderates are more abundant than the extremists in the economy. So, we have $L^M(z) \leq L^E(z), \forall z$. We impose the following structure on non-loyalist politicians' leader preferences.

Assumption 1 (Non-loyalist politicians' leader preferences)**.** *Let $u^{jj'}(\eta)$ denote the payoff of an ideology-j politician with a type-j' leader. The following holds*

i) $\frac{\partial u^{jj'}(\eta)}{\partial \eta} > 0$, $\forall j, j'$,

[4] More generally, we assume that there are two types of politicians, one type being more common. While we use the moderate-extremist types as an example throughout the text, our results apply to other cases where politicians differ along a key policy issue.

ii) $u^{jj}(\eta) > u^{jj'}(\eta)$ *for* $j \neq j'$,

iii) $u^{ME}(\eta') \geq u^{MM}(\eta)$ *for* $\eta' \geq \eta^{ME}(\eta)$,

iv) $u^{EM}(\eta') \geq u^{EE}(\eta)$ *for* $\eta' \geq \eta^{EM}(\eta)$.

Assumption 1 states that non-loyalist politicians prefer more competent leaders and tend to choose ideologically aligned leaders, unless competence differentials are significant.

All members of a party aggregate their assets to produce the party's electoral rents. We impose the following structure on a party's rent production technology.

Assumption 2 (A party's electoral rents)**.** *Party k's electoral rents,* $f_k(x_k;\eta)$ *are determined as a function of the total assets of the party,* x_k*, and the party leader's type. We assume that*

$$f_k(x_k;\eta,j) = x_k^{\eta}.$$

Assumption 2 states that a party's campaign production increases with its members' total assets and its leader's productivity.[5]

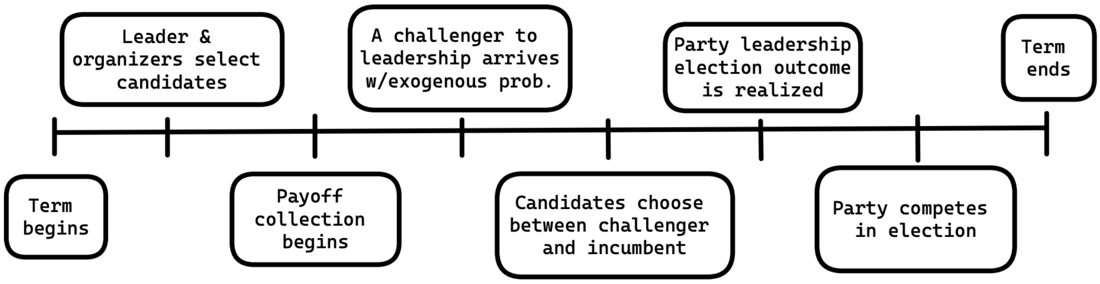
The timing of the events is summarized in Figure 1. At the beginning of a term, a party leader selects the candidates, potentially with the help of local party-branches, immediately after which the party structure occurs and all agents begin collecting their political payoffs. Then, a moderate (extremist) challenger to party leadership arrives with some probability. If there is a challenger, each candidate votes to choose between the incumbent and the challenger. The challenger wins the leadership position if she earns at least half of the candidates' votes. Otherwise, the incumbent maintains her status as the leader. The party competes in the elections with its newly selected leader, after which the term ends.

3.2 Party structure

This section derives the party structures —the distribution of politician types in each party— given party leaders' delegation rates. Each party has measure 1 positions. To fill these positions, the leader of a party may either appoint her

[5] Similar to span-of-control models, we assume that a more capable leader can lead more resources (Akcigit et al. (2018); Lucas Jr (1978); Eeckhout and Kircher (2012)).

Figure 1: Timing of the political game



loyalists or ask for the help of local organizers. The party leader does not have information about the politicians' assets so she randomly samples from the asset distribution of loyal politicians, $L^L(z)$. By the law of large numbers, $L^L(z) > 0$ also corresponds to the asset distribution among the loyalist members of the party.

Party organizers have complete information about the asset distribution of regular politicians at the local level. When party leaders delegate the recruitment of politicians to the local branches, organizers of different parties compete with each other to recruit the most able politicians for the party positions the leaders delegate them to fill. The local party organizers' competition over regular politicians is essentially a many-to-one matching problem where a continuum of politicians match with a finite number of parties. While party organizers aim to recruit the politicians with the highest amounts of resources, politicians' preference-ranking of the parties depend on both the competence and the ideology of the party leader. The stable equilibrium of this matching competition is characterized by cutoff politician types recruited in equilibrium by each party. Specifically, ideology-j members of party k's assets lie between $\underline{z}_k^j$ and $\bar{z}_k^j$, for $j \in \{M, E\}$. Appendix A.1 proves the existence and uniqueness theorems of such an equilibrium by following the exact same steps of Azevedo and Leshno (2016), henceforth AL.[6] The

[6] The intuition underlying the proofs is the Deferred Acceptance (DA) algorithm of Gale and Shapley (1962). DA is essentially a propose-and-reject algorithm. First, politicians submit their ranked preferences for parties. Local party organizers, who observe the types of politicians, rank them in order of priority by the amounts their resources. Then, an iterative process begins. In each round, the politician applies to his most preferred party that has not rejected him yet. A party considers all the proposals it received along with its current match from the previous round (if any). The party then "holds onto" the politicians with the highest amount of resources up to its capacity, and rejects the rest. A politician who gets rejected becomes unmatched and will propose to the next party on their list in the next round.

following proposition characterize the structure and of a party, the proof of which is provided in Apppendix A1.2.

Proposition 1 (Party structure)**.** *The structure of party k who recruits a share ϕ_k of members by organizers, and recruits loyalists for the remaining $1-\phi_k$ of the positions is as follows. The asset distribution of loyal politicians is $(1-\phi_k)\ell^L(z)$ and the asset distribution of non-loyalist politicians is such that their total density in the party allows the organizers to fill the share ϕ of positions*

$$\phi_k = \int_{\underline{z}_k^M}^{\bar{z}^M} \ell^M(s)ds + \int_{\underline{z}_K^E}^{\bar{z}^E} \ell^E(s)ds. \tag{3.1}$$

Let $\psi^j = \frac{\int_{\underline{z}_k^j}^{\bar{z}^j} \ell^j(s)ds}{\phi_k}$ be the share of j-ideology politicians among the regular party members. Then density of a type-z politician in party k is

$$g_k(z) \quad = \quad \begin{cases} (1-\phi)\ell^L(z) \ \ if\ z \notin (min\{\underline{z}_k^M, \underline{z}_k^E\}, max\{\bar{z}_k^M, \bar{z}_k^E\}) \\ \\ (1-\phi)\ell^L(z) + \phi\Big[\psi^M\ell^M(z) + (1-\psi^M)\ell^E(z)\Big] \ \ if\ z \in (\underline{z}_k^j, \bar{z}_k^j), j\in \{M, E\} \\ \\ (1-\phi)\ell^L(z) + \phi\psi^M\ell^j(z) \ \ if\ z \in (\underline{z}_k^j, \bar{z}_k^j), j \in \{M, E\}\ and\ z \notin (\underline{z}_k^{j'}, \bar{z}_k^{j'}), j' \neq j. \end{cases} \tag{3.2}$$

Corollary 1 (Party's assets)**.** *The expected total size of party k who recruits a share ϕ_k of members by organizers, and recruits loyalists for the remaining $1-\phi_k$ of the positions is*

$$x_k = \phi_k z_k^L + (1-\phi_k)z_k^D \tag{3.3}$$

where

$$z_k^L = \int_0^{z_L^{max}} s\ell^L ds \tag{3.4}$$

is the mean asset of a loyal politician, and

$$z_k^D = \psi^M \int_{\underline{z}_k^M}^{\bar{z}^M} s\ell^M(s)ds + (1-\psi^M)\int_{\underline{z}_K^E}^{\bar{z}^E} s\ell^E(s)ds. \tag{3.5}$$

is the mean assets of politicians recruited by the local organizers.

3.3 Leadership contest

This section introduces the leadership competition between a challenger and an incumbent leader. When there is a leadership competition, each party member casts their vote for their preferred leader, and the leader that obtains the most votes wins party leadership. While a loyalist always sides with the incumbent, a non-loyalist chooses the leader that provides the best payoff to him. A leader's winning probability of such a challenge is the expected probability that a member is going to choose her, and it is given by

$$\underbrace{\pi^{jj'}}_{\substack{\text{prob. that an ideology-j}\\ \text{leader wins against}\\ \text{an ideology-}j'\text{ challenger}}} = \underbrace{(1-\phi)}_{\substack{\text{share }(1-\phi)\text{ loyal}\\ \text{politicians choose}\\ \text{incumbent w/prob. 1}}} + \underbrace{\phi}_{\substack{\text{share of politicians}\\ \text{recruited by local}\\ \text{organizers}}} \underbrace{\Omega^{jj'}}_{\substack{\text{prob. that a recruited}\\ \text{politician chooses ideology-j}\\ \text{incumbent over ideology-}j'\\ \text{challenger}}} \tag{3.6}$$

for $j, j' \in \{E, M\}$.[7] In equation 3.6, the probability that an ideology-j incumbent wins against an ideology-j' challenger depends on the incumbent's delegation rate, ϕ. The share $(1-\phi)$ of politicians are loyalists who always choose the incumbent over any challenger. The remaining share ϕ of politicians vote for the leader with probability $\Omega^{jj'}$.

$\Omega^{jj'}$ is the probability that a non-loyalist party member will choose the incumbent over a challenger. It is found by deriving the expected probability that a non-loyalist will choose the incumbent over a challenger. Politicians' preferences stated in Assumption 1 determine how a politician votes for a leader. If both leaders have the same ideology, that is, $j = j'$, then each politician unanimously prefers the leader that has higher competence. Because leaders' competence, η, is

[7] This probability function assumes that a leader expects that a randomly chosen member of the parliament will determine her survival chances.

distributed uniformly, we have that $\Omega^{jj} = \eta$ for a type-(j, η) leader for $j \in \{M, E\}$.

On the other hand, when the incumbent and the challenger have different ideologies, i.e., $j' \neq j$, politicians may have heterogeneous leader preferences depending on their ideological match. Under Assumption 1, a non-loyalist has a trade-off between ideology and competence such that he may prefer a leader with a worse ideological match if she is sufficiently superior in terms of competence. These preferences are assumed to be translated into threshold competence types $\eta^{ME}(\eta)$ and $\eta^{EM}(\eta)$ such that a non-loyalist M-ideology party member will choose an E leader over a type-(M, η) incumbent if the E leader's competence is above $\eta^{ME}(\eta)$. Similarly, an E-member will vote for a type-(M, η) incumbent over an E challenger if the E challenger's competence is less than $\eta^{EM}(\eta)$. Given that M members' share in the party is ψ^M, we can summarize the probability that a non-loyalist will choose the incumbent over a challenger as follows

$$\Omega^{jj'}(\eta) = \begin{cases} \eta & j \in \{M, E\}, j = j' \\ \underbrace{\psi^j}_{\substack{\text{share of} \\ \text{j-ideology} \\ \text{members}}} \underbrace{\eta^{jj'}(\eta)}_{\substack{\text{j-members} \\ \text{prefer the j-leader if} \\ \text{j'-leader's competence} \\ \text{is less than } \eta^{jj'}(\eta)}} + \underbrace{(1 - \psi^j)}_{\substack{\text{share of} \\ \text{j'-ideology} \\ \text{members}}} \underbrace{\eta^{j'j'}(\eta)}_{\substack{\text{j'-members} \\ \text{prefer the j-leader if} \\ \text{j'-leader's competence} \\ \text{is less than } \eta^{j'j'}(\eta)}} & j \in \{M, E\}, j \neq j' \end{cases} \tag{3.7}$$

3.4 Value function of a party leader

This section presents the value function of a leader. Note that because the delegation decision and candidate recruitment takes place at the beginning of an election term, the leader's value function stays constant throughout the election cycle, as long as she stays at the helm of the party. As a result, we drop time indicators from the value functions. The value of an ideology-j incumbent leader, $V^{incumbent,j}$, is

$$\underbrace{V^{incumbent,j}}_{\substack{\text{value of}\\ \text{leader}}} = \tau \underbrace{f_k(x_k;\eta)}_{\substack{\text{flow}\\ \text{payoff}}} + \tau \underbrace{\frac{1}{1+\rho}}_{\text{discounter}} \Bigg\{ \underbrace{(1 - \sum_{j' \in \{M,E\}} \alpha^{j'})}_{\text{no challenge}} \underbrace{V^{incumbent,j}}_{\substack{\text{value of}\\ \text{leader}}}$$

$$+ \sum_{j \in \{M,E\}} \underbrace{\alpha^{j'}}_{\substack{\text{j'-ideology}\\ \text{challenger}}} \underbrace{V^{challenge,j'}}_{\substack{\text{value of incumbent}\\ \text{when challenged by}\\ \text{an ideology-}j\text{ leader}}} + o(\tau) \Bigg\}. \tag{3.8}$$

Reading from left to right, an ideology-j incumbent leader's lifetime value from party leadership, $V^{incumbent}$, is the sum of a flow payoff and a continuation value that she receives for an infinitesimally small time period τ, plus a term $o(\tau)$. The flow payoff, $f_k(x_k;\eta,j)$, consists of the party's electoral rents introduced in Assumption 2. The continuation value, which the leader discounts at rate ρ, weights the expected values of three mutually exclusive events: not having any challengers; having a challenger with ideology M, which occurs at rate α^M; and having a challenger with ideology E, which occurs at rate α^E.[8] When there is no challenger, the leader receives the value of leadership for the next period, $V^{incumbent,j}$. When a challenger of ideology j', for $j' \in \{M,E\}$, arrives, the expected value of the leader is

$$\underbrace{V^{challenge^j}}_{\substack{\text{value of incumbent}\\ \text{when challenged by}\\ \text{an ideology-}j\text{ leader}}} = \underbrace{\pi^{jj'}}_{\substack{\text{incumbent wins}\\ \text{the challenge}}} \underbrace{V^{incumbent,j}}_{\substack{\text{value of}\\ \text{leader}}} + \underbrace{(1-\pi^{jj'})}_{\substack{\text{incumbent loses}\\ \text{the challenge}}} \underbrace{V^{exit}}_{0}. \tag{3.9}$$

Reading from left to right, the value of an incumbent leader with an ideology-j challenger, $V_k^{challenge,j}$, weights the expected values of winning and losing the challenge. If the incumbent wins the challenge, which occurs with probability $\pi^{jj'}$, she continues to receive the leadership value, $V_k^{incumbent,j}$. If she loses leadership, then she exits politics and receives the exit value, V_k^{exit}, which is normalized to zero.

Finally, imposing equation 3.9 into equation 3.8, and taking the limits, we

[8] New leader arrival rates, α^M and α^E, are allowed to be different from the distribution of leader ideologies, p^M and $1-p^M$, respectively.

obtain

$$V^{incumbent,j} = \frac{(1+\rho) f_k(x_k;\eta)}{1+\rho-(1-\alpha_M-\alpha_E)-(\alpha^M\pi^M+\alpha^E\pi^E)}. \tag{3.10}$$

3.5 Equilibrium analysis

After having characterized a party's total expected assets and a leader's survival chances, we can now work our way through the leader's value function to obtain the optimal delegation rule, ϕ^*. The latter is found by maximizing a leader's lifetime value function (described in equation 3.10) given her probability of remaining the leader (characterized in equations 3.3 to 3.5) and party's size (characterized in equations 3.6 and 3.7), which, in turn, depend on the extent of delegation.

We begin by providing the definition of a stable equilibrium, then introduce the existence theorem.

Definition 1 (A stable equilibrium)**.** A stable equilibrium with cutoffs is a collection of a) a set of equilibrium delegation rates ϕ_k, and b) a set of party structures $\Theta_k = [z, g_k(z_k)]\ for\ \forall k \in K$ such that the following statements hold:

i) Given the vector of delegation rates $\phi_k, \forall k \in K$, each party's optimal party structure maximizes each leader's value function in equation 3.10.

ii) Each party's optimal choice of party structure induces the vector of delegation rates $\phi_k, \forall k \in K$.

Theorem 1 (Existence of a stable equilibrium)**.** *A stable equilibrium exists.*

The proof in Appendix A shows that the value function satisfies the necessary properties for a stable fixed point to exist. Intuitively, an equilibrium exists because a leader's value function has two components with different returns to delegation rate. The optimal delegation rate balances a leader's prospects of staying at the helm of her party with party growth. The following theorem formally characterizes the optimal delegation rule.

Theorem 2 (Optimal delegation rule)**.** *The type-(j,η) leader of party k chooses to delegate a share of positions equal to the following expression*

$$\phi_k^* = \begin{cases} -\frac{z_k^L}{z_k^D - z_k^L + \eta(z_k^L + \Delta_k)} - \frac{\rho\eta(z_k^L + \Delta_k)}{(\alpha^j(1-\eta) + \alpha^{j'}(1-\Omega^{jj'}))(z_k^D - z_k^L + \eta(z_k^L + \Delta_k))} & \textit{if } \phi \in (0,1) \\ 0 & \textit{if } \phi \leq 0 \\ 1 & \textit{if } \phi \geq 1 \end{cases} \tag{3.11}$$

The expression of the optimal rule captures the trade-off we have described in the introduction: higher levels of delegation increase the party size and, in turn, the leader's rents due to more non-loyalists, but, at the same time, increase the leader's own survival prospects (captured by the denominator in the right fraction). Thus, changes in any of the two due to exogenous shocks will vary with the optimal delegation rule.

The next proposition characterizes the properties of the optimal delegation rule, and the following corollaries present how optimal delegation rule changes with exogeneous changes in the economy.

Proposition 2 (Fragility of a liberal democracy)**.** *The equilibrium delegation rule $\phi^*(\cdot)$ is increasing in leaders' competence, η. Moreover, keeping competence constant, M-ideology leaders have a higher delegation rate than E-ideology leaders. This has two implications: i) Less competent, more extremist leaders have a longer expected tenure as a party leader, and ii) extremist and less competent leaders select lower-quality politicians, so their parties shrink.*

Proposition 2 characterizes a pattern we call, the 'slippery slope' of democratic, intraparty politics. The less competent and the more extremist leaders tend to lose democratic leadership competitions because they cannot provide a high party membership value to politicians. To prevent the loss of their position, such leaders tend to fill the ranks in their parties with loyalists. Loyalists typically have fewer political resources than the politicians who are recruited by local party organizers. At the same time, loyalists always side with the incumbent leader when a challenger arrives. As a result, less competent and more extremist party leaders tend to stay at the helm of their parties for a longer time than more moderate and more competent leaders, despite these more extremist leaders causing their parties to shrink under their leadership.

Moreover, ideology critically exacerbates this result. Moderate (and competent) leaders endogenously place higher instrumental value on market-based recruitment because they can 'kill two birds with one stone'. As their brand –their (γ, η)-type– is more attractive to ordinary politicians, they can get better matches: not only they can recruit the more competent candidates, but at the same time, they can also fill the party ranks with ideological lookalikes (moderates). Thus, the matching market is relatively more lucrative for them which, in turn, incentivizes them to assume relatively higher risks (i.e. higher replacement probability) and delegate more.

Corollary 2 (Political talent pool)**.** *A party leader's optimal rate of delegation decreases (increases) if the mean assets of loyal politicians (non-loyal members recruited by the party) increases.*

Corollary 3 (Introduction of new parties)**.** *The introduction of a new party influences an existing party leader's delegation rule only if the new party changes the ordinal preference ranking of the incumbent party for a politician of any ideology type. If this ranking is altered, the incumbent party's total assets decreases. However, the distribution of ideologies among the non-loyalist members also changes, which can either increase or decrease the leadership survival probability. Consequently, the overall effect on the optimal delegation rule is ambiguous.*

3.6 Remarks on welfare and efficiency

Next, we comment on welfare analysis. Given that in our set-up voters are introduced last, and mostly serve as mechanical (non-strategic) agents, it is safe to assume that maximizing social welfare is tantamount to having a benevolent 'impartial spectator' who chooses ϕ in order to maximize the overall quality of all those who become politicians. That is:

$$\mathcal{W}(\phi) = \sum_{k=1}^{K} \mathbb{E}[x_k(\phi_k)|\Phi] \tag{3.12}$$

where $\Phi = [\phi_1, \phi_2, ..., \phi_K]$ and x_k is defined in equation 3.3.

In other words, it is as if we are assuming that voters' well-being increases in the average quality of politicians recruited across all parties. Then, we can state the following observation.

Theorem 3 (Welfare analysis)**.** *Voters' (social) welfare increases in leaders' delegation rate ϕ, as long as the political assets of the last recruited non-loyalist politician are greater than $\mathbb{E}[\ell^L]$. This implies that welfare benefits from the superior information of local party organizers would be limited if: a) the difference between the average assets of non-loyalist and loyalist politicians is small and, b) the number of parties is large.*

To see the intuition behind this statement, consider first what happens to overall welfare (i.e., the quality of recruited politicians) if we were to take a partial equilibrium approach and study each party in isolation, ceteris paribus.[9] Then, by applying a logic similar to the envelope theorem (i.e. assume that a $\partial \phi_k^* / \partial \phi_{j-k} = 0$), it is straightforward to see that the socially optimal level of delegation is $\phi_k^O = 1$. The latter follows directly from the fact that $z_k^D > z_k^L$. But when there are many parties, politicians who aspire to be selected as candidates also rank parties in terms of their preference to join them.

Now, consider that moderate politicians (whose talent pool is drawn from $L^M(z)$) have a certain preference ranking of parties, and similarly, extreme politicians (who are drawn from $L^E(z)$) have a different ranking of their own. Because these rankings are constant, given a politician's ideology, parties recruit continuously from the said distributions. That is, M politicians' first-ranked party gets the top end (right tail) of the $L^M(z)$ distribution, and so on. An analogous argument applies for E politicians. Hence, given that we have two *separate* distributions of non-loyal politicians, for a given ideology parties recruit *continuously* from them. This, in turn, implies that for some party k when the level of delegation is greater than a certain threshold $\bar{\phi}_k < 1$, the last recruited non-loyal (M or E) politician will have (in expectation) a level of political assets/skill lower than the expected value (λ^L) of the average loyal candidate. In other words, when there is excessive demand for politicians (due to a large number of parties who delegate a lot), so-

[9] Notice that ϕ_k^* is determined in equilibrium; that is, it also depends on the choices of the remaining $k-1$ actors.

ciety is better of the least desired (by politicians) parties recruit their candidates from the pool of loyalists at their leaders' discretion instead of participating in the 'market for political talent'. Thus, full delegation and complete decentralization of the selection process is generically not socially optimal.

There are two interesting things to point out from this. First, while in principle more delegation allows parties (and society) to harness all available information regarding overall candidate quality, which strictly improves social welfare, fully resolving the issue of asymmetric information can occasionally lead to lower welfare. This is due to the nature of the two-sided matching process: once the most appealing parties select from the upper tail of the talent pool, less appealing parties can potentially be better off if they withdraw from the competitive 'candidates' market' and simply recruit the best among the pool of loyalists. Thus, harnessing the information that the party middlemen posses is useless and potentially welfare decreasing. This is a novel insight that contrasts findings in other markets (e.g., insurance) where resolving information asymmetries results always in welfare improvements.

Second, and equally unexpected, the above negative effect of acquiring more information on welfare is amplified as electoral competition becomes more intense (i.e. when there are more parties competing to recruit politicians in the local markets). This might sound paradoxical at first, but given the nature of political competition, where parties enjoy a quasi-monopoly on ideology, it makes sense from the perspective of aggregate welfare maximization that the least appealing parties withdraw from the recruiting market when they face tougher competition.

Finally, it is worth noting that the above result of the potential suboptimality of full delegation is further strengthened if one assumes that, in addition to maximizing politicians' assets, voters also care about (intra-party) political stability (fewer leadership contests) which, in turn, is strictly increasing in the share of loyal politicians recruited by parties. Thus, ceteris paribus, a society that is even weakly averse to costly political conflict should be strictly better-off with less than full intra-party democracy (decentralization). The latter observation can also partly explain the variation in intraparty constitutions and processes that we observe around the world.

4 Discussion and Conclusion

Our work brings to surface two points that may have broader implications that extend beyond intraparty politics and selection. Clearly, the ideology-competence trade-off we have identified provides a strategic link between authoritarianism, or a leader's desire for more control, and ideological extremism. This, in turn, affects the quality of political selection and democratic representation (see e.g. Dal Bó et al. (2017); Besley (2005); Besley et al. (2011)).[10] But this 'fragility of democracy' may have broader implications as well in firms, bureaucracies, and organizations with multiple aims. A potential application is in span-of-control literature, where a senior manager's or a CEO's priorities such as corporate responsibility, environmental protection, or combating sexual harassment and discrimination can affect the quality of personnel recruitment and the firm's overall output.

Moreover, our work makes a contribution to modelling a two-sided political labor market. Our model's main results straightforwardly extend to cases where the organization has multiple dimensions of ideological or identity concerns. Suppose that the vector of leader ideologies γ is instead a vector $(\gamma_1, \gamma_2, ..., \gamma_N)$ where in addition to left-right ideology one cares about liberal rights (e.g., abortion), gender/racial discrimination, ethnicity, religion, environmental degradation, and other elements of identity. Our model can identify the exchange rates between ability and these other dimensions. Moreover, it can lend itself to structural estimation which will allow us to endogenously determine the relative importance of each of these γ-dimensions on organizational decisions and structure.

On a more normative note, our work also offers insights with respect to the so-called 'slippery slope' of decentralized (democratic) decision-making processes. For instance, our findings highlight an apparent tension between higher levels of delegating or outsourcing key decisions downstream while, at the same time, retaining competent leadership at the helm –a trade-off that is directly linked to organizational success. The latter has clear implications for optimal constitutional design

[10]This conclusion also echoes the well-known Fearon (1999)-Besley (2005) critique when there is significant agent (politician or candidate) heterogeneity and principals (voters) have non-common values. In our context as well, because some leaders if given authority over selection cannot credibly commit to not choosing a loyal candidate over one that generates more electoral rents, delegation can be welfare improving.

and/or organizational structure regarding the desired degree of (de)centralization in decision-making. Put simply, in most set-ups characterized by significant agent and institutional heterogeneity, the equilibrium level of decentralization will generically be different from the one that maximizes organizational success. The latter will only be attained under a very strict set of conditions, if at all. Contrast this with the institutional arrangement of full decentralization under any conditions preferred by ordinary organization members (i.e., those members with no career concerns). Thus, our findings highlight the apparent tension between career concerns (at any level of management) and optimal constitutional design in heterogeneous organizations. They also point to the fact that this slippery slope is driven by the career concerns of leaders and politicians alike. Hence, optimal constitutional designs should take these concerns and the forces they generate into account.

References

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding "choice" in school choice," American Economic Journal: Microeconomics, 7, 1–42.

Acemoglu, D. and J. A. Robinson (2020): The narrow corridor: States, societies, and the fate of liberty, Penguin.

Aghion, P. and J. Tirole (1997): "Formal and real authority in organizations," Journal of political economy, 105, 1–29.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018): "Dancing with the stars: Innovation through interactions," Tech. rep., National Bureau of Economic Research.

Azevedo, E. M. and J. D. Leshno (2016): "A supply and demand framework for two-sided matching markets," Journal of Political Economy, 124, 1235–1268.

Bandiera, O., M. C. Best, A. Q. Khan, and A. Prat (2021): "The allocation of authority in organizations: A field experiment with bureaucrats," The Quarterly Journal of Economics, 136, 2195–2242.

Barbera, S. and M. O. Jackson (2004): "Choosing how to choose: Self-stable majority rules and constitutions," The Quarterly Journal of Economics, 119, 1011–1048.

Besley, T. (2004): "Paying Politicians: Theory and Evidence," Journal of the European Economic Association, 2, 193–215.

——— (2005): "Political Selection," Journal of Economic Perspectives, 19, 43–60.

Besley, T. and S. Coate (1997): "An Economic Model of Representative Democracy," Quarterly Journal of Economics, 112, 85–114.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011): "Do Educated Leaders Matter?" Economic Journal, 121, 205–227.

Buisseret, P., O. Folke, C. Prato, and J. Rickne (2022): "Party nomination strategies in list proportional representation systems," American Journal of Political Science, 66, 714–729.

Cakir, S. (2019): "Choosing Your Pond: A Structural Model of Political Power Sharing," Available at SSRN 3467761.

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

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

Chattopadhyay, R. and E. Duflo (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," Econometrica, 72, 1409–1443.

Clots-Figueras, I. (2012): "Are Female Leaders Good for Education? Evidence from India," American Economic Journal: Applied Economics, 4, 212–244.

Dal Bó, E., F. Finan, N. Y. Li, and L. Schechter (2021): "Information technology and government decentralization: experimental evidence from Paraguay," Econometrica, 89, 677–701.

Dal Bó, E. and F. Finan (2018): "Progress and Perspectives in the Study of Political Selection," Annual Review of Economics, 10, 541–575.

Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017): "Who Becomes A Politician?" Quarterly Journal of Economics, 132, 1877–1914.

Dewan, T., A. Galeotti, C. Ghiglino, and F. Squintani (2015): "Information aggregation and optimal structure of the executive," American Journal of Political Science, 59, 475–494.

Eeckhout, J. and P. Kircher (2012): "Assortative matching with large firms: Span of control over more versus better workers," Universitat Pompeu Fabra (Mimeo).

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

Funke, M., M. Schularick, and C. Trebesch (2020): "Populist leaders and the economy," .

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

Hyytinen, A., J. Meriläinen, T. Saarimaa, O. Toivanen, and J. Tukiainen (2018): "Public Employees as Politicians: Evidence from Close Elections," American Political Science Review, 112, 68–81.

Invernizzi, G. M. and C. Prato (2019): "Power Sharing, Mobilization, and Party Organization," Mobilization, and Party Organization (September 30, 2019).

Kirkland, P. A. (2020): "Business Owners and Executives as Politicians: The Effect on Public Policy," Journal of Politics, forthcoming.

Lee, D. S., E. Moretti, and M. J. Butler (2004): "Do Voters Affect or Elect Policies? Evidence from the U. S. House," Quarterly Journal of Economics, 119, 807–859.

Lucas Jr, R. E. (1978): "On the size distribution of business firms," The Bell Journal of Economics, 508–523.

Matakos, K., R. Savolainen, O. Troumpounis, J. Tukiainen, and D. Xefteris (2018): "Electoral Institutions and Intraparty Cohesion," SSRN Scholarly Paper ID 3212638, Social Science Research Network, Rochester, NY.

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

——— (2015): "Mediocracy," Journal of Public Economics, 130, 32–44.

Meriläinen, J. (2022): "Political Selection and Economic Policy," Economic Journal, forthcoming.

Osborne, M. J. and A. Slivinski (1996): "A Model of Political Competition with Citizen-Candidates," Quarterly Journal of Economics, 111, 65–96.

Pande, R. (2003): "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India," American Economic Review, 93, 1132–1151.

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

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

A Proofs

This appendix contains the mathematical proofs of the theorems and propositions presented in the main text. It primarily focuses on the existence and uniqueness of a stable two-sided matching equilibrium between politicians and parties, as well as additional proofs concerning the party structure, optimal delegation rules, and welfare implications. These results leverage standard matching theory and extend classical results to the context of political recruitment and party competition.

A.1 Two-sided matching equilibrium

We begin by defining the stable matching equilibrium in the local market. In this setting, non-loyalist politicians are matched with parties, and the equilibrium is considered stable if no party and politician have an incentive to form a new match that would make both better off. After presenting the existence and uniqueness theorems for such an equilibrium, we provide the proofs in sections A.1.1 and A.1.2, respectively.

Definition A.1 (Stable equilibrium of matching between regular politicians and parties)**.** A matching between non-loyal politicians and parties is pairwise stable if no pair of a politician and a party prefer to break away from a match partner and match to each other.

Theorem A.1 (Existence)**.** *Given the delegation rates of each of K party's leaders, $\phi_1, \phi_2, ..., \phi_K$, there exists a stable equilibrium of the matching between parties and politicians recruited by the local party organizers. This equilibrium is characterized by cutoff politician types recruited by each party.*

Theorem A.2 (Uniqueness)**.** *Given the delegation rates of each of K party's leaders, $\phi_1, \phi_2, ..., \phi_K$, the stable equilibrium with cutoff politician types recruited in each party is the unique equilibrium of the matching between local organizers and politicians. Specifically, local organizers of party k recruit j-ideology politicians whose assets are between $(\underline{z}_k^j, \bar{z}_k^j)$, for $j \in \{M, E\}$, $\forall k$.*

A.1.1 Existence proof

In this section we prove Theorem A.1., the existence of a stable equilibrium of the matching between parties and non-loyalist politicians following the exact same steps of Azevedo and Leshno 2016 (henceforth AL).[11] Note that our model differs from AL by explicitly modeling the multi-layered heterogeneity in politicians' party preferences. Nevertheless, we show that politicians' 'demand' for parties still satisfies the key properties for a fixed point in 'party-membership cutoffs' (market-clearing cutoffs) to exist. Following AL, we begin by providing a definition for politicians' demand for a party, then discuss its properties. Next, we introduce a mapping function for party membership cutoffs , and show that it satisfies the necessary properties for applying Tarski's fixed point theorem, which establishes the existence of a stable equilibrium.

Definition A.2 (Party demand)**.** A politician's demand for a party is defined as his highest-preferred party that accepts him. Let K be the total set of parties. The aggregate demand for a set of political parties $K' \subseteq$ K is defined as

$$\sum_{k \in K'} D_k(z) \tag{A.1}$$

where z_k is the "cutoff" for joining party k; that is, the minimum amount of assets a politicians needs to "pay" to the party to affiliate with the party.

Remark A.1 (Monotonicity, gross substitutability, and continuity of demand)**.** *Aggregate demand for any subset of parties $K' \subseteq K$, given in A.1, is 1) monotone nonincreasing in z_k for all $k' \in K'$, 2) monotone nondecreasing in $z_{k''}$ for any $k'' \in K \setminus K'$ so it satisfies the gross substitutes property, and 3) continuous in cutoff politician types.*[12]

[11] Our paper does not have any contributions to the proofs of the existence and the uniqueness of a two-sided stable matching equilibrium between non-loyalist politicians and parties. One can instead read Appendices A and B of AL to read the original proofs written for a more general continuum-to-discrete matching problem. We only adjust the notation of AL to our problem to facilitate following the proofs for our example. Moreover, using the structure we impose on the types and preferences of political agents, Proposition 1 derives a more detailed characterization of party structures in a stable equilibrium.

[12] AL do not impose structure on students' preferences and assumes that their demand for schools is continuous in cutoff student types. In our model, we explicitly lay out politicians'

Definition A.3 (Contraction mapping function)**.** Given z_{-k}, let $I_k(z_{-k})$ be the interval of possible cutoffs for party k that clear the market for k

$$I_k(z_{-k}) = \{p \in [0,1] : D_k(z_k, z_{-k}) \leq \phi_k \text{ and } D_k(z_k, z_{-k}) = \phi_k \text{ if } z_k > 0\}. \tag{A.2}$$

Define $T(z)$ as the function that prompts party k to adjust its cutoff to the nearest point within $I_k(z_{-k})$ relative to z_k

$$T_k(z) = argmin_{z \in I_k(z_{-k})} |z - z_k|. \tag{A.3}$$

Proposition A.1. *The map T is monotone non-decreasing with respect to the partial order on the interval $[0,1]^K$. The set of fixed points of T corresponds to the set of market clearing cutoffs.*

Proof. Because D is continuous and $D_k(1, z_{-k}) = 0$, we have either $0 \in I_k(z_{-k})$ or there exists $z_k \in [0,1]$ such that $D_k(z_k, z_{-k}) = \phi_k$. Note that $I_k(z_{-k})$ is nonempty in either case. Moreover, monotonicity and continuity of demand makes $I_k(z_{-k})$ a compact interval. This establishes that T is well-defined.

To show the monotonicity of T, consider two values $z_k \leq z'_k$ with corresponding $t_k = T_k(z_k)$, and $t'_k = T_k(z'_k)$. For a contradiction, assume that $t'_k < t_k$. Given $t_k > 0$ and using the monotonicity and gross substitutes properties of demand, we have

$$\phi_k = D_k(t_k, z_{-k}) \leq D_k(t'_k, z_{-k}) \leq D_k(t'_k, z'_{-k}) \leq \phi_k \tag{A.4}$$

and

$$\phi_k = D_k(t_k, z_{-k}) \leq D_k(t_k, z'_{-k}) \leq D_k(t'_k, z'_{-k}) \leq \phi_k \tag{A.5}$$

Equations A.4 and A.5 imply that $D_k(t_k, z'_{-k}) = D_k(t'_k, z_{-k})$. Thus, $|t'_k, t_k| \subseteq I_k(z_{-k}) \cap I_k(z'_{-k})$. Because t_k is the closest point to z_k in $I_k(z_{-k})$, we have that $z_k \geq t_k$. Therefore, $z'_k \geq t_k$. So, we have $|t_k - z'_k| < |t'_k - z'_k|$, which contradicts the

preferences for parties. Because all politicians have strict preferences over parties, the demand for a party is continuous in "party membership cutoff" even though politicians with different ideologies differ in their preference-rank of parties.

assumption that $z'_k = T_k(z')$. This contradiction establishes that T is monotone.

Finally, note that z^* is a fixed point of T if and only if each $z^*_k \in I_k(z^*_{-k})$ for all k, which is equivalent to z^* being a market clearing cutoff, which establishes that the set of fixed points of T coincides with the market clearing cutoffs. $\square$

Corollary A.1. *At least one stable matching exists.*

Proposition A1 and Tarski's Theorem establishes Corollary 1. That is, at least one stable matching exists, and stable matchings are the fixed points of T.

A.1.2 Uniqueness proof

This section establishes that the stable matching between parties and non-loyalist politicians is unique (Theorem A.2) by following the exact same steps outlined by Azevedo and Leshno in 2016 (hereafter referred to as AL). The proof is structured into four main steps. We first demonstrate that the set of market clearing cutoffs forms a complete lattice. This structure ensures that any subset of market clearing cutoffs has both a supremum and an infimum within the lattice, facilitating the analysis of extremal conditions. Next, we establish the existence of the smallest and largest market clearing cutoffs. This step allows us to define bounds within which all other market clearing cutoffs must lie. We then show how AL extends the rural hospitals theorem, a classic result from matching theory, to the continuum setting. This theorem asserts that if a party fails to fill its quota in one stable matching, it will fail to do so in all stable matchings. It also confirms that the total measure of unmatched politicians remains constant across all stable matchings. Additionally, the demand for a subset of parties remains invariant under both the largest and smallest market clearing cutoffs. Finally, we show that the smallest and largest market clearing cutoffs are identical when they have full support, proving the uniqueness of the stable equilibrium in the model.

Given a set of cutoffs $X \subseteq 2^{([0,1]^K)}$, define

$$(\sup X)_k = \sup_{z \in X} z_k$$

and

$$(\inf X)_k = \inf_{z \in X} z_k,$$

as lattice operators on cutoffs.

Theorem A.3 (Lattice theorem)**.** *The set of market clearing cutoffs is a complete lattice.*

The proof of Theorem A.3 follows from Tarski's theorem and proposition A.1, which together imply that the set of market clearing cutoffs is a lattice.

Corollary A.2. *There exists smallest and greatest market-clearing cutoffs.*

Proposition A.2. *(Politician- and party- optimal cutoff adjustment algorithms) The limits*

$$\begin{aligned} z^- &= lim_{k \to \infty} T^i(0) \\ z^+ &= lim_{k \to \infty} T^i(1) \end{aligned}$$

exist, and equal the smallest and largest market clearing cutoffs.

Proof. First, consider the iterative application of the map T starting from the initial vector of zeros. Since $0 \leq T(0)$, the sequence generated by successive iterations of T satisfies the inequality $T^i(0) \leq T^{i+1}(0)$, meaning that the sequence $T^i(0)$ is monotone non-decreasing. By the monotonicity and bounded nature of sequence, the limit z^- exists. The continuity of demand ensures that z^- is indeed a market clearing cutoff. Specifically, for any i and with $z^i = T^i(0)$, the following inequality holds

$$D(z_k^{i+1}, z_{-k}^i) \leq \phi_k,$$

with equality if $z_k^i > 0$. Taking the limit, we have

$$D_k(z^-) \leq \phi_k$$

with equality if $z_k^- > 0$. This confirms that z^- is a market clearing cutoff.

If z^* is any other market clearing cutoff, then by definition $0 \leq z^*$. Therefore, applying the iterative process to z^*, we have

$$T^i(0) \leq T^i(z^*) = z^*.$$

Taking the limit as $k \to \infty$, we get $z^- \leq z^*$. The proof for z^+ is analogous. $\square$

Next, we present AL's proof of how the rural hospitals theorem of classic matching theory extends to the continuum setting. This result implies that a party that does not fill its quota in one stable matching does not fill its quota in any other stable matching. Moreover, the measure of unmatched politicians is the same in every stable matching.

Theorem A.4 (Rural hospitals theorem)**.** *The measure of politicians matched to each party is the same in any stable matching. Furthermore, if a party does not fill its capacity, it is matched to the same set of politicians in every stable matching, except for a set of politicians with measure 0.*

Part 1: in any stable matching, the same measure of politicians matches to each party. Consider two market-clearing cutoffs z and z', and let $z^+ = zVz'$. Take a party k, and assume without loss of generality that $z_k \leq z'_k$. Due to the gross substitutes property, we must have that $D_k(z^+) \geq D_k(z')$, as $z^+_k = z'_k$ and the cutoffs of other parties are higher under z^+. In addition, if $z'_k > 0$, then $D_k(z^+) = \phi_k \geq D_k(z)$. Furthermore, if $z'_k = 0$, then $z_k = z'_k$, and $D_k(z^+) \geq D_k(z)$. Therefore, in either case, we can conclude

$$D_k(z^+) \geq max\{D_k(z), D_k(z')\}.$$

Additionally, the demand for being unmatched $1 - \sum_{k \in K} D_k(\cdot)$ must be at least as large under z^+ than under z or z'. Since the total demand for being unmatched and for all parties always sums to 1, it follows that for every party $D_k(z^+) = D_k(z) = D_k(z')$. $\square$

Part 2: a party with unfilled capacity matches the same set of politicians in any stable matching. Consider two stable matchings μ and μ' with corresponding market-clearing cutoffs $z = z_\mu$ and $z' = z_{\mu'.}$ Let $z^+ = z \vee z'$ and define $\mu^+ = \mu(z^+)$. Now, consider a party k such that $\int_{z \in \mu(k)} z dz < \phi_k$. Consequently, we have $0 = z_k = z'_k = max\{z_k, z'_k\} = z^+_k = 0$. By the definition of demand, it follows that $\mu(k) \subseteq \mu^+(k)$ and $\mu'(k) \subseteq \mu^+(k)$. From the first part of the theorem, we know that the measures of $\mu(k), \mu'(k)$, and $\mu^+(x)$ are equal, completing the proof.

□

Having shown the existence of the smallest and the largest market clearing cutoffs, and having extended the rural hospitals theorem to the continuum case, we now proceed to show that the demand for a specific subset of parties remains the same under both the largest and smallest market-clearing cutoffs, which follows directly from the rural hospitals theorem. Finally, we show how this implies that the smallest and largest market-clearing cutoffs are identical (Abdulkadiroğlu et al. (2015), Azevedo and Leshno (2016)).

Denote the excess demand given a vector of cutoffs z by

$$D^{excess}(z) = D(z) - \phi$$

Proposition 1 (Uniqueness of the market clearing cutoffs)**.** *Under the full support assumption, there exists a unique vector of market clearing cutoffs.*

Proof. Let $z^- \leq z^+$ represent the smallest and greatest market-clearing cutoffs, with corresponding stable matchings μ^-, μ^+. Define $K^+ = \{k \in K : z^+_k \neq z^-_k\}$. Notably, for all parties in K^+, we have $z^+_k > 0$. Let $K^0 = K \setminus K^+$. Let $\theta = (z, j)$ denote a student's type. Given that for all parties $k \in K^0$, we have $z^+_k = z^-_k$, and for all parties k in K^+ we have $z^+_k > z^-_k$, we can infer that

$$\{\theta \in \Theta : \mu^+(\theta) \in C^+\} \subseteq \{\theta \in \Theta : \mu^-(\theta) \in C^+\}.$$

The rural hospitals theorem implies that the difference between these two sets has measure 0.

Let $\succ^+$ be a fixed preference relation that ranks all parties in K^+ more highly than those in K^0. Consequently, the set described must include all politicians with preference $\succ^+$ and resources $z^-_k \leq z < z^+_k$ for all $k \in K^+$. That is

$$\{(\succ^+, z) \in \Theta : z_k^- \leq z < z_k^+ \forall k \in K^+\} \subseteq \{\theta \in \Theta : \mu^-(\theta) \in K^+\} \setminus \{\theta \in \Theta : \mu^+(\theta) \in K^+\}$$

Thus the former set has measure 0 :

$$\{(\succ^+, z) \in \Theta : z_k^- \leq z < z_k^+ \forall k \in K^+\} = 0.$$

Given full support assumption and since $z_k^- < z_k^+$ for all k in K^+, this can occur if K^+ is the empty set. This implies that $z^- = z^+$, establishing that there exists a vector of market clearing cutoffs that is unique. □

Finally, we are ready to conclude the proof of Theorem 2 in the main text with the following two claims. The first claim asserts that in an economy with multiple stable matchings, either the demand function is not differentiable, or the derivative matrix of the demand function is not invertible. This claim is supported by the observation that demand remains unchanged for cutoffs between the smallest and largest market-clearing cutoffs. This constancy is a result of the monotonicity of demand and the rural hospitals theorem.

Claim 1. *If there is more than one stable matching, then there exists at least one market clearing cutoff z^* such that either demand is not differentiable, or the derivative matrix $\partial D(z^*|E)$ of the demand function is singular.*

Proof. If there is at least one market-clearing cutoff where the demand function is not differentiable, the claim is proven. Now consider the case where the demand function is differentiable at all market-clearing cutoffs. By the lattice theorem, there exists smallest and largest market-clearing cutoffs, denoted z^- and z^+ respectively, with $z^- \leq z^+$. Define

$$K^+ = \{k : z_k^- < z_k^+\}.$$

The set K^+ is nonempty, given that there exists more than one market clearing cutoff. Let F be the subspace of R^K where all coordinates corresponding to parties not in K^+ are zero, i.e.,

$$F = \{v \in R^K : v_k = 0, \forall k \notin K^+\}.$$

Consider $z \in [z^-, z^+]$. For any party $k \notin K^+$ we have $z_k^+ = z_k^- = z_k$. Thus, by the gross substitutes property,

$$D_k(z^+) \geq D_k(z) \geq z_k(z^-).$$

The rural hospitals theorem implies that $D_k(z^-) = D_k(z^+)$. Therefore, $D_k(\cdot)$ is constant within the interval $[z^-, z^+]$. In particular, for any $k \notin K^+$ and $k' \in K^+$ we have

$$\partial_{k'} D_k(z^-) = 0,$$

which means that the the derivative matrix ∂D takes the subspace F into itself.

Moreover, for all $z \in [z^-, z^+]$, by the monotonicity of demand, we have

$$\sum_{k \in K^+} D_k(z^-) \geq \sum_{k \in K^+} D_k(z) \geq \sum_{k \in K^+} D_k(z^+).$$

The rural hospitals theorem implies that $D_k(z^-) = D_k(z^+)$ for all $k \in K^+$. Thus, $\sum_{k \in K^+} D_k(z)$ is constant within the interval $[z^-, z^+]$. This implies that

$$\sum_{k \in K^+} \partial_{k'} D_k(z^-) = 0$$

for all $k, k' \in K^+$. Consequently, the linear transformation $\partial D(z^-)$ restricted to the subspace F is not invertible. Because $\partial D(z^-)$ maps F into itself, $\partial D(z^-)$ is not invertible, proving the claim. Furthermore, according to the Sard's Theorem, for almost all vectors K, the demand function is continuously differentiable with a nonsingular derivative at all market clearing cutoffs.

□

Claim 2. *For almost every $\phi \in R_+^K$ with $\sum_k \phi_k < 1$, at every market clearing cutoff z^*, demand is continuously differentiable and the derivative matrix $\partial D(z^*)$*

is invertible.

Proof. The assumption that $\sum_k \phi_k < 1$ implies that all market clearing cutoffs z^* satisfy $D(z^*) = \phi$. Moreover, all market clearing cutoffs are strictly greater than 0. The assumption that the supply of all parties is strictly positive implies that market clearing cutoffs are strictly lower than 1. So, any vector of market clearing cutoffs lies in the open set $(0,1)^K$.

Define the closure of the set of points where the demand function is not differentiable as

$$NDP = closure\Big(z \in (0,1)^K : D(\cdot) \text{ is not continuously differentiable at } z\Big).$$

Note that, by the definition of a regular distribution of types, the image of NDP under $D(\cdot|)$ has measure 0. In particular, for almost every ϕ, demand at every associated market clearing cutoff is continuously differentiable.

Moreover, restricted to the open set $(0,1)^K \setminus NDP$, the demand function is continuously differentiable. Consequently, by Sard's theorem, the set of critical values of $D(\cdot)$ restricted to $(0,1)^K \setminus NDP$ has measure 0. That is, for almost all S, there are no vectors z in $(0,1)^K \setminus NDP$ such that $D(\cdot) = \phi$ and $\partial D(z)$ is singular. Together, these observations imply that, for almost all ϕ, demand at associated market-clearing cutoffs is both continuously differentiable and has an invertible derivative matrix.

□

As AL shows, the proof of the uniqueness of the stable equilibrium (Theorem A.2) follows from Claims 1 and 2. Take ϕ such that $\sum_k \phi_k < 1$ and assume that there is more than one stable matching. Claim 1 implies that demand is either non-differentiable, or has a singular derivative in at lest one of the market-clearing cutoffs. However, Claim 2 shows that this only holds for a measure 0 set of vectors ϕ. Thus, the set of vectors ϕ such that $\sum_k \phi_k < 1$ and there is more than one stable matching has measure 0, completing the proof.

Given that the stable matching equilibrium is unique, we can express the market-clearing cutoffs for each party using the structure of our model. Parties rank all politicians vertically by their assets, and given a constant leader ideology, all politicians prefer a more competent leader. Thus, the asset distribution of non-loyalist politicians in a given party is continuous for a given ideology. Consequently, equilibrium is characterized by each party having two cutoffs, where a politician of a given ideology can join the party only if their assets exceed the cutoff value. Therefore, the assets of party-k's members are within the range $[\underline{z}_k^j, \bar{z}_k^j], \forall \text{j} \in \{M, E\}$.

A.2 Proofs related to party structure

This section details the structure of political parties under the stable matching equilibrium. Proposition 1 outlines how party leaders recruit politicians based on their assets and ideological alignment.

Proof. (Proposition 1) A party leader randomly samples a share $1 - \phi$ of politicians from the pool of loyal politicians, $L^L(z)$. The remaining share ϕ of positions are filled by party organizers. Theorem 1 demonstrates the existence of a cutoff equilibrium where local organizers allocate ϕ of the positions to the highest-ability politicians available. This equilibrium is defined by the lowest-ability politicians recruited by each party. If all party leaders had the same ideology, every politician, regardless of their own ideology, would prefer to join the party led by the most competent leader. Given non-loyalist politicians' strict preference for parties based on the leader's competence, they would join parties according to their abilities. Consequently, the party led by the most competent leader would attract the highest-ability politicians from the pool of non-loyal politicians. The second-best leader would recruit from the remaining top-ability politicians, and this pattern would continue accordingly.

When there is an additional dimension of heterogeneity in politicians' preferences for parties—specifically, an ideological match—the cutoff equilibrium is characterized by two cutoffs for each party. These cutoffs represent the lowest asset politician recruited by each party for two ideologies: moderate and extremist. This occurs because, while party leaders still rank politicians based on their abilities, a politician might prefer a leader with a better ideological match over a more competent leader. Specifically local organizers of party k recruit j-ideology politicians whose assets fall between $(\underline{z}_k^j, \bar{z}_k^j)$, for $j \in \{M, E\}$, $\forall k$. The total density of non-loyalist politicians sum to ϕ, so we have

$$\phi_k = \int_{\underline{z}_k^M}^{\bar{z}^M} \ell^M(s)ds + \int_{\underline{z}_K^E}^{\bar{z}^E} \ell^E(s)ds. \tag{A.6}$$

The party structure follows. □

Corollary 1. The proof follows directly from Proposition 1. □

A.3 Proof of Theorem 1: Existence of a stable equilibrium

The proof of Theorem 1 demonstrates the existence of a stable equilibrium in the political recruitment process. This finding relies on the continuity of a party leader's value function. The proof's framework is grounded in established results from dynamic programming.

Proof. (Theorem 1) Given equation 3.9, equation 3.10 defines a contraction mapping for the value function of a leader, $V^{incumbent}$. The leader's flow payoff, $f(\cdot)$ is strictly increasing. The existence of the value function follows from standard theorems in dynamic programming. □

A.4 Proofs that characterize the optimal delegation rule

The optimal delegation rule determines how much authority a party leader delegates to local organizers when recruiting politicians. Theorem 2 derives the closed-form solution of this rule. Then, we prove Proposition 2, which demonstrates that more competent leaders tend to delegate more authority to local organizers, while moderate leaders, due to their abundance in the political spectrum, tend to delegate at least as much as extremist leaders. This result implies that the stability of a democracy depends on the competence and ideology of its political leaders. Finally, the proofs of corollaries 2 and 3 show how optimal delegation rule varies with external factors such as the availability of talented politicians, and intensified inter-party political competition.

Proof. (Theorem 2) Finding the first-order condition of a leader's value function in equation 3.10 with respect to the delegation rule, we obtain

$$\begin{aligned}
\frac{\partial V^{incumbent,j}}{\partial \phi} &= \frac{(1+\rho)f'}{\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'})} + \frac{(1+\rho)f(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})}{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))^2} = 0 \\
\implies \quad \frac{f'}{f} &= -\frac{(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})}{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))} \\
\implies \quad \frac{\eta x_k^{\eta-1}(z_k^D - z_k^L + \phi \frac{\partial z_k^D}{\partial \phi})}{x_k^\eta} &= -\frac{(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})}{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))} \\
\implies \quad \frac{\eta(z_k^D - z_k^L - \phi\frac{1}{\phi}(z_k^D + \sum_{j\in\{M,E\}}\Delta_k^j)}{\phi(z_k^D - z_k^L) + z_k^L} &= -\frac{(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})}{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))} \\
\implies \quad \frac{\phi(z_k^D - z_k^L) + z_k^L}{\eta(-z_k^L - \sum_{j\in\{M,E\}}\Delta_k^j)} &= -\frac{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))}{(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})} \\
\implies \quad \frac{\phi(z_k^D - z_k^L)}{\eta(z_k^L + \sum_{j\in\{M,E\}}\Delta_k^j)} + \frac{z_k^L}{\eta(z_k^L + \sum_{j\in\{M,E\}}\Delta_k^j)} &= \frac{(\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'}))}{(\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi})}
\end{aligned}$$

where the third line substitutes party size (equation 3.3) into the production function defined in Assumption 2. Note that the expected assets of non-loyalist politicians, z_k^D, depend on the extent of the leader's delegation through the relationship $\frac{\partial z_k^D}{\partial \phi} = \frac{1}{\phi}\left[z_k^D + \sum_{j\in\{M,E\}}\Delta_j I_k^j\right]$, where for $j \in \{M,E\}$ we have $\Delta_j = \frac{\partial \underline{z}_k^j(\phi)}{\partial \phi}\ell^j(\underline{z}_k^j(\phi))(\int_{\underline{z}_k^j(\phi)}^{\bar{z}_k^j} s\ell^j(s)ds - \underline{z}_k^j(\phi)\int_{\underline{z}_k^j(\phi)}^{\bar{z}_k^j}\ell^j(s)ds)$. Rearranging the terms, we obtain

$$\phi = -\frac{z_k^L}{z_k^D - z_k^L} + \frac{\eta(z_k^L + \sum_{j\in\{M,E\}}\Delta_k^j)}{z_k^D - z_k^L}\frac{\rho + \alpha^j(1-\pi^{jj}) + \alpha^{j'}(1-\pi^{jj'})}{\alpha^j \frac{\partial \pi^{jj}}{\partial \phi} + \alpha^{j'}\frac{\partial \pi^{jj'}}{\partial \phi}}. \tag{A.7}$$

Next, we substitute for a leader's probability of winning an election defined in equations 3.6 and 3.7 in A.7. Rearranging the terms, we obtain the closed-form solution for the optimal delegation rule

$$
\begin{aligned}
\phi &= -\frac{z_k^L}{z_k^D - z_k^L + \eta(z_k^L + \sum_{j\in\{M,E\}} \Delta_k^j)} \\
&- \frac{\rho\eta(z_k^L + \sum_{j\in\{M,E\}} \Delta_k^j)}{(\alpha^j(1-\eta) + \alpha^{j'}(1-\Omega^{jj'}))(z_k^D - z_k^L + \eta(z_k^L + \sum_{j\in\{M,E\}} \Delta_k^j))}
\end{aligned}
$$

Proof. (Proposition 2. Fragility of a liberal democracy) The proof proceeds by showing that i) the optimal delegation rule is increasing in a leader's competence level, η and ii) a moderate leader's delegation rate is no less than that of an extremist, ceteris paribus.

The effect of competence on delegation rate:

Let $A = \alpha^j(1-\eta) + \alpha^{j'}(1-\Omega_k^{jj'})$. Then $\frac{\partial A}{\partial \eta} = -\alpha^j - \alpha^{j'}\frac{\partial \Omega^{jj'}}{\partial \eta}$

$$
\begin{aligned}
\frac{\partial \phi_k}{\partial \eta} &= \frac{1}{(1-\eta)^2}\frac{z_k^L}{(z_k^L - z_k^D)} + \frac{\rho(1-\eta)A - \eta\rho\left(-A + (1-\eta)\frac{\partial A}{\partial \eta})\right)}{(1-\eta)^2A^2} \\
&= \frac{1}{(1-\eta)^2}\frac{z_k^L}{(z_k^L - z_k^D)} + \frac{\rho A - (1-\eta)\frac{\partial A}{\partial \phi}}{(1-\eta)^2A^2} \\
&= \frac{1}{(1-\eta)^2}\frac{z_k^L}{(z_k^L - z_k^D)} + \frac{\rho}{(1-\eta)^2A} + \frac{\eta\rho}{(1-\eta)^2A} - \frac{\eta\rho}{(1-\eta)^2A} - \frac{(1-\eta)\frac{\partial A}{\partial \eta}}{(1-\eta)^2A^2} \\
&= \frac{1}{1-\eta}\left(\frac{1}{(1-\eta)}\frac{z_k^L}{(z_k^L - z_k^D)} + \frac{\eta\rho}{(1-\eta)A}\right) + \frac{(1-\eta)\rho}{(1-\eta)^2A} - \frac{\frac{\partial A}{\partial \eta}}{(1-\eta)A^2} \\
&= \frac{1}{1-\eta}\left(\frac{1}{(1-\eta)}\frac{z_k^L}{(z_k^L - z_k^D)} + \frac{\eta\rho}{(1-\eta)A}\right) + \frac{\rho}{(1-\eta)A} - \frac{\frac{\partial A}{\partial \eta}}{(1-\eta)A^2} \\
&= \frac{1}{1-\eta}\left(\phi_k + \frac{\rho}{A} + \frac{\alpha^j + \alpha^{j'}\frac{\partial \Omega^{jj'}}{\partial \eta}}{A^2}\right)
\end{aligned}
$$

Because $\phi_k \geq 0$ and $\frac{\partial \Omega^{jj'}}{\partial \eta} \geq$ by assumption, we have that $\frac{\partial \phi_k}{\partial \eta} \geq 0$.

Comparison of moderate and extremist leaders' delegation decisions

We show that both terms in the optimal delegation rule, given below, are at least as great for an M- leader as an E-leader

$$\phi_k = -\frac{1}{(1-\eta)}\frac{z_k^L}{(z_k^D - z_k^L)} + \frac{\eta}{(1-\eta)}\frac{\rho}{\alpha^j(1-\eta) + \alpha^{j'}(1-\Omega_k^{jj'})}.$$

Consider two party leaders with identical competence levels but different ideologies. Since moderate (M) politicians are more abundant at each skill level, $L^M(z) < L^E(z), \forall z$, we expect the average assets of non-loyalists, z_k^D to be at least as high with and M leader compared to an E leader. This is because a politician might prefer a less competent leader with a better ideological match over a more competent leader with a worse ideological match. Thus, when E is the leader, available M politicians may choose to join the party of a less competent M leader instead. Similarly, when M is the leader, top E politicians who would have joined the party with the E leader might instead join the party of the next best E leader with lower competence. However, because M politicians are more abundant than E politicians, the party loses more in terms of non-loyalist politician assets with an M leader. Therefore, z_k^D is expected to be at least as high for the M leader as it is for the E leader. Consequently, the first fraction is as high for the M leader as for the E leader.

Next, we compare the term $\alpha^j(1-\eta) + \alpha^{j'}(1-\Omega_k^{jj'})$ for both ideology types of leaders. First, notice that $\alpha^M > \alpha^E$ simply because the M-type is more abundant.

Now, consider an M-incumbent. If an M challenger comes, all non-loyalists choose the more competent one, so the probability of losing for the incumbent is $1-\eta$. If an E-challenger comes, the incumbent's probability of losing is $1-\Omega_k^{ME} = 1 - \psi^M\eta^{ME}(\eta) - (1-\psi^j)\eta^{EE}$ (see equation 3.7). Note that because $\eta^{ME}(\eta) \geq \eta$ and $\eta^{EE}(\eta) \leq \eta$, (i.e., M members prefer and E challenger only if they are sufficiently more competent than the M incumbent, but E members can choose the E challenger even when the challenger is less competent than the incumbent) and $\psi^M > 0.5$ because M members are more abundant in the economy, we have that $1-\eta > 1-\Omega_k^{ME}$.

Next, consider an E incumbent. If an E challenger arrives, all non-loyalists choose the higher competence leader, so the incumbent's probability of losing is $1-\eta$. If an M-challenger comes, his probability of losing is 1-$\Omega_k^{EM} = 1 - \psi^M\eta^{MM}(\eta) - (1-\psi^j)\eta^{EM}$. Again, we have $\eta^{MM}(\eta) < \eta$ and $\eta^{EM}(\eta) \geq \eta$. Because

$\psi^M > 0.5$, we have $1 - \eta < 1 - \Omega_k^{EM}$.

Putting together, the denominator of the second fraction for an M incumbent, $\alpha^M(1-\eta) + \alpha^E(1-\Omega_k^{ME})$, is smaller than that of an E-incumbent, $\alpha^E(1-\eta) + \alpha^M(1-\Omega_k^{EM})$, and so the second term is also larger for an M leader, which completes the proof. □

Proof. (Corollary 2).

$$\frac{\partial \phi_k}{\partial z_k^L} = -\frac{1}{(1-\eta)} \frac{z_k^D}{(z_k^L - z_k^D)^2} < 0$$

$$\frac{\partial \phi_k}{\partial z_k^D} = -\frac{1}{(1-\eta)} \frac{z_k^L}{(z_k^L - z_k^D)^2} > 0$$

□

Proof. (Corollary 3) An examination of the optimal delegation rule in equation 3.11 shows that the introduction of a new party may impact a leader's optimal delegation rule primarily by changing the composition of non-loyalist politicians recruited by local organizers. If the new party appeals to some of the non-loyalists who would have been recruited in its absence, the organizers will need to recruit from the lower-ability politicians available. This leads to two implications. First, the average assets of non-loyalist party members, z_D, would decrease. Second, the distribution of ideology among non-loyalists may be affected, which would change the probability that non-loyalists choose the incumbent during a leadership challenge, $\Omega_k^{jj'}$. While the total assets of the non-loyalists are expected to decrease, prompting the leader to lower her delegation rate, the effect on the leadership survival probability can vary.

For example, if the new party shares the same ideology as the incumbent party but has a more competent leader, it can attract both moderate and extremist politicians from the incumbent party. Consequently, local organizers will aim to fill the party with the best available politicians, and the ideological composition of the newly recruited politicians will depend on their availability in the economy. The new share of moderate non-loyalists in the party could be either lower or higher.

If the newly introduced party does not alter the incumbent party's preference ranking for both ideology types of politicians, the party leader's optimal delegation rule will remain unaffected.

Other terms in the optimal delegation rule are not affected by the introduction of a new party. These include exogenous parameters (such as the discount rate, rent production technology, and leadership challenge arrival rate) or factors that govern within-party leadership competition, such as the distribution of loyal politicians' assets and the incumbent leader's competence level. □

A.5 Proof of Theorem 5: Welfare analysis

This section analyzes the welfare implications of the political recruitment process. Voters' welfare is assumed to increase with the average quality of politicians. We show that while more delegation can improve welfare by recruiting higher-quality non-loyalist politicians, the returns to delegation are diminishing, and beyond a certain point, further delegation may not lead to welfare improvements.

Proof. (Theorem 5.) Voters' welfare is assumed to be increasing in the average quality of all politicians. This implies that more delegation is welfare improving as long as the expected asset of a loyalist politician, $E[\ell^L(z)]$ is greater than the expected asset of the last recruited non-loyalist politician. However, this may not hold for all delegation levels because delegation has diminishing marginal returns for voters' welfare. Because parties rank politicians vertically by their assets, the organizers recruit from the right-tail of the distribution of non-loyalist politicians, $L^M(z)$ and $L^E(z)$. After the richest politicians are recruited by the politicians' top-ranked parties, the remaining parties begin to recruit from the remaining truncated distribution of non-loyalist politicians. Depending on the number of competing parties and the actual shapes and locations of the loyalist and non-loyalist politicians, average expected asset of a loyalist politician may be greater than the asset of the marginally-recruited non-loyalist politician. □