

Grunewald, Andreas; Hansen, Emanuel; Pönitzsch, Gert

**Conference Paper**

## Political Selection and the Concentration of Political Power

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Collective Decision Making I, No. A06-V1

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

Suggested Citation: Grunewald, Andreas; Hansen, Emanuel; Pönitzsch, Gert (2014) : Political Selection and the Concentration of Political Power, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Collective Decision Making I, No. A06-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<http://hdl.handle.net/10419/100339>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Political Selection and the Concentration of Political Power

Andreas Grunewald<sup>1</sup>, Emanuel Hansen<sup>1,2</sup>, and Gert Pönitzsch<sup>1,3</sup>

<sup>1</sup>University of Bonn

<sup>2</sup>University of Cologne

<sup>3</sup>Kiel Institute for the World Economy

February 2014

## Abstract

This paper studies the effects of power-concentrating institutions on the quality of political selection, i.e., the voters' capacity to identify and empower competent politicians. In our model, candidates are privately informed about their abilities and are driven by office rents as well as welfare considerations. We show that variations in power concentration involve a trade-off. On the one hand, higher concentration of power enables the voters' preferred politician to enforce larger parts of his agenda. On the other hand, higher power concentration increases electoral stakes and thereby induces stronger policy distortions. We identify a negative relation between the optimal level of power concentration and the extent of office motivation. In particular, full concentration of power is desirable if and only if politicians are mostly welfare-oriented. The results of an empirical analysis are in line with this prediction.

**Keywords:** Elections, Constitutional Design, Selection, Asymmetric Information

**JEL classification:** D72, D82, H11

We are grateful to Rafael Aigner, Felix Bierbrauer, Paul Heidhues, Martin Hellwig, Mark Le Quement, Gilat Levy, and seminar participants in Bonn, Cologne, Tübingen, and at the LSE for very fruitful discussions. Financial support from the Bonn Graduate School of Economics is gratefully acknowledged.

# 1 Introduction

In representative democracies, political power is exercised by elected politicians. The role of institutions is to enforce the voters' interests within the political process. From the founding of modern democracies in the 18th century to recent constitutional drafts in Egypt and Lybia, political thinkers have been engaged in finding the best institutions. A central question in the debate has been whether political power should be concentrated on one group of political agents, typically the party winning the general election, or dispersed between different groups. Strikingly, there are pronounced cross-country differences along this dimension, with classical extreme cases being the United Kingdom (concentrated power), and Switzerland (dispersed power).<sup>1</sup>

The Federalist Papers highlight two channels through which constitutions affect social welfare: the selection of competent politicians into office and the disciplining of politicians in office.<sup>2</sup> The economic literature on the second channel consistently finds that power-dispersing institutions increase welfare as they help to discipline egoistic incumbents. In contrast, economists have little to say about the first channel, political selection (see Besley 2005). It cannot be taken for granted that voters are able to identify and empower the most competent politicians. Since voters base their ballot on their perceptions of candidates' competencies (Stokes et al. 1958, King 2002, Pancer et al. 1999), politicians exert considerable effort to appear competent and virtuous during electoral campaigns. This impedes the voters' capacity to empower able candidates. A comprehensive appraisal of political institutions thus has to account for whether or not institutions enforce the selection of competent candidates for office.

The aim of this paper is to study the effects of power-concentrating institutions on the politicians' campaign behavior and on political selection. We consider a pre-election setup in which candidates are privately informed about their quality and partly motivated by office rents. Voters infer candidates' qualities from their policy proposals. We identify a trade-off that arises for changes in the level of power concentration. On the one hand, higher concentration of power implies a better allocation of power to competent candidates, as long as political campaigns provide at least some information to voters. We refer to this positive effect on

---

<sup>1</sup>For a discussion of this crucial issue and its relation to various specific institutions, see Lijphart (2012), Lijphart (1999), and Tsebelis (2002).

<sup>2</sup>“The aim of every political Constitution is, or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of society; and in the next place, to take the most effectual precautions for keeping them virtuous whilst they continue to hold their public trust” (Madison 1788*b*).

welfare as the *empowerment effect* of power concentration. On the other hand, more concentration of power increases the desire of office-motivated candidates to win the election. Mimicking of competent candidates becomes more profitable, resulting in increasingly distorted policy choice. Thus, campaigns convey less information about the candidates' competence, and voters are less able to select high-ability politicians. We label this negative effect on welfare the *behavioral effect* of power concentration.

We formalize our argument by a simple model in which two candidates compete in a public election by making binding policy proposals. In particular, they can either commit to a risky reform or to the (riskless) status quo. Candidates differ in their abilities, which are unobservable to the electorate. Only highly able candidates can increase expected welfare by implementing a reform, while less able candidates should stick to the status quo. Voters observe policy proposals, draw inferences about the candidates' abilities, and vote accordingly. In equilibrium, a reform proposal is associated with high ability, and reforming candidates win the election more often than those proposing the status quo. Politicians not only care about welfare, but are also office-motivated. This creates incentives for low-ability candidates to mimic the policy choice of their more able counterparts at the cost of adopting inefficient policies.

Variations in the level of power concentration induce the empowerment effect and the counteracting behavioral effect. The relative sizes of these effects depend on the importance of office motivation in politicians' preferences. The optimal institution balances both effects. We find that the optimal level of power dispersion is higher, the more politicians are driven by office rents. If and only if politicians care predominantly about implementing efficient policies, it is optimal to concentrate power completely in the hands of the election winner. Conversely, if office rents are a strong component of candidates' motivation, some dispersion of political power enhances voter welfare.

The basic intuition behind this result is the following. Candidates' office motivation induces mimicking and distorts policy choices. Higher concentration of power strengthens the electoral incentives and, consequently, aggravates these distortions. For particularly high office motivation, it is optimal to reduce the resulting inefficiencies by decreasing the concentration of power, even though this involves giving some power to low-ability candidates. Hence, power-dispersing institutions are beneficial if and only if politicians are mainly office-motivated.

Many important dimensions of the political process are intentionally omitted from the basic model to build up the intuition behind the results. However, the underlying intuition is very generally applicable to the process of political selection.

To demonstrate this we show that qualitatively identical findings arise in a variety of model extensions. First, we introduce a continuous policy space by assuming that candidates may choose the magnitude of reform they propose rather than restricting their choice set to a reform and the status quo. Second, we relax the assumption of binding policy commitments by introducing the possibility to sometimes withdraw a proposed reform after the election. Third, we allow for heterogeneous policy preferences in the electorate. In this setting, we additionally show that increasing power dispersion reduces inequality in the society. Fourth, we show that our results also hold if candidates are heterogeneous with respect to their office motivation.

Data from international surveys like the *International Social Survey Panel* indicate considerable differences across countries in how voters assess the motives of their politicians. Assuming that these differences mirror actual heterogeneity in politicians' motivation, our theoretical analysis gives rise to a testable hypothesis: Countries in which politicians are policy-motivated benefit from power concentration. In contrast, countries with predominantly office-motivated politicians suffer from reduced welfare if they concentrate political power.

In a cross-country design, we investigate whether the welfare effect of power concentration depends on politicians' motivation. We combine data on the perceived motivation of politicians with measures of political institutions.<sup>3</sup> As a measure for the performance of the political system, we use growth in real GDP per capita. Due to data availability restrictions, our analysis is restricted to eighteen established democracies. For this set of countries, the data provide support for our hypothesis.

The paper proceeds as follows. The next section reviews the related literature. Section 3 presents the model. Section 4 delivers the benchmark of perfect information. Thereafter, we analyze the equilibrium behavior of privately informed politicians in Section 5. We proceed by examining the effects of institutions in Section 6. In Section 7, we present design and results of the empirical analysis. Section 8 provides a number of modifications and extensions of the basic theoretical model and Section 9 concludes. All formal proofs are provided in the Appendix.

## 2 Related literature

In this paper, we identify the economic effects of power-dispersing institutions, which limit the office-holders discretion. Many economists have addressed this question for a homogeneous set of politicians, thereby abstracting from political selection. With

---

<sup>3</sup>We use Lijphart's index of the executive-parties dimension, which orders political systems according to the implied dispersion or concentration of power, considering five categories of political institutions (Lijphart 1999).

homogenous politicians, power-dispersing institutions increase voter welfare. For example, Lizzeri & Persico (2001) demonstrate in a pre-election setting that office-motivated politicians provide more of an efficient public good and less pork barrel under proportional representation than under plurality voting. In a post-election setting, Persson & Tabellini (2003) show that voters are more able to discipline an incumbent if power is separated between multiple political agents.

These papers abstract from any heterogeneity in candidate quality and thus from the role of political selection.<sup>4</sup> The importance of incorporating selection into the analysis of political institutions is demonstrated by Besley (2005). Selecting competent politicians has two aspects. First, to choose among competing candidates the one who holds most promise to design and implement efficient policies. This aspect is based on the candidates' campaign behavior. Second, to keep in office only politicians who perform adequately during the term. In other words, politicians are screened both before and after an election.

The role of institutions for political selection has so far only been studied in post-election settings. A first model addressing this question is Maskin & Tirole (2004). It investigates conditions under which the voter prefers political decisions to be taken by accountable "politicians" instead of non-accountable "judges". Maskin & Tirole (2004) argue that holding public officials accountable in reelections provides incentives to pander to public opinion and is thus not optimal for some kinds of political decisions. While they do not compare alternative democratic institutions, this approach is taken by Smart & Sturm (2006). They study variations in the level of accountability through the introduction of term limits. Depending on the share of public-spirited politicians, a limit of two terms as applied in many modern democracies is shown to be optimal. In a dynamic framework, Acemoglu et al. (2010) compare political selection in democracies and autocracies, distinguishing these regimes according to the veto power that incumbents have over changes in government. They find that perfect democracy always leads to the emergence of the most competent government, while even the least able government can persist forever with any degree of incumbency veto power.

Closest to our paper is the analysis by Besley & Smart (2007), who study the effects of several fiscal restraints on political selection in a post-election setting. Similar to Maskin & Tirole (2004) and Smart & Sturm (2006), they identify a trade-off between disciplining incumbents and improving political selection. Whenever

---

<sup>4</sup>The assumption that candidates differ in a quality dimension, sometimes referred to as "valence issue", is applied in a large number of papers, including Adams (1999), Ansolabehere & Snyder (2000), Aragones & Palfrey (2002), Messner & Polborn (2004), Sahuguet & Persico (2006), and Krassa & Polborn (2011).

an institution allows the disciplining of bad incumbents, i.e., makes them adopt welfare-enhancing policies, this prevents effective political selection because voters are unable to distinguish a disciplined but bad politician from a good one. Our pre-election model produces a different trade-off. If voters have to infer the ability of candidates from their campaigns, dispersing power leads to both better policy choice and better selection, but comes at the cost of giving some political power to low-ability candidates. Besley & Smart (2007) consider four fiscal restraints that limit the office-holders' discretion, including limits on government size and transparency. Our focus, in contrast, is on power-dispersing institutions, such as proportional representation, federalism, or public referenda. Interestingly, Besley & Smart (2007) find that three of the four restraints only increase voter welfare if there are sufficiently many benevolent politicians. This contrasts our result according to which power dispersion is optimal if and only if the candidates are strongly driven by egoistic motives.

A different selection aspect from the one we study is the voters' decision whether to enter politics and thus the pool of political candidates. This aspect has been studied by Mattozzi & Merlo (2008) and in citizen-candidate models by Caselli & Morelli (2004) and Messner & Polborn (2004). While these models capture the effects of financial incentives on the quality of the pool of candidates, they assume that power is equally distributed among elected politicians. Thus, political institutions that affect the concentration of power cannot be analyzed. The analysis of Iaryczower & Mattozzi (2008) is closer to ours, as it deals with voting systems and their effect on the quality of political candidates. They show that neither the majoritarian electoral system nor proportional representation clearly brings forth a higher-quality pool of candidates. Our model analyzes a broader range of institutions and their effect on the competence in government, but takes the pool of candidates as given.

Finally, we relate to a growing empirical literature on democratic systems and their effects on fiscal policy. The analyses often focus on specific political institutions (see, e.g., Feld & Voigt 2003, Persson & Tabellini 2004, Enikolopov & Zhuravskaya 2007, Blume et al. 2009, Voigt 2011). In contrast, we apply a classification of political systems based on the implied dispersion of political power, thus encompassing a broad range of institutions. Using the same classification, Lijphart (1999, 2012) as well as Armingeon (2002) examine the influence of power dispersion on various political and economic outcomes. While Lijphart (1999) finds no effect of power dispersion on measures of economic performance, Armingeon (2002) finds a negative effect of power dispersion on unemployment and inflation. Complementing these findings, we show that the effect of power dispersion on growth in real GDP per

capita positively depends on the strength of politicians' office motivation.

### 3 The model

Our model studies the effects of institutions on candidates' campaigns and political selection. Candidates differ in quality, more precisely in the ability to implement welfare-enhancing policies. They are privately informed about their abilities and commit to a policy prior to the election. Voters observe candidates' campaigns and vote based on the welfare they expect each candidate to provide. We depict political institutions in reduced form, by means of how much political power is concentrated in the political system. With higher concentration of power, the candidate receiving a majority of votes is more capable to enforce his agenda.

The game consists of three stages. At the first stage, nature independently draws both candidates' abilities  $a_1$  and  $a_2$ , which are privately revealed to the candidates. At the second stage, both candidates simultaneously make binding policy proposals,  $x_1$  and  $x_2$ . At the third stage, voters observe the proposals, update their beliefs about the candidates' abilities and cast their votes. Based on the election result, political power is divided between both candidates according to a power allocation rule. Finally, a policy decision is taken.

While the basic model serves to clarify the main arguments, we discuss a number of modifications in Section 8. In particular, we allow for a continuous policy space, a form of limited commitment and heterogeneity in the voters' policy preferences. Importantly, these modifications do not alter the main results of the basic model.

#### 3.1 Voters

There is a continuum of fully rational and risk-neutral voters of mass one who have preferences both over policy and candidates. In the considered policy field an amount of a reform ( $x \in [0, 1]$ ) can be implemented, where  $x = 0$  is associated with the status quo and  $x = 1$  with a complete reform. All voters receive the same positive payoff from the reform if and only if it is successful. More precisely, we assume that a successful reform yields a return of  $x$  to each voter while a failed reform yields a return of zero. Whenever a reform is adopted, all voters bear a cost of  $c$ . Maintaining the status quo is costless and yields a certain payoff of zero.

Voters might also care about other policy fields and about the candidates' ideologies or personal characteristics other than ability. We account for these preferences in the tradition of the probabilistic voting model by assuming that voters have heterogeneous candidate preferences (Lindbeck & Weibull 1987). If policy is set by



candidate 1, voter  $k$  receives an additional utility of  $\mu_k$ , while we normalize the additional utility if candidate 2 determines policy to zero. Let  $\mu_k$  be distributed according to some continuous pdf that is symmetric around zero and has full support on the interval  $[-1, 1]$ . This guarantees heterogeneity in the resulting voting preferences.<sup>5</sup>

Altogether, if candidate  $i$  is in power and sets policy  $x_i$ , voter  $k$  receives a utility of

$$V_k(x_i, i) = \begin{cases} \mathbb{1}(i)\mu_k + x_i - c & \text{reform succeeds} \\ \mathbb{1}(i)\mu_k - c & \text{if reform fails} \\ \mathbb{1}(i)\mu_k & \text{status quo is maintained,} \end{cases} \quad (1)$$

where  $\mathbb{1}(i)$  denotes the indicator function which is one if  $i = 1$  and zero otherwise. Voter  $k$  prefers candidate 1 if and only if he expects  $V_k(x_1, 1)$  to be larger than  $V_k(x_2, 2)$ . Voters vote sincerely, i.e., each voter casts his vote for his preferred candidate. Hence, candidate  $i$ 's vote share depends positively on the voters' belief about the payoff candidate  $i$  provides, and negatively on the belief about the payoff provided by his opponent.

## 3.2 Candidates

Two candidates run for office. Each candidate  $i$  can commit to an amount of a reform  $x_i \in [0, 1]$ . More able candidates design better reforms, i.e., reforms that are more likely to succeed. We measure candidate  $i$ 's ability by the implied probability of a successful reform,  $a_i \in [0, 1]$ . Hence, policy  $x_i$  set by candidate  $i$  provides an expected payoff of  $x_i(a_i - c)$  to each voter. Since candidate preferences are symmetrically distributed around zero, this is candidate  $i$ 's welfare contribution.

Nature independently draws both candidates' abilities from the cumulative distribution  $\Phi$ . Let the corresponding density function  $\phi$  have full support on  $[0, 1]$  and be continuously differentiable. After observing his ability, each candidate  $i$  commits to policy  $x_i \in [0, 1]$ . Thus, the strategy  $X_i$  of politician  $i$  is a mapping from abilities to policy proposals.

Each candidate cares about his amount of political power (office motivation) as well as about his expected welfare contribution (policy motivation). The utility function of politician  $i$  is given by

$$U_i(a_i, x_i) = f(v_i, \rho) [\theta + x_i(a_i - c)], \quad (2)$$

---

<sup>5</sup>Note that our results are independent of whether these candidate preferences are subject to an additional aggregate shock as in Lindbeck & Weibull (1987).

where  $\theta > 0$  denotes the relative weight of office motivation. In particular,  $\theta$  describes the marginal increase in the candidates' office utility due to an increase in the amount of power.<sup>6</sup> Candidate  $i$ 's power  $f(v_i, \rho) \in (0, 1)$  describes his influence on the policy that is implemented after the election. This policy is a compromise between both campaign proposals. More precisely, we assume that the implemented policy is a linear combination of the proposals with weights being equal to the power of the candidates. Each candidate's power depends on his vote share  $v_i$  and the parameter of power concentration  $\rho$ , representing the set of political institutions. To simplify notation, this utility function is formulated at an ex interim stage, i.e., taking the expected payoff after the election but before the reform outcome has been realized. For readability, we have also omitted the dependence of  $v_i$  on both candidates' strategies and actions.

Note that candidate  $i$  only cares about how expected welfare is affected by his policy choice, not about voter welfare in general. This way to formulate policy preferences of politicians has been introduced by Maskin & Tirole (2004), who phrase it *legacy motive*. It captures politicians' desire to leave a positive legacy to the public.<sup>7</sup>

### 3.3 Political institutions

We model political institutions by a power allocation function  $f$  that translates election results into an allocation of political power, i.e., each proposal's influence on the policy that is implemented after the election. Formally, candidate  $i$ 's power  $f(v_i, \rho)$  depends on his vote share  $v_i$  and on the level of power concentration  $\rho$  implied by the set of political institutions.

**Definition 1.** *The continuously differentiable function  $f : [0, 1] \times \mathbb{R}_+ \rightarrow [0, 1]$  is a power allocation function if it satisfies:*

1. *symmetry in  $v_i$ :  $f(v_i, \rho) = 1 - f(1 - v_i, \rho)$*
2. *monotonicity in  $v_i$ :  $\frac{\partial f(v_i, \rho)}{\partial v_i} \geq 0 \forall \rho$*
3. *piece-wise monotonicity in  $\rho$ :  $\frac{\partial f_i(v_i, \rho)}{\partial \rho} > 0 \forall v_i \in (1/2, 1)$ .*
4.  *$\lim_{\rho \rightarrow \infty} f(v_i, \rho)$  is a step function with a discontinuity at  $v_i = \frac{1}{2}$ .*

---

<sup>6</sup>We assume that  $\theta$  mirrors the candidates' intrinsic utility of power exertion and can thus not be manipulated by institutions.

<sup>7</sup>Alternatively, we could assume that candidates directly care for welfare. While complicating the analysis, this assumption would not change the qualitative results.

The first property establishes anonymity, i.e., the constitution does not treat candidates differently. The second rules out that candidates receive a larger amount of political power if they gain less votes in the election. The third allows us to interpret any rise in  $\rho$  as an unambiguous increase in the concentration of power. The higher  $\rho$ , the larger is the amount of power assigned to the election winner, i.e., the candidate who gains more than half of the votes. The fourth, implies that in the case of full power concentration, the election winner receives all power that is allocated through the corresponding election.

This modeling approach allows to study a large variety of institutional differences. Figure 1 illustrates how political institutions can be represented by power allocation functions. Each panel depicts two examples. Throughout, the solid line represents institutions that concentrate power more strongly than those corresponding to the dashed line.

Panel I depicts two stylized allocation rules frequently used to compare electoral systems in the theoretical literature (see, e.g., Lizzeri & Persico 2005). The solid line represents institutions that fully concentrate power in the hands of the election winner. This step function is the standard way to model plurality voting. The dashed line represents proportional representation, which implies a lower concentration of political power and is often modeled by the identity function  $f(v_i, \rho) = v_i$ .

A less simplistic representation of these two systems is shown in Panel II. Here, the winner's amount of power depends on his margin of victory, e.g., because delegates might vote against the party lines. Plurality voting tends to generate clear-cut majorities, as the winning party typically receives a share of parliamentary seats beyond its vote share. In contrast, the allocation of seats corresponds closely to vote shares under proportional representation. Thus, the dashed curve for the proportional system is flatter than the one for plurality voting.

In Panel III, the dashed line represents a political system with a supermajority requirement for certain policy decisions (as employed in Germany and the US). This requirement generates additional steps in the power allocation function, since some policies can only be enforced after a landslide victory. In contrast, the solid line corresponds to a system as applied in the UK, where any decision can be taken by a simple majority.

Finally, the dashed line in Panel IV depicts the use of direct democratic institutions as employed for example in Switzerland. Even after a landslide victory in the election, the winning party cannot always implement its agenda. The opposition party can block policies via a referendum or even enforce its own proposals. Thus, only a limited part of political power is at stake in the parliamentary election (similar

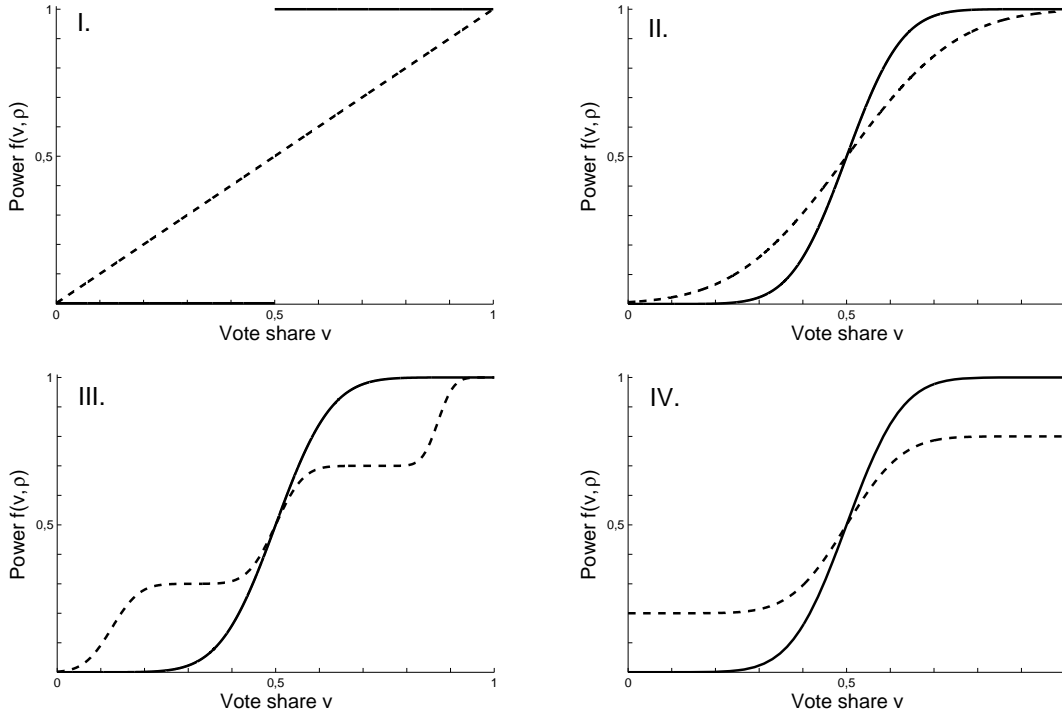


Figure 1: Political institutions and the corresponding power allocation functions.

arguments can be made with respect to federalism, bicameralism or a constitutional court).

### 3.4 Equilibrium concept and normative criterion

To solve the game, we study Perfect Bayesian equilibria. Thus, an equilibrium of the game consists of a strategy profile and a belief system such that (1) both candidates play mutually best responses when announcing their policy proposals, anticipating the winning probabilities for each vector  $(x_1, x_2)$  that are implied by the voters' beliefs, and (2) the voters' belief system  $\sigma$  is derived from the candidates' strategies  $X_1, X_2$  according to Bayes' rule everywhere on the equilibrium path. In the main part of the paper we refine the set of equilibria by applying the D1 criterion proposed Cho & Kreps (1987), which restricts the feasible set of out-of-equilibrium beliefs. Intuitively, D1 requires that each deviation from equilibrium actions be attributed to the type that profits most of it.

In the following, we analyze the effects of changes in power concentration, i.e., in parameter  $\rho$ . As normative criterion, we use a utilitarian welfare function in ex

ante perspective, i.e., expected welfare before candidates' abilities are drawn:

$$W(\rho, \theta) = \int_0^1 \int_0^1 \phi(a_1)\phi(a_2) \sum_{i=1}^2 f(v_i, \rho) X_i(a_i)(a_i - c) da_2 da_1. \quad (3)$$

Welfare is, hence, given by the weighted sum of the politicians' welfare contributions, integrated over all possible combinations of candidates' ability. The weights correspond to the candidates' power,  $f(v_i, \rho)$ . Note that welfare is calculated using equilibrium strategies, which are functions of the parameters  $\rho$  (power concentration) and  $\theta$  (candidates' office motivation).

## 4 Benchmark case: perfect information

If individual abilities are observable to the electorate, voters condition their ballot on candidates' abilities and reform proposals. In particular, the fraction of citizens voting for a candidate is increasing in his welfare contribution.

Under perfect information, candidates' motives are fully aligned: Each candidate maximizes his power by proposing the policy with the highest welfare contribution. Hence, a complete reform ( $x_i = 1$ ) is proposed by high-ability candidates with  $a_i \geq c$ . In contrast, a candidate with ability  $a_i < c$  gains more power by proposing the status quo instead of a reform with a negative welfare contribution. Thus, candidates' policy choices are undistorted: A politician proposes to implement a complete reform if and only if the reform enhances welfare. As a consequence, candidates with higher ability, i.e., those who propose to reform, receive higher vote shares in the election.

This result has a direct welfare implication. While variations in power concentration  $\rho$  do not distort candidates' behavior, a higher concentration of power allocates more power to candidates with higher welfare contribution. Hence, welfare strictly increases with the level of power concentration. The following Proposition summarizes these results.

**Proposition 1.** *Under perfect information, candidates propose a complete reform if and only if  $a_i \geq c$ . Welfare is maximized if political power is completely concentrated.*

## 5 Imperfect information

In the following, we derive the equilibrium properties under the assumption that candidates are privately informed about their abilities. First, we argue that candidates always play cutoff strategies in D1 equilibria. These involve either a complete reform proposal or a status quo proposal if the ability is above or below a certain cutoff, respectively. In a second step, we derive existence of D1 equilibria and show that strategies are symmetric in equilibrium. In a third step, we establish uniqueness of equilibria.

Under imperfect information, voters form beliefs about politicians' welfare contribution on the basis of their policy proposals. The vote share of candidate  $i$  thus depends on the proposals only, and not on candidates' abilities. However, the higher a candidate's ability  $a_i$ , the higher is his welfare contribution from every increment of a reform. Hence, the reform magnitude candidates propose must be weakly increasing in ability.

**Lemma 1.** *Given any belief system  $\sigma$  and any strategy of the opponent  $X_j$ , the optimal strategy of candidate  $i$  is increasing in his ability.*

The D1 criterion yields that only complete reforms or the status quo are played in equilibrium. In particular, the most able candidates choose to propose a complete reform. They would only be inclined to propose an intermediate reform, if a complete proposal is associated with low ability. This, however is not feasible under D1 since able candidates profit most of a reform proposal. As a consequence, all agents with ability above  $c$  propose a complete reform. They profit from the higher welfare contribution as well as from the higher office utility generated by this proposal. If an intermediate reform is proposed in equilibrium, it is thus associated with a negative welfare contribution. Consequently, such a proposal leads to a smaller vote share than the status quo proposal, and candidates below  $c$  strictly prefer the latter. Overall, only the two extremes of the action space are played in a D1 equilibrium.

Hence, the decision of candidates boils down to either propose a complete reform or the status quo. Clearly, a candidate's relative incentive to propose a complete reform is monotonically increasing in his ability. Hence, in every D1 equilibrium the behavior of candidates exhibits the cutoff property, i.e., all candidates with ability greater than or equal to some cutoff value  $\alpha_i$  propose a complete reform, while candidates with lower ability propose the status quo.

**Lemma 2.** *In every D1 equilibrium, the optimal strategy of candidate  $i$  can be*

characterized by a unique cutoff  $\alpha_i \in [0, 1]$  such that

$$X_i(a_i) = \begin{cases} 1, & \text{if } a_i \geq \alpha_i \\ 0, & \text{if } a_i < \alpha_i. \end{cases}$$

In the following, we denote strategies only by their corresponding cutoffs. Candidates compare the utility of proposing a reform and of proposing the status quo, given the strategy of the opponent and the voters' belief system  $\sigma$ . They propose a reform if the following utility difference is positive:

$$R_i(a_i, \alpha_i, \alpha_j, \rho) = E[U_i(a_i, x_i = 1)|\alpha_i, \alpha_j, \sigma] - E[U_i(a_i, x_i = 0)|\alpha_i, \alpha_j, \sigma].$$

We refer to this utility difference as the reform incentive function  $R_i$ .

**Proposition 2.** *D1 equilibria always exist and exhibit symmetric cutoffs  $\alpha_i = \alpha_j = \alpha$  smaller than  $c$ .*

There always exist perfect Bayesian equilibria with cutoffs smaller than one ( $\alpha \in [0, 1)$ ). Intuitively, high types always want to reform, if low ability types play the status quo. In this situation, voters associate a high welfare contribution with a reform and high ability types gain from a higher vote share as well as from their expected welfare contribution. Hence, there only two possible situations. If office motivation is high, low able candidates have strong incentives to mimic more able ones and it is an equilibrium that all candidates propose a reform ( $\alpha = 0$ ). If in contrast office motivation is lower, candidates with low abilities prefer to play the status quo whenever voters expect all candidates to propose a reform. In this case  $\alpha = 0$  is thus not an equilibrium. However, in such a situation continuity of the reform incentive implies that there is at least one interior ability type who is indifferent between proposing a reform and the status quo if voters expect him to be the cutoff type.

These equilibria clearly satisfy the D1 criterion.  $\alpha \in (0, 1)$  implies that all actions are played in equilibrium and every perfect Bayesian equilibrium with this property satisfies D1. For  $\alpha = 0$ , even the least able candidates propose a reform. If any candidate deviates to the status quo, his associated welfare contribution is always zero independently of his ability. Out-of-equilibrium beliefs about abilities are thus irrelevant for the voting decision.<sup>8</sup>

---

<sup>8</sup> There can be perfect Bayesian equilibria with at least one cutoff being equal to one. These however, do not satisfy the D1 criterion, because they base on very pessimistic beliefs about reforming candidates.

To see that the equilibrium cutoff is characterized by  $\alpha \leq c$ , note that a reform proposal is always associated with a positive welfare contribution. Hence, a candidates' expected vote share increases if he proposes a reform. If voters associated a negative welfare contribution with a reform proposal, candidates with ability below  $c$  would never propose a reform. Otherwise, they would suffer from a negative welfare contribution as well as from a loss in expected office utility. Clearly, this is a contradiction: If only candidates with ability above  $c$  were to propose a reform, the associated welfare contribution could not be negative. Hence, a reform is associated with a higher vote share than the status quo. It follows that candidates with ability above  $c$  always choose to reform. They gain not only from their positive welfare contribution, but also from an increase in expected office rewards. Thus, the equilibrium cutoffs must be below  $c$ . The cutoff type is hence willing to make a negative welfare contribution to increase his chances to enter office.

**Definition 2.** *An equilibrium with  $\alpha > 0$  is an informative equilibrium.*

Next, we derive the equilibrium condition for informative equilibria.<sup>9</sup> In any informative equilibrium, candidates with high ability choose to propose a reform while candidates with low ability choose the status quo. Policy proposals hence represent informative signals about the candidates' abilities. The cutoff type is indifferent between proposing a reform or the status quo. Thus, the equilibrium cutoff  $\alpha$  is implicitly defined by  $R_i(\alpha, \alpha, \alpha, \rho) = 0$ . The resulting equilibrium condition is

$$R(\alpha, \rho) = \underbrace{\theta \left( f(v^r, \rho) - \frac{1}{2} \right)}_{\text{Change in office utility}} + \underbrace{\left[ \frac{1}{2} + \Phi(\alpha) \left( f(v^r, \rho) - \frac{1}{2} \right) \right]}_{\text{Change in welfare contribution}} (\alpha - c) = 0, \quad (4)$$

where  $v^r$  represents the vote share of a reforming candidate, when facing an opponent who proposes the status quo.

Equation (4) separates both aspects of the politicians' preferences. The change in office utility reflects the expected increase in office rewards due to a reform proposal. This term is always positive, since a reform proposal is associated with a higher vote share. However, the politician also cares about his welfare contribution. The change in welfare contribution stands for his influence on the implemented compromise if he proposes a reform. Note that this term represents a loss, since the ability of the cutoff type  $\alpha$  is below the cost  $c$ . In the next step, we establish the uniqueness of D1 equilibria.

---

<sup>9</sup>We discuss the effects of power dispersion in non-informative equilibria in Subsection 8.1.



**Assumption 1.** *The ability distribution  $\phi(a)$  is bounded from above with  $\phi(a) < \frac{1+\Phi(a)}{c-a}$  for all  $a < c$ .*

Assumption 1 is a regularity assumption on the ability distribution, which is fulfilled, e.g., for the uniform distribution. It ensures that the reform incentive is monotonically increasing in the cutoff. Hence, both incentive functions cannot intersect more than once and there is a unique equilibrium. For the remainder of the paper, we take Assumption 1 as given.<sup>10</sup>

**Proposition 3.** *There is a unique D1 equilibrium. Moreover, it exists  $\tilde{\theta}(\rho) \in \mathbb{R}^+ \cup \{\infty\}$  such that this equilibrium is informative if and only if  $\theta < \tilde{\theta}(\rho)$ .*

For any given level of power concentration, the existence of informative equilibria depends on the level of office motivation. If office motivation is not too large, i.e.,  $\theta < \tilde{\theta}(\rho)$ , cutoffs are larger than zero and the unique D1 equilibrium is informative. Otherwise, all candidates propose a reform, and the equilibrium is non-informative. If the average ability is smaller than the costs, the unique equilibrium is informative for all  $\theta < \infty$ . In this case, a reform proposal by all candidates would imply a negative welfare contribution from a reform, which cannot be true in equilibrium.

## 6 Effects of power-concentrating institutions

Empirically, democratic countries differ strongly in their political institutions and the implied power concentration. As we have argued in Subsection 3.3, our framework allows to represent these differences by means of an appropriate power allocation function  $f(v_i, \rho)$ . In this section, we study the effects of variations in power concentration  $\rho$ .

### 6.1 Effects on candidates' behavior

The power allocation function  $f$  determines the electoral incentives of political candidates. Under perfect information, variations in power concentration leave the behavior of candidates unaffected and policy choice is efficient (see Proposition 1).

With asymmetric information and office-motivated candidates, in contrast, policy choice is distorted as some low-ability candidates propose welfare-reducing reforms. Political institutions affect the magnitude of these policy distortions.

---

<sup>10</sup>If Assumption 1 is not given, multiple equilibria may arise. The following analysis is still valid, if we restrict our attention to the welfare optimal equilibrium.

**Proposition 4.** *In any informative equilibrium, increasing power concentration  $\rho$  leads to the proposal of more inefficient reforms:  $\frac{d\alpha}{d\rho} < 0$ .*

Consider some level of power concentration  $\rho_0$ . By construction, the cutoff type with ability  $a_i = \alpha_0 < c$  is indifferent between proposing the reform and the status quo. We find that after an increase in power concentration, the cutoff type strictly prefers to propose a reform. In particular, his utility of proposing the status quo decreases while his utility of proposing a reform increases.

If the cutoff type proposes the status quo, his welfare contribution is equal to zero. Thus, he only draws utility from office rents. With increasing power concentration, the office rents are reduced because he receives less power when running against a reforming opponent.

If the cutoff type proposes a reform, he again receives office rents, but also incurs a utility loss due to his negative welfare contribution. His overall utility is given by the sum of these two components. This sum is positive because it must be equal to the utility from a status quo proposal. With increasing power concentration, both the office rents and the negative welfare contribution increase by the same factor. Hence, his utility from a reform proposal also increases by this factor.

Consequently, with higher levels of power concentration, status quo proposals yield lower utility while reform proposals become more attractive. The equilibrium cutoff thus decreases with the level of power concentration.

## 6.2 Effects on welfare

In the following, we study the effects of power-concentrating institutions on ex ante welfare. With privately informed candidates, the relation between power concentration and welfare is not as clear-cut as under perfect information.

On the one hand, there is still a positive *empowerment effect* of power concentration. Whenever both policies are proposed, the majority of votes goes to the reforming candidate, who provides higher expected welfare than the candidate proposing the status quo (see Section 5). Consequently, any increase in power concentration  $\rho$  assigns more power to the appropriate candidate.

On the other hand, the previous section demonstrated a negative *behavioral effect* of power concentration. By reinforcing the electoral stakes, stronger concentration of power induces the proposal of more inefficient reforms. This reduces the information revealed during the campaigns and limits the voters' capacity to allocate power to high-ability candidates.

We assume that the following regularity condition is satisfied.

**Assumption 2.** *The ability distribution has a non-decreasing reversed hazard rate  $\Phi(a)/\phi(a)$ .*

**Lemma 3.** *The welfare function  $W$  is strictly quasi-concave in  $\rho$ .*

Lemma 3 implies that the welfare function has a unique maximum in  $\rho$ . Its proof involves analyzing how power concentration influences the empowerment effect and the behavioral effect.

First, consider the positive empowerment effect. With increasing  $\rho$ , a reforming candidate receives more power if he runs against an opponent proposing the status quo. The average reform payoff determines how much welfare is increased by this reallocation of power. An increase in power concentration also induces more inefficient reforms and thus diminishes the average reform payoff. Consequently, the empowerment effect is strictly decreasing in  $\rho$ , as illustrated by the solid line in Figure 2.

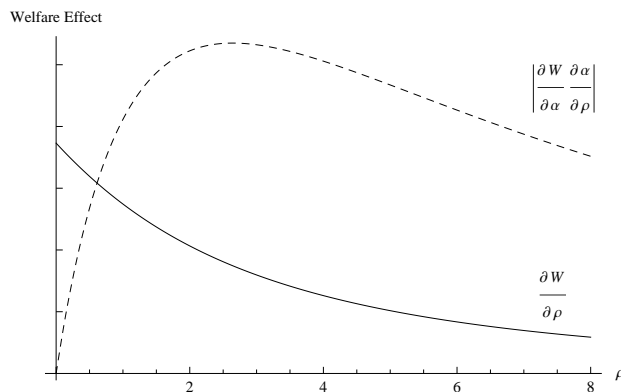


Figure 2: *The welfare effects of a change in  $\rho$  for a logistic power allocation function with mean  $\mu = 0.5$  and scale parameter  $\beta = 1/\rho$ ,  $\mu_k$  distributed according to  $\mathcal{N}(0, 0.5)$ , a uniform ability distribution,  $\theta = 1$ , and  $c = 0.6$ . The solid line represents the (positive) empowerment effect, the dashed line represents the (negative) behavioral effect. The optimal level of  $\rho$  is attained at the intersection of both lines.*

Second, consider the negative behavioral effect. It results because increasing power concentration leads to a reduction in the cutoff  $\alpha$ . The size of this effect depends on, first, the marginal welfare loss from a decline in  $\alpha$ , and second, the sensitivity of  $\alpha$  with respect to changes in power concentration. Both factors are affected differently by increasing power allocation.

Regarding the first factor, higher power concentration induces the cutoff to depart further from its efficient level  $c$ . As increasingly inefficient reforms are proposed, the marginal welfare loss from reductions in  $\alpha$  increases with  $\rho$ . Regarding the second factor, the sensitivity of  $\alpha$  depends on the additional vote share a candidate gains by proposing a reform, which is directly related to the average reform payoff.

For higher levels of power concentration, the average reform payoff becomes smaller and so does the additional vote share. Thus, higher levels of  $\rho$  come along with a reduced sensitivity of  $\alpha$ , which attenuates the behavioral effect. As a consequence, the behavioral effect is non-monotonic in  $\rho$  (see the dashed line in Figure 2).

The sign of the overall effect of power concentration on welfare depends on the relative sizes of both effects. Under the monotone hazard rate assumption, the ratio of empowerment effect and behavioral effect is strictly increasing in  $\alpha$  and decreasing in  $\rho$  at every local extremum, as we show in Appendix A. Thus, the welfare function cannot have an interior minimum and at most one interior maximum in  $\rho$ , which corresponds to the definition of quasi-concavity.

**Proposition 5.** *If and only if office motivation is below some threshold level  $\bar{\theta}$ , welfare is maximized by full concentration of power. If instead  $\theta > \bar{\theta}$ , it is optimal to disperse power,  $\rho^*(\theta) \in (0, \infty)$ , and the optimal concentration of power is strictly decreasing in the candidates' office motivation,  $\frac{d\rho^*}{d\theta} < 0$ .*

Proposition 5 establishes a relation between candidates' motivation and the optimal level of power concentration. Intuitively, higher office motivation makes mimicking more attractive and induces more inefficient reforms. Allocating power to reforming candidates is consequently less beneficial, so that the positive empowerment effect decreases in  $\theta$ . Furthermore, higher office motivation reinforces the negative behavioral effect, since candidates respond more strongly to the electoral incentives.

Regarding the optimal constitution, we have to distinguish two cases. First, consider the case of mainly policy-oriented candidates,  $\theta < \bar{\theta}$ , in which mimicking is not prevalent and the average reform payoff is large. In this case, the negative behavioral effect is sufficiently small to be dominated by the positive empowerment effect for all levels of  $\rho$ . Consequently, welfare is maximized by full concentration of power. Second, consider the case of mainly office-motivated candidates,  $\theta > \bar{\theta}$ , in which mimicking is widespread. Hence, the behavioral effect is reinforced relative to the empowerment effect. It is then optimal to attenuate electoral incentives by decreasing power concentration. Both effects outbalance each other at some interior level  $\rho^* \in (0, \infty)$  that represents the optimal institution. By the same logic, the optimal level of power concentration is reduced with any further increase in the level of office motivation.

So far we have established the optimal democratic institution. However, our model also allows to investigate whether democratic selection of political leaders is desirable at all. A similar question has been addressed by Maskin & Tirole (2004), who compare decision-making by accountable "politicians" and non-accountable

“judges”. A non-democratic regime allocates all political power to a randomly chosen dictator. While such a regime obviously rules out selection, it also eliminates incentives for inefficient policy choice. In our model, this non-democratic system yields the same welfare as the limiting case of a democratic system with fully dispersed power. Consequently, Proposition 5 implies that democratic systems with appropriately chosen power concentration always dominate the non-democratic alternative. This contrasts with the result of Maskin & Tirole (2004) that, under certain circumstances, political decisions should rather be delegated to “judges” than to “politicians”.

## 7 Empirical analysis

In this section, we analyze whether data for established democracies support our model predictions. Proposition 5 states that power concentration is always conducive to the implementation of efficient policies if politicians exhibit low levels of office motivation,  $\theta < \bar{\theta}$ . At higher levels of office motivation, in contrast, it is optimal to disperse power. Moreover, the optimal degree of power concentration declines for further increases in office motivation. The implications of our model can be summarized in the following Hypothesis.

**Hypothesis.** *The effect of power concentration on welfare depends on the level of politicians’ office motivation. Power concentration has positive effects on welfare if politicians are mainly policy-motivated. In contrast, if politicians are mainly office-motivated, the welfare effect of power concentration is significantly smaller or even negative.*

We face several restrictions to data availability. In particular, there exists yet no objective measure for politicians’ office motivation or for the ability of empowered politicians. A rigorous empirical test is hence infeasible. Nevertheless, we propose an empirical strategy to illustrate the consistency of our model predictions with the data.

### 7.1 Operationalization

The empirical analysis requires three basic measures. As the dependent variable, we need a measure of efficient policies. Key independent variables are the degree of power concentration within the political system and the extent of politicians’ office motivation.

As a measure for *efficient policies*, we use growth in real GDP per capita (World Bank). It provides a concise and objective measure of developments that bear the potential of welfare improvements for the general public. Growth has been used as outcome variable by a number of other empirical studies on political institutions as Feld & Voigt (2003) and Enikolopov & Zhuravskaya (2007). Other frequently used outcome measures relate to fiscal policy (see Voigt 2011), which is not addressed in our model.

Several measures of democratic institutions have been discussed in the literature. Lijphart’s index of the executive-parties dimension captures the *concentration of power* that is implied by the set of political institutions (Lijphart 1999). This well-established measure quantifies how easily a single party can take complete control of the government. Low values of the index correspond to high concentration of power within the political system. Index values are provided for 36 economically developed countries with a long democratic tradition. The measure is based on the period 1945-1996. New Zealand underwent major constitutional changes after 1996 and is thus excluded from the analysis. Its inclusion, however, does not change the qualitative results.

While *office motivation* cannot be measured objectively, indication for it may come from voter surveys. The International Social Survey Programme (ISSP) includes questions on voters’ opinions about politicians.<sup>11</sup> The item relevant to our study was included in its 2004 survey (ISSP Research Group 2012), which was performed in most democratic states: “Most politicians are in politics only for what they can get out of it personally.” Agreement with this statement is coded on a five point scale. We use the mean points of all survey participants in a country as our measure for the importance of office motivation. That means we assume that differences in this item reflect differences in politicians’ motives.<sup>12</sup>

For an easy interpretation of regression results, we normalize the indices for both office motivation and power concentration to range between zero and one. High values indicate pronounced office motivation of political leaders or a strong concentration of political power, respectively.

---

<sup>11</sup>Other surveys such as the World Values Survey, the Global Barometer Survey, the Eurobarometer, or the European Value Survey query trust or confidence in institutions, such as the political parties and the national parliament. Such questions only indirectly relate to politicians’ motivation.

<sup>12</sup>Alternatively, one could use measures that are based on experts’ assessments like the Corruption Perception Index from Transparency International and the Worldwide Governance Indicators from Kaufmann et al. (2009). However, these indices focus on rent extraction and not on private motivations of politicians in general.

## 7.2 Design

Our analysis focuses on countries with a similar degree of democratization. We require that all countries be established democracies as identified by the 2002 Polity IV Constitutional Democracy index (Marshall & Jaggers 2010). All countries have to feature an index of 95 or higher, which excludes Venezuela from the sample. The remaining 18 countries in the sample are similar with respect to their economic characteristics. In particular, all countries are economically highly developed, as classified by the World Bank. They furthermore feature a Human Development Index (HDI) of at least 0.9 in 2004, which places them in the top quintile of all countries.<sup>13</sup>

The time-invariant regressors require a cross-country analysis. All explanatory variables correspond to 2004 or earlier years. To address problems of reverse causality, our explained variable captures growth after 2004. I.e., the explanatory variables cannot be affected by our explained variable. To test for an interaction effect between our main explanatory variables, power concentration and office motivation, we include an interaction term between both variables in the regression.<sup>14</sup> We control for variables that may be correlated with both our explanatory variables and our explained variable. Most notably, past economic performance affects growth (see, e.g., Barro 1991, Sala-i-Martin 1994) and may also alter voters' perception of politicians. We hence control for GDP per capita in 2004. Also other variables have been shown to robustly affect growth, such as capital accumulation, school enrollment rates, life expectancy, or openness of the economy (see, e.g., Sala-i-Martin 1997). To capture such influences and to keep the number of explanatory variables low, we add past growth in real GDP per capita (from 1991 to 2004) to the regression.<sup>15</sup> Controlling for past economic performance, ensures that effects of the voters' perception of politicians do not capture the quality of past economic outcomes.

## 7.3 Results

To broadly describe the data, we split the country set at the median value of politicians' office motivation. Figure 3 plots growth against concentration of power for the two sets of countries. The left panel depicts the relationship for countries in

---

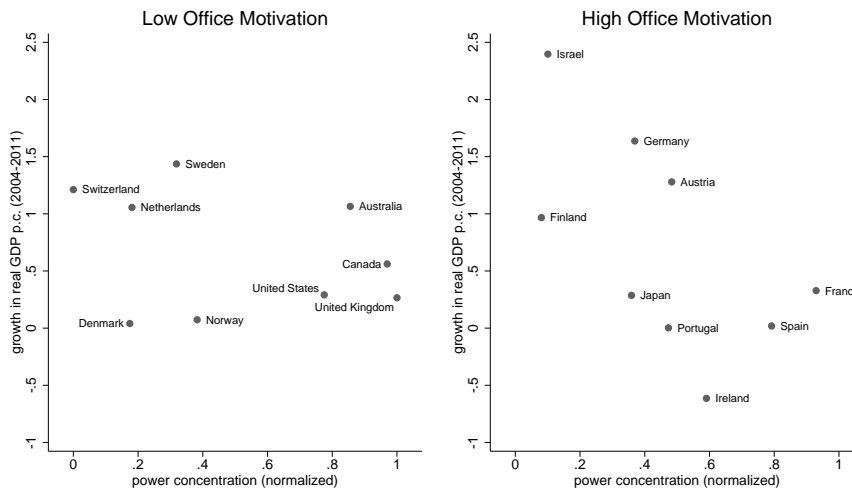
<sup>13</sup>The similarity in socioeconomic development was formulated as a major prerequisite for cross-country analyses in Armingeon (2002).

<sup>14</sup>The analysis of an interaction effect can be problematic if the interacting variables are highly correlated. However, we find no correlation between power dispersion and office motivation (Pearson's correlation coefficient  $\rho = -0.199$ ,  $p = 0.427$ ). This also suggests that historically developed political institutions within a country do not exhibit the optimal level of power concentration.

<sup>15</sup>Descriptive statistics for all variables are provided in Appendix C.

which politicians' office motivation is below its median value, while the right panel depicts the relationship for countries in which politicians' office motivation is above its median value. The figure suggests that power concentration has only small effects if politicians are mainly policy-motivated, whereas power concentration is harmful for growth if politicians are mainly office-motivated. For both groups of countries, the bivariate correlations between power concentration and growth support this observation.<sup>16</sup>

Figure 3: Relationship between power concentration and growth



For a statistical test of the effects of power concentration on economic growth, we use office motivation as a continuous explanatory variable in an OLS regression. We also control for relevant covariates. Table 1 presents the regression results. Test statistics are based on White heteroscedasticity-consistent standard errors.

Column (a) displays the results of a regression model without interaction term. In this regression, the coefficient of power concentration estimates the effect on economic growth under the assumption that this effect does not depend on the level of office motivation. We find that this coefficient is insignificant.

This picture changes if the interplay between power concentration and politicians' motivation is taken into account. Column (b) reports the corresponding regression results. Most importantly, the coefficient of the interaction term between power concentration and office motivation is negative and significant. Thus, power concentration is more negatively related to growth, the more office-motivated politicians

<sup>16</sup>For countries with high levels of politicians' office motivation, there is a negative and weakly significant relationship between growth and power concentration (Pearson's correlation coefficient,  $\rho = -0.618$ ,  $p = 0.076$ ), while there is no significant relationship between the two variables for countries with low levels of politicians' office motivation (Pearson's correlation coefficient,  $\rho = -0.291$ ,  $p = 0.447$ ).



Table 1: OLS regression results

	Growth in real GDP per capita (2004-2011)	
	(a)	(b)
Power concentration	-0.852 (0.490)	3.565* (1.637)
Office motivation	-0.125 (1.090)	2.566* (1.182)
Power concentration · office motivation		<b>-8.948**</b> (3.520)
Real GDP per capita in 2004 (in \$ 1000)	-0.025 (0.025)	-0.0436* (0.021)
Growth in GDP per capita (1991-2004)	-0.267*** (0.079)	-0.352*** (0.059)
Constant	2.382** (1.090)	1.845** (0.806)
adjusted $R^2$	0.19	0.49
$F$	4.38	14.17
$N$	18	18

Standard errors are provided in brackets. \*\*\*, \*\*, \* indicate significance at the 1-, 5-, and 10-percent level, respectively.

are. The inclusion of the interaction term in the regression also strongly increases the explanatory power of the econometric model. The adjusted  $R^2$  increases from 0.19 to 0.49, even though no additional information is used.

Table 2: Effect of power concentration

	lowest office motivation	highest office motivation
Coefficient	3.565*	-5.382**
Standard error	1.637	1.955

The table depicts the coefficient of power concentration for the lowest level of office motivation ( $\theta = 0$ ) and for the highest level of high office motivation ( $\theta = 1$ ). \*\*\*, \*\*, \* indicate significance at the 1-, 5-, and 10-percent level, respectively.

As it turns out, power concentration comes along with either increased or decreased growth prospects depending on the level of politicians' office motivation. The conditional effect of power concentration at the lowest and the highest level of office motivation in our country set are reported in Table 2. At the lowest level of office motivation, power concentration is positively related to growth. In contrast,

this relationship is negative at the highest level of office motivation. Our analysis thus leads to the following result.

**Result.** *The higher is office motivation, the more negative is the relation between power concentration and growth. Furthermore, power concentration is negatively related to growth if politicians' office motivation is high. If politicians are mainly policy-motivated, power concentration comes along with increased growth.*

We conclude that the data is in line with the model presented in this paper. We do not only observe a negative and significant interaction effect, but also that the effect of power concentration changes its sign as suggested by the theory. Moreover, taking this interaction effect into account increases explanatory power considerably. In the analysis of political institutions, neglecting the interplay between power concentration and politicians' office motivation thus conceals actual patterns and yields misleading conclusions.

## 7.4 Discussion of empirical results

We conduct several robustness checks for our empirical analysis. In the following, we discuss the use of different indicators for our main variables, a possible impact of the financial crisis on our results, and an alternative explanation for our result.

First, we check whether the negative and significant interaction term between power concentration and politicians' office motivation is robust to the use of different measures for our key variables. Instead of politicians' motivation from the ISSP, we also use confidence in political parties as contained in the third wave of the World Values Survey (WVS) concluded in 1998 (WVS 2009). Using this measure and adjusting the GDP and growth variables to the survey date, the interaction effect remains negative and significant ( $p=0.009$ ,  $F=1141.31$ ,  $N=10$ ). Unfortunately, the set of countries covered both by the third wave of the WVS as well as by Lijphart is small. Other surveys on politicians' office motivation have been conducted only very recently and are thus not applicable within our research design.

The measure for power concentration by Lijphart (1999) is available in a more current version from Armingeon et al. (2011). The use of this indicator yields a highly significant interaction term ( $p=0.009$ ,  $F=9.95$ ,  $N=17$ ). Armingeon et al. (2011) also provide a modified index that focuses on institutional factors only. It is based on the variables "electoral disproportionality" and "number of parties" and is invariant to behavioral factors such as "absence of minimal winning coalitions" included in the original index. Using this measure instead, the results remain significant ( $p=0.061$ ,  $F=15.99$ ,  $N=17$ ). We also use three different measures that

capture important aspects of power concentration and find similar patterns. For the index for checks and balances (Keefer & Stasavage 2003) and a plurality electoral system dummy (Beck et al. 2001), the interaction effect shows the expected sign ( $p = 0.046, F = 17.26, N = 18$  and  $p = 0.087, F = 8.91, N = 18$ , respectively). For the nine-categorical type of electoral system (IDEA 2004), however, the coefficient is insignificant ( $p = 0.159, F = 8.53, N = 18$ ).

Second, one might fear that our result is influenced by the financial crisis which affected output beginning in 2008. To ensure that the financial crisis does not drive patterns in the data, we may restrict explained GDP growth to the years 2004-2007. For this shorter period, the interaction term between power concentration and office motivation remains weakly significant ( $p=0.092, F=4.20, N=18$ ). Using the World Values Survey for our measure of politicians' office motivation we can expand explained GDP growth to the years 1998-2007. These data provide a similar picture ( $p=0.062, F=127.54, N=10$ ). An alternative approach to deal with the financial crisis is to exclude countries that were particularly affected. The result is robust to the exclusion of any one country from the analysis (all p-values below 0.068, F above 4.82) and to the exclusion of any subset of the countries Ireland, Spain, and Portugal (all p-values below 0.031, F above 4.82), which were hit most severely by the financial crisis.

Finally, we discuss whether the empirical result could be explained by a different channel. It could be argued that the main role of political institutions is to discipline rent-seeking politicians. In particular, power-dispersing institutions may restrict rent extraction in office, which would be more important the more politicians value rents. However, the altered behavior of politicians would also affect the possibility to screen politicians in office and to reelect only good ones. Empirically, we cannot distinguish between our explanation and this alternative, since measures for politicians' office motivation may capture not only preferences for power per se, but also for rent extraction. Theoretically, however, this alternative channel has been discussed in a post-election model by Besley & Smart (2007). They investigate the effects of four fiscal restraints that limit the office holders' discretion. Three of these constraints enhance welfare only if the share of benevolent politicians is sufficiently large. This suggests that power dispersion enhances welfare only if office motivation is low, which is in contrast to our model and the empirical findings.

## 8 Modifications and extensions

Our model is flexible enough to incorporate several modifications and extensions. In a first step, we extend the analysis to non-informative equilibria. In a second step, we generalize the setting in three dimensions. First, we introduce a continuous policy space by assuming that candidates may propose the magnitude of reform rather than limiting their choice set to a complete reform and the status quo. Second, we show that limited commitment, i.e., the possibility to withdraw a proposal after the election with a certain probability, does not change the results. Third, we allow for heterogeneous policy preferences of voters in the sense that reforms may benefit some voters and harm others. None of these modifications alter the qualitative results of the model.

### 8.1 Non-informative equilibria

In the main part of the paper, we focus on the effects of power dispersion in informative equilibria. However, changes in power concentration may also allow to move from a non-informative to an informative equilibrium, thereby increasing welfare.

Proposition 3 establishes that the unique D1 equilibrium is informative if and only if  $\theta$  is below the threshold  $\tilde{\theta}(\rho)$ . If instead  $\theta \geq \tilde{\theta}(\rho)$ , the unique equilibrium satisfying D1 is non-informative: Even the least able candidate proposes a reform, although it will fail with certainty ( $\alpha = 0$ ). However,  $\tilde{\theta}(\rho)$  is strictly decreasing in  $\rho$  and approaches infinity for  $\rho \rightarrow 0$ . Thus, it is always possible to move from non-informative equilibria to informative ones by implementing political institutions that induce more power dispersion.

**Proposition 6.** *For any level of  $\theta$ , there exists  $\tilde{\rho}(\theta) > 0$  such that the unique D1 equilibrium is informative if and only if power is sufficiently dispersed,  $\rho < \tilde{\rho}(\theta)$ . All informative equilibria strictly welfare-dominate non-informative equilibria.*

For the last statement, note that in any informative equilibrium low ability candidates with  $a_i \in [0, \alpha)$  refrain from proposing inefficient reforms. Hence, these candidates receive less power, while more power is allocated to candidates with  $a_i \geq \alpha$ , who provide on average a positive welfare contribution.

### 8.2 Continuous policy space

Until now, we have assumed that candidates can either propose a reform or the status quo. However, many policy decisions are inherently continuous and politicians can choose “how much” of a reform to implement. Suppose that the action space of

the candidates is  $x_i \in [0, 1]$ , with  $x_i$  representing the proposed share of a complete reform. As before, the welfare contribution of candidate  $i$  is given by  $x_i(a_i - c)$ .

Given the continuous policy space, it is possible to assume that the implemented policy is a compromise between the candidates' agendas instead of a lottery between the proposals. Then, candidates with larger amounts of power  $f(v_i, \rho)$  are able to enforce larger parts of their agenda, while candidates with less power only slightly influence the political decision.

**Proposition 7.** *Let the action space of candidates be continuous with  $x_i \in [0, 1]$ . There is a unique D1 equilibrium, which is outcome equivalent to the one resulting for a binary action space. As a consequence, Propositions 1 to 5 hold.*

Since reform incentives are still monotonically increasing in the ability of the candidates, so is the magnitude of reform they propose in equilibrium. D1 yields that only complete reforms or the status quo are played in equilibrium. To see this, first note that complete reforms are always proposed in a D1 equilibrium. If this was not the case, D1 would require that a deviation to a complete reform would have to be attributed to the most able candidate. Given this belief, a complete reform would yield a profitable deviation for the most able candidates. As a consequence, all agents with ability above  $c$  propose a complete reform instead of proposing only a share of a reform, since they profit from the higher welfare contribution as well as from the higher office utility generated by this proposal. If an intermediate reform is proposed in equilibrium, voters thus associate it with a negative welfare contribution. Consequently, such a proposal leads to a smaller vote share than the status quo proposal, and candidates below  $c$  strictly prefer the latter. Overall, only the two extremes of the action space are played in a D1 equilibrium.

We conclude that restricting the action space of candidates to the status quo policy and a complete reform does not impose a loss of generality. The unique equilibrium satisfying D1 is outcome equivalent in the sense that all agents choose the same actions and hold the same beliefs on the equilibrium path. All proofs carry over to this setup.

### 8.3 Limited commitment

The assumption of full commitment is widely used to ensure tractability of models (see, e.g., Persson & Tabellini 2003). However, it may seem too restrictive that politicians can never change or adapt their agenda. In our setting, candidates with ability lower than  $c$  have an incentive to withdraw a reform proposal when they gain power. A straightforward way to introduce limited commitment into the model is

to assume that, with probability  $\lambda > 0$ , the environment changes after the election and politicians may deviate from their proposal. For example, this could be due to an unexpected shock in the policy field or a major event in another policy field. With probability  $1 - \lambda$ , on the contrary, they have to carry out their proposal.

**Proposition 8.** *Suppose policy proposals are binding with probability  $\lambda$ . Then Propositions 1 to 5 continue to hold.*

This form of limited commitment increases incentives to propose a reform for low ability candidates, since they may be able to withdraw their proposal after the election. However, this only affects the level of equilibrium cutoffs and not the qualitative results.

Note that the welfare effect of reduced commitment is ambiguous. On the one hand, all candidates with ability  $a_i < c$  withdraw their reforms with probability  $\lambda$ , thereby increasing welfare. On the other hand, as limited commitment diminishes the negative welfare contribution of a reform proposal for low ability candidates, more inefficient reforms are proposed. Thus, reform proposals become less informative to the voters, and high-ability candidates receive less political power. The worse selection of politicians as well as the more inefficient reform proposals per se represent negative effects on welfare.

## 8.4 Heterogeneous preferences

In political philosophy as well as public debate, a major virtue of power dispersion is seen in the political representation of minorities and the prevention of a tyranny of the majority. For example, James Madison argues in the Federalist #51 that "the rights of the minority will be insecure" without proper checks and balances (Madison 1788a). So far, our analysis has abstracted from this aspect of political institutions in order to emphasize effects of power dispersion that are independent of minority rights.

To incorporate heterogeneity in voters' policy preferences into our model, we may assume that voters differ in their benefit from a reform rather than in their candidate preferences. In particular, voter  $k$  receives a payoff of  $\mu_k$  if a reform is successfully implemented. Let the preference parameter  $\mu_k$  be symmetrically distributed according to the pdf  $\xi(\mu)$  and the cdf  $\Xi(\mu)$  with full support on some interval  $[\underline{\mu}, \bar{\mu}]$ . We assume that the mean preference is larger than the reform cost  $c$ , while  $\underline{\mu} \in (0, c)$ . This implies that a majority of voters is in favor of the reform, as long as it is adopted by a sufficiently able candidate, while a minority unambiguously prefers the status quo.

**Proposition 9.** *If the voters have heterogeneous policy preferences according to distribution  $\Xi(\mu)$ , Propositions 1 to 5 continue to hold.*

Essentially, the proofs for all previous results hold whenever the expected vote share of a reforming candidate  $i$  is increasing in the average ability of candidates that propose a reform, i.e., in the equilibrium cutoff  $\alpha_i$ . The basic model can be seen as the special case with a degenerate distribution function with  $\mu_k = 1$  for all voters.<sup>17</sup>

Given these heterogeneous policy preferences, our model allows to reconsider Madison’s conjecture. Increasing power dispersion leads to higher amounts of power for candidates proposing the status quo, which is the minority’s preferred option. As a consequence, the status quo is proposed more often yielding an additional increase in the minority’s welfare.

**Lemma 4.** *In any informative equilibrium, the utility of each minority voter  $k$  with  $\mu_k \leq c$  is strictly decreasing in the concentration of political power.*

The quote above suggests that the Founding Fathers of the United States were interested in the protection of minority rights per se. Formally, this objective can be captured by introducing inequality aversion into the welfare function, using a strictly increasing, strictly concave, and twice continuously differentiable weighting function  $w$ :

$$W_{IA} = \int_{\underline{\mu}}^{\bar{\mu}} w(V(\mu_k, \rho)) \xi(\mu_k) d\mu_k.$$

In this function,  $V(\mu_k, \rho)$  represents the expected utility of a voter with preference  $\mu_k$ . Following Atkinson (1973) and Hellwig (2005), the relative curvature of  $w$  can be interpreted as a measure of inequality aversion. Compared to the inequality-neutral welfare function,  $W_{IA}$  puts higher weights on voters with low expected utility.

**Proposition 10.** *Any welfare function  $W_{IA}$  with inequality aversion is maximized at a lower level of power concentration than the inequality-neutral function  $W$ .*

Intuitively, power-dispersing institutions reduce the discretion of the election winner, who is chosen by the majority. The expected utility of the majority of

---

<sup>17</sup>Our model also allows for additional (ideological) heterogeneity with respect to the candidates. Let the reforms advocated by both candidates be targeted towards different groups of voters and let  $\mu_{ki}$  denote the payoff to voter  $k$  from a successful reform by candidate  $i$ . If both parameters share the unconditional distribution  $\Xi(\mu)$  defined above, Proposition 9 continues to hold for any correlation between  $\mu_{k1}$  and  $\mu_{k2}$ . With negative correlation, the candidates’ reform proposals differ strongly or are even diametrically opposed (as in a stylized left-right policy space).

voters is hence reduced while the minority is better off. The utility of the minority is valued strongly by an inequality averse constitutional designer. Thus, he will choose to disperse power more strongly than if he was inequality-neutral.<sup>18</sup>

## 8.5 Heterogeneous office motivation

Throughout the paper, we have used that the degree of politicians' office motivation is reflected by a level of office motivation that is identical for all politicians. The following extension establishes that our results continue to hold if the degree of politicians' office motivation is instead reflected by the fraction  $\mu$  of politicians that are motivated by office as opposed to motivated by welfare.

Let a fraction  $\mu$  of candidates be office-motivated with  $\theta_h > \bar{\theta}$ , while a fraction  $1 - \mu$  is purely welfare-motivated with  $\theta_l = 0$ . Moreover, suppose Assumption 2 holds for the weighted distribution  $K(a) = (1 - \mu)\Phi(c) + \mu\Phi(a)$ . The general features of the equilibrium convey to the new setup.

**Proposition 11.** *If there are low and high motivation politicians with  $\theta_l = 0$  and  $\theta_h > \bar{\theta}$ , Propositions 1 to 4 continue to hold.*

We can determine how the optimal allocation of power is affected by the composition of the set of political candidates.

**Proposition 12.** *There exists a  $\bar{\mu}$  such that for all  $\mu < \bar{\mu}$ , welfare is maximized by full concentration of power. If instead  $\mu > \bar{\mu}$ , it is optimal to disperse power,  $\rho^*(\mu) \in (0, \infty)$ , and the optimal concentration of power is strictly decreasing in the fraction of candidates with high office motivation,  $\frac{d\rho^*}{d\mu} < 0$ .*

The intuition from our main model carries over to this setup. A higher fraction of office-motivated candidates  $\mu$  causes more distortions in policy choice. The positive empowerment effect is reduced, as reforming candidates on average provide lower welfare. Furthermore, the negative behavioral effect is reinforced, since more candidates respond inefficiently to the electoral incentives.

## 9 Conclusion

We have investigated how the level of power concentration affects campaign behavior of politicians and social welfare if candidates are office-motivated and privately

---

<sup>18</sup>Note that  $W_{IA}$  is maximized at a strictly lower level than  $W$  for any  $\theta > \bar{\theta}$ . For the opposite case, even constitutional designers with small degrees of inequality aversion will prefer to concentrate power completely.



informed about their ability. Increasing the concentration of power has two effects. On the one hand, it has a positive *empowerment effect* because more power is given to election winners, who provide higher welfare in expectation. On the other hand, it also has a negative *behavioral effect*. Stronger concentration of political power reinforces the incentive for low-ability candidates to mimic more able ones. This limits the voters' capacity to identify and empower high-ability politicians.

The optimal institutional design balances both effects. We have established that the optimal level of power concentration is negatively related to the extent of office motivation. If politicians care mainly about welfare, power concentration yields strictly positive effects. In politicians are mainly office-motivated, on the contrary, welfare is maximized by institutions that divide power between election winner and loser. Intuitively, the concentration of power induces distortions in policy choice of office-motivated candidates. The more office-motivated the candidates are, the more beneficial it is to disperse power and thereby to reduce the distortions in policy choice.

In the empirical part, we have confronted these predictions with data for eighteen established democracies. Our findings are in line with the theoretically derived hypothesis. In a regression with economic growth as dependent variable, we find a negative and significant interaction effect between office motivation and power concentration. For the highest levels of office motivation, power-concentrating institutions come along with significantly lower economic growth, while we find a positive correlation for countries with the lowest levels of office motivation.

Our analysis provides insight into how political institutions affect the selection of competent candidates into office. Due to concerns about rent-seeking or corruption, voters may at the same time wish to select politicians with high degrees of policy motivation. Moreover, political institutions can also affect the pool of candidates that run for office. The analysis of the interaction between these three aspects of selection is a plausible next step in the study of political institutions.

## References

- Acemoglu, D., Egorov, G. & Sonin, K. (2010), 'Political Selection and Persistence of Bad Governments', *Quarterly Journal of Economics* **125**(4), 1511–1575.
- Adams, J. (1999), 'Policy Divergence in Multicandidate Probabilistic Spatial Voting', *Public Choice* **100**(1), 103–122.

- Ansolabehere, S. & Snyder, J. M. (2000), 'Valence politics and equilibrium in spatial election models', *Public Choice* **103**, 327–336.
- Aragones, E. & Palfrey, T. R. (2002), 'Mixed equilibrium in a downsian model with a favored candidate', *Journal of Economic Theory* **103**(1), 131–161.
- Armingeon, K. (2002), 'The Effects of Negotiation Democracy: A Comparative Analysis', *European Journal of Political Research* **41**, 81–105.
- Armingeon, K., Careja, R., Potolidis, P., Gerber, M. & Leimgruber, P. (2011), *Comparative Political Data Set III 1990-2009*, Institute of Political Science, University of Berne.
- Atkinson, A. (1973), How Progressive Should Income Tax Be?, in M. Parkin & R. Nobay, eds, 'Essays on Modern Economics', Longman, London.
- Barro, R. J. (1991), 'Economic Growth in a Cross Section of Countries', *Quarterly Journal of Economics* **106**(2), 407–43.
- Beck, T., Clarke, G., Groff, A., Keefer, P. & Walsh, P. (2001), 'New Tools in Comparative Political Economy: The Database of Political Institutions', *World Bank Economic Review* **15**(1), 165–176.
- Besley, T. (2005), 'Political Selection', *Journal of Economic Perspectives* **19**(3), 43–60.
- Besley, T. & Smart, M. (2007), 'Fiscal Restraints and Voter Welfare', *Journal of Public Economics* **91**(3-4), 755–773.
- Blume, L., Müller, J., Voigt, S. & Wolf, C. (2009), 'The Economic Effects of Constitutions: Replicating-and Extending-Persson and Tabellini', *Public Choice* **139**(1), 197–225.
- Caselli, F. & Morelli, M. (2004), 'Bad Politicians', *Journal of Public Economics* **88**(3-4), 759–782.
- Cho, I.-K. & Kreps, D. M. (1987), 'Signaling Games and Stable Equilibria', *Quarterly Journal of Economics* **102**(2), 179–222.
- Enikolopov, R. & Zhuravskaya, E. (2007), 'Decentralization and Political Institutions', *Journal of Public Economics* **91**(11-12), 2261–2290.

- Feld, L. P. & Voigt, S. (2003), 'Economic Growth and Judicial Independence: Cross-Country Evidence Using a New Set of Indicators', *European Journal of Political Economy* **19**(3), 497–527.
- Hellwig, M. F. (2005), 'A Utilitarian Approach to the Provision and Pricing of Excludable Public Goods', *Journal of Public Economics* **89**, 1981–2003.
- Iaryczower, M. & Mattozzi, A. (2008), Ideology and competence in alternative electoral systems, Levine's Working Paper Archive 122247000000002387, David K. Levine.
- IDEA (2004), *Handbook of Electoral System Design*, International Institute for Democracy and Electoral Assistance.
- ISSP Research Group (2012), 'International Social Survey Programme 2004: Citizenship I (ISSP 2004)', GESIS Data Archive, Cologne, ZA3950 Data file Version 1.3.0, doi: 10.4232/1.11372.
- Kaufmann, D., Kraay, A. & Mastruzzi, M. (2009), Governance matters VIII : Aggregate and individual governance indicators 1996-2008, World Bank Policy Research Working Paper No. 4978.
- Keefer, P. & Stasavage, D. (2003), 'The Limits of Delegation: Veto Players, Central Bank Independence and the Credibility of Monetary Policy', *American Political Science Review* **97**(3), 407–423.
- King, A. (2002), Do Leaders' Personalities Really Matter?, in A. King, ed., 'Leaders' Personalities and the Outcomes of Democratic Elections', Oxford University Press, Oxford, pp. 1–42.
- Krasa, S. & Polborn, M. (2011), A Political-Economy Model of Taxation and Government Expenditures with Differentiated Candidates, Working paper, The Selected Works of Mattias K Polborn.
- Lijphart, A. (1999), *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*, New Haven: Yale University Press.
- Lijphart, A. (2012), *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*, 2 edn, New Haven: Yale University Press.
- Lindbeck, A. & Weibull, J. W. (1987), 'Balanced-Budget Redistribution as the Outcome of Political Competition', *Public Choice* **52**(3), 273–297.

- Lizzeri, A. & Persico, N. (2001), ‘The Provision of Public Goods under Alternative Electoral Incentives’, *American Economic Review* **91**(1), 225–239.
- Lizzeri, A. & Persico, N. (2005), ‘A Drawback of Electoral Competition’, *Journal of the European Economic Association* **3**(6), 1318–1348.
- Madison, J. (1788*a*), ‘Federalist No. 51, The Structure of the Government Must Furnish the Proper Checks and Balances Between the Different Departments’, accessed through [http://thomas.loc.gov/home/histdox/fed\\_51.html](http://thomas.loc.gov/home/histdox/fed_51.html).
- Madison, J. (1788*b*), ‘Federalist No. 57, The Alleged Tendency of the New Plan to Elevate the Few at the Expense of the Many Considered in Connection with Representation’, accessed through [http://thomas.loc.gov/home/histdox/fed\\_57.html](http://thomas.loc.gov/home/histdox/fed_57.html).
- Marshall, M. G. & Jagers, K. (2010), ‘Polity IV Project: Dataset Users’ Manual’, accessed through <http://www.systemicpeace.org/inscr/p4manualv2010.pdf>.
- Maskin, E. & Tirole, J. (2004), ‘The Politician and the Judge: Accountability in Government’, *American Economic Review* **94**(4), 1034–1054.
- Mattozzi, A. & Merlo, A. (2008), ‘Political careers or career politicians?’, *Journal of Public Economics* **92**(3-4), 597–608.
- Messner, M. & Polborn, M. (2004), ‘Paying Politicians’, *Journal of Public Economics* **88**(12), 2423–2445.
- Pancer, S. M., Brown, S. D. & Barr, C. W. (1999), ‘Forming Impressions of Political Leaders: A Cross-National Comparison’, *Political Psychology* **20**(2), 345–368.
- Persson, T. & Tabellini, G. (2003), *The Economic Effects of Constitutions*, The MIT Press, Cambridge.
- Persson, T. & Tabellini, G. (2004), ‘Constitutional Rules and Fiscal Policy Outcomes’, *American Economic Review* **94**(1), 25–45.
- Sahuguet, N. & Persico, N. (2006), ‘Campaign Spending Regulation in a Model of Redistributive Politics’, *Economic Theory* **28**(1), 95–124.
- Sala-i-Martin, X. (1994), ‘Cross-Sectional Regressions and the Empirics of Economic Growth’, *European Economic Review* **38**(3-4), 739–747.
- Sala-i-Martin, X. (1997), ‘I Just Ran Two Million Regressions’, *American Economic Review* **87**(2), 178–83.

- Smart, M. & Sturm, D. M. (2006), Term Limits and Electoral Accountability, CEP Discussion Papers dp0770, Centre for Economic Performance, LSE.
- Stokes, D. E., Campbell, A. & Miller, W. E. (1958), ‘Components of electoral decision’, *The American Political Science Review* **52**(2), pp. 367–387.
- Tsebelis, G. (2002), *Veto Players. How Political Institutions Work*, Princeton University Press/Russell Sage Foundation.
- Voigt, S. (2011), ‘Positive Constitutional Economics II – A Survey of Recent Developments’, *Public Choice* **146**(1), 205–256.
- World Bank (2012), ‘World Development Indicators’, July 2012, accessed through <http://data.worldbank.org/indicator/NY.GDP.PCAP.KD>.
- WVS (2009), ‘World Values Survey 1981-2008 official aggregate v.20090902, 2009’, World Values Survey Association ([www.worldvaluessurvey.org](http://www.worldvaluessurvey.org)), Aggregate File Producer: ASEP/JDS Data Archive, Madrid, Spain.

## Appendix A: Proofs for main model

### Proof of Proposition 1

#### Efficient policy choice

Since voters can directly observe candidates’ abilities as well as their policies, voters have the belief system  $\sigma = a$  and are able to fully anticipate the difference in payoffs. The vote share of candidate 1 is given by

$$v_1(x_1, x_2, \sigma) = 1 - \Omega(x_2(\hat{a}_2(x_2) - c) - x_1(\hat{a}_1(x_1) - c)) = 1 - \Omega(x_2(a_2 - c) - x_1(a_1 - c)).$$

Candidate 1 chooses  $x_1$ , taking into account his opponent’s strategy  $X_2$ , to maximize

$$U_1(x_1, a_1) = \int_0^1 \phi(a_2) f(v_1(x_1, x_2, \sigma), \rho) (\theta + x_1(a_1 - c)) da_2.$$

As  $f$  is strictly increasing in  $v_1$ , which in turn is strictly increasing in the difference in welfare contributions, candidate 1 is only interested in maximizing his welfare contribution. Clearly, the dominant strategy is given by

$$X_1(a_1) = \begin{cases} x_i = 0 & \text{for } a_i < c \\ x_i = 1 & \text{for } a_i \geq c. \end{cases}$$

The reasoning for candidate 2 is analogous.

## Positive welfare effect of increasing power concentration

To simplify notation, we denote the welfare contribution of player  $i$  by  $\pi_i(x_i, \sigma) = x_i(\hat{a}_i(x_i) - c)$ . We suppress the dependence on the belief system and the action whenever possible without creating confusion. Moreover, let  $g(\pi_1 - \pi_2, \rho)$  be the expected power share above one half

$$g(\pi_1 - \pi_2, \rho) \equiv f(v_1(x_1, x_2, \sigma), \rho) - \frac{1}{2}.$$

For any candidate, there are two cases. He can either face an opponent with a reform proposal or one who proposes the status quo. The ex ante welfare is the weighted average of these two alternatives. Given the optimal behavior identified above, welfare in the full information case is given by

$$\begin{aligned} W(\rho) &= \\ & \int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[ \frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[ \frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1 \\ & + \int_0^1 \int_{a_1}^1 \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[ \frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[ \frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1 \\ & = 2 \int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) \left\{ \pi(X_1(a_1), \sigma) \left[ \frac{1}{2} + g(\pi_1 - \pi_2, \rho) \right] + \pi(X_2(a_2), \sigma) \left[ \frac{1}{2} - g(\pi_1 - \pi_2, \rho) \right] \right\} da_2 da_1. \end{aligned}$$

The derivative of the welfare function with respect to power concentration  $\rho$  is given by

$$\frac{dW}{d\rho} = 2 \int_0^1 \int_0^{a_1} \phi(a_1)\phi(a_2) [\pi(X_1(a_1), \sigma) - \pi(X_2(a_2), \sigma)] \frac{\partial g(\pi_1 - \pi_2, \rho)}{\partial \rho} da_2 da_1 > 0.$$

As  $a_1 > a_2$  under the integral, the payoff difference  $\pi_1 - \pi_2$  and  $g(\pi_1 - \pi_2, \rho)$  are throughout positive, so that we have  $\frac{\partial g(\pi_1 - \pi_2, \rho)}{\partial \rho} > 0$  due to the properties of the function  $f(v, \rho)$ .

## Proof of Lemma 2

We only need to deal with the case of a politician with ability lower than  $c$ , since candidates with ability above  $c$  always choose to reform. Candidate 1 chooses to reform if and only if

$$\begin{aligned} & prob(x_2 = 1)f(v_1(x_1 = 1, x_2 = 1, \sigma), \rho)(\theta + a_1 - c) \\ & + (1 - prob(x_2 = 1))f(v_1(x_1 = 1, x_2 = 2, \sigma), \rho)(\theta + a_1 - c) \\ & - prob(x_2 = 1)f(v_1(x_1 = 0, x_2 = 1, \sigma), \rho)\theta - (1 - prob(x_2 = 1))\theta \frac{1}{2} > 0. \end{aligned}$$

It can easily be seen that this reform incentive function is strictly monotonically increasing in the individual ability  $a_1$ . The same argument holds for Candidate 2. Thus, the optimal strategy of each candidate will always be a cutoff strategy.

## Proof of Proposition 2

Depending on the parameter values, there may exist equilibria in which the cutoff is equal to 1. As argued in the text, these cannot satisfy the D1 criterion. In the following, we show that there exists a unique D1 equilibrium exhibiting cutoffs different from 1 and that this equilibrium features symmetric cutoffs.

### Symmetry of cutoffs

Using the insight from above, we can write the incentive function of candidate 1 as

$$\begin{aligned} R_1(a, \alpha_1, \alpha_2, \rho) = & (1 - \Phi(\alpha_2))f(\rho, v_1(x_1 = 1, x_2 = 1, \sigma))(\theta + a - c) \\ & + \Phi(\alpha_2)f(\rho, v_1(x_1 = 1, x_2 = 0, \sigma))(\theta + a - c) \\ & - (1 - \Phi(\alpha_2))f(\rho, v_1(x_1 = 0, x_2 = 1, \sigma))\theta - \Phi(\alpha_2)\frac{\theta}{2}. \end{aligned}$$

Using the definition of  $g(\pi_1 - \pi_2, \rho)$ , this simplifies to

$$\begin{aligned} R(a, \alpha_1, \alpha_2) = & (1 - \Phi(\alpha_2)) \left[ \left( g(\pi_1 - \pi_2, \rho) + \frac{1}{2} \right) (\theta + a - c) \right] \\ & + \Phi(\alpha_2) \left[ \left( g(\pi_1, \rho) + \frac{1}{2} \right) (\theta + a - c) \right] - (1 - \Phi(\alpha_2)) \left( \frac{1}{2} - g(\pi_2, \rho) \right) \theta - \Phi(\alpha_2)\theta\frac{1}{2}. \end{aligned}$$

In equilibrium, the reform incentive is zero for the cutoff-type  $\alpha_1$ .

$$\begin{aligned} R_1(\alpha_1, \alpha_1, \alpha_2) = & 0 \\ \Leftrightarrow & \frac{\theta [\Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2)) [g(\pi_1 - \pi_2, \rho) + g(\pi_2, \rho)]]}{c - \alpha_1} = \\ & \frac{1}{2} + \Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2))g(\pi_1 - \pi_2). \end{aligned}$$

Subtracting the corresponding equation for  $R_2$ , we get

$$\begin{aligned} & \frac{\theta [\Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2)) [g(\pi_1 - \pi_2, \rho) + g(\pi_2, \rho)]]}{c - \alpha_1} \\ & - \frac{\theta [\Phi(\alpha_1)g(\pi_2, \rho) + (1 - \Phi(\alpha_1)) [-g(\pi_1 - \pi_2, \rho) + g(\pi_1, \rho)]]}{c - \alpha_2} = \\ & \Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2))g(\pi_1 - \pi_2, \rho) - \Phi(\alpha_1)g(\pi_2, \rho) + (1 - \Phi(\alpha_1))g(\pi_1 - \pi_2, \rho) \\ \Leftrightarrow & \left[ \frac{\theta\Phi(\alpha_2)}{c - \alpha_1} - \frac{\theta(1 - \Phi(\alpha_1))}{c - \alpha_2} - \Phi(\alpha_2) \right] g(\pi_1, \rho) \\ & - \left[ \frac{\theta\Phi(\alpha_1)}{c - \alpha_2} - \frac{\theta(1 - \Phi(\alpha_2))}{c - \alpha_1} - \Phi(\alpha_1) \right] g(\pi_2, \rho) \\ & + \underbrace{\left[ (1 - \Phi(\alpha_2)) \left( \frac{\theta}{c - \alpha_1} - 1 \right) + (1 - \Phi(\alpha_1)) \left( \frac{\theta}{c - \alpha_2} - 1 \right) \right]}_{>0} g(\pi_1 - \pi_2, \rho) = 0. \end{aligned}$$

If  $\alpha_1 = \alpha_2$ , this condition is trivially fulfilled. Assuming wlog  $\alpha_1 > \alpha_2$ , the equality above can only be satisfied if

$$\left[ \frac{\theta\Phi(\alpha_2)}{c - \alpha_1} - \frac{\theta(1 - \Phi(\alpha_1))}{c - \alpha_2} - \Phi(\alpha_2) \right] g(\pi_1, \rho) < \left[ \frac{\theta\Phi(\alpha_1)}{c - \alpha_2} - \frac{\theta(1 - \Phi(\alpha_2))}{c - \alpha_1} - \Phi(\alpha_1) \right] g(\pi_2, \rho).$$

However, we have  $\pi_1 > \pi_2$  by assumption, which implies  $g(\pi_1, \rho) > g(\pi_2, \rho)$ . Furthermore, we can show that the factor before  $g(\pi_1, \rho)$  is larger than the one before  $g(\pi_2, \rho)$ :

$$\begin{aligned} \frac{\theta}{c - \alpha_1}\Phi(\alpha_2) - \frac{\theta}{c - \alpha_2}(1 - \Phi(\alpha_1)) - \Phi(\alpha_2) &> \frac{\theta}{c - \alpha_2}\Phi(\alpha_1) - \frac{\theta}{c - \alpha_1}(1 - \Phi(\alpha_2)) - \Phi(\alpha_1) \\ \Leftrightarrow \frac{\theta}{c - \alpha_1} + \Phi(\alpha_1) &> \frac{\theta}{c - \alpha_2} + \Phi(\alpha_2). \end{aligned}$$

The last inequality is clearly fulfilled, generating a contradiction. Thus, the reform incentive functions  $R_1$  and  $R_2$  cannot simultaneously attain zero for different cutoffs, and there are only symmetric equilibria.

## Existence

Let  $\pi$  denote the difference in welfare contributions between a reform and a status quo proposal. Making use of the symmetric cutoffs, the incentive function simplifies to

$$R(\alpha, \rho) = \left[ \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right] (\alpha - c) + \theta g(\pi, \rho) = 0.$$

Note that  $R(1, \rho)$  is always positive. If  $R(0, \rho) < 0$ , the reform incentive is equal to zero at least once due to continuity, and there exists an interior equilibrium. If  $R(0, \rho) \geq 0$ , it is an equilibrium that candidates of all abilities choose to reform. Hence, there is at least one equilibrium.

## Proof of Proposition 3

Next, we establish uniqueness. The derivative of the incentive function with respect to  $\alpha$  is

$$\frac{\partial R}{\partial \alpha} = \underbrace{(\theta + (\alpha - c)\Phi(\alpha))g_\pi(\pi, \rho)}_A \frac{\partial \pi}{\partial \alpha} + \underbrace{\left( \frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho) \right)}_B.$$

The reform incentive function yields that  $A$  is always larger than zero in equilibrium for the cutoff type.  $B$  is also larger than zero, due to Assumption 1. The reform incentive is thus throughout increasing in the cutoff. Consequently, the reform incentive attains zero for at most one cutoff value.



We use implicit differentiation to prove that there is a unique  $\tilde{\theta}(\rho)$  such that the unique equilibrium is informative if and only if  $\theta < \tilde{\theta}(\rho)$ . If  $\theta = \tilde{\theta}(\rho) < \infty$ , the reform incentive is exactly zero for  $\alpha = 0$ . In an informative equilibrium, the derivative of the cutoff in  $\theta$  is given by

$$\frac{d\alpha}{d\theta} = -\frac{g(\pi, \rho)}{(\theta + (\alpha - c)\Phi(\alpha))g_\pi(\pi, \rho)\frac{\partial\pi}{\partial\alpha} + \left(\frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho)\right)} < 0.$$

The denominator is positive (see above), as is the numerator. Thus, this derivative is strictly negative in any informative equilibrium, and  $\alpha > 0$  for any  $\theta < \tilde{\theta}(\rho)$ . Moreover, the reform incentive function implies that  $\alpha \rightarrow c$  if  $\theta \rightarrow 0$ . By continuity, there is a unique  $\tilde{\theta}(\rho) > 0$  such that the unique equilibrium is informative if  $\theta < \tilde{\theta}(\rho)$ .

## Proof of Proposition 4

Again, we use implicit differentiation to evaluate the derivative.

$$\frac{d\alpha}{d\rho} = -\frac{\frac{\partial R}{\partial \rho}}{\frac{\partial R}{\partial \alpha}} = -\frac{(\theta + (\alpha - c)\Phi(\alpha))g_\rho(\pi, \rho)}{(\theta + (\alpha - c)\Phi(\alpha))g_\pi(\pi, \rho)\frac{\partial\pi}{\partial\alpha} + \left(\frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho)\right)} < 0.$$

While the numerator is unambiguously positive, the positive sign of the denominator follows from Assumption 2. Hence, the overall effect is negative.

## Proof of Lemma 3

Using the symmetry in equilibrium, welfare can be simplified considerably.

$$\frac{W(\rho)}{2} = \underbrace{\int_\alpha^1 \phi(a)(a - c)da}_{z(\alpha)} \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right).$$

Note that there is a direct effect on welfare, since the function  $g(\pi, \rho)$  depends on  $\rho$ , and an indirect effect, since  $\rho$  changes the strategies of the politicians. Hence, we evaluate the total derivative of  $W(\rho)$ :

$$\frac{dW}{d\rho} = \frac{\partial W}{\partial \rho} + \frac{\partial W}{\partial \alpha} \frac{d\alpha}{d\rho}.$$

In the following, we denote by  $D > 0$  the denominator of the derivative of  $\alpha$  with respect to  $\rho$ .

$$\begin{aligned} \frac{dW}{d\rho} = & \Phi(\alpha)z(\alpha)g_\rho(\pi, \rho) + \\ & + \left\{ (c - \alpha)\phi(\alpha) \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) + z(\alpha) \left( \phi(\alpha)g(\pi, \rho) + \Phi(\alpha)g_\pi(\pi, \rho)\frac{\partial\pi}{\partial\alpha} \right) \right\} \frac{d\alpha}{d\rho} \end{aligned}$$

$$\begin{aligned}
&= \left\{ \Phi(\alpha)z(\alpha) [\theta + (\alpha - c)\Phi(\alpha)] g_\pi \frac{d\pi}{d\alpha} + \Phi(\alpha)z(\alpha) \left[ \frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho) \right] \right. \\
&\quad \left. - \left[ (c - \alpha)\phi(\alpha) \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) + z(\alpha) \left( \phi(\alpha)g(\pi, \rho) + \Phi(\alpha)g_\pi(\pi, \rho) \frac{\partial \pi}{\partial \alpha} \right) \right] \right. \\
&\quad \left. [\theta + \Phi(\alpha)(\alpha - c)] \right\} \frac{g_\rho(\pi, \rho)}{D} \\
&= \frac{g_\rho(\pi, \rho)}{D} \left\{ \Phi(\alpha)z(\alpha) \left[ \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) + (\alpha - c)\phi(\alpha)g(\pi, \rho) \right] \right. \\
&\quad \left. - [\phi(\alpha)\theta g(\pi, \rho) + z(\alpha)\phi(\alpha)g(\pi, \rho)] \frac{c - \alpha}{2g(\pi, \rho)} \right\} \\
&= \frac{g_\rho(\pi, \rho)}{D} \left\{ \Phi(\alpha)z(\alpha) \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) - \phi(\alpha)(c - \alpha) \left[ \frac{\theta}{2} + z(\alpha) \left( \frac{1}{2} + \Phi(\alpha) \right) \right] \right\} \\
&= \frac{g_\rho(\pi, \rho)}{D} \left\{ \Phi(\alpha) \frac{W(\rho)}{2} - \phi(\alpha)(c - \alpha) \left[ \frac{\theta}{2} + \frac{W(\rho)}{2} \right] \right\} \\
&= \frac{g_\rho(\pi, \rho)}{2D} \{ \Phi(\alpha)W(\rho) - \phi(\alpha)(c - \alpha)(\theta + W(\rho)) \}
\end{aligned}$$

In any interior ( $\rho > \infty$ ) extreme value of the welfare function, the term in brackets has to equal zero, since its factor is positive. Rearranging, we get the following necessary and sufficient condition for interior extreme values of the welfare function:

$$h(\rho) \equiv \frac{\Phi(\alpha)}{\phi(\alpha)(c - \alpha)} - \left( 1 + \frac{\theta}{W(\rho)} \right) = 0.$$

Next, we prove that function  $h$  has at most one root in  $\rho$ , i.e., the welfare function attains at most one maximum. Assumption 2 is a sufficient condition for the first term to be decreasing in  $\rho$  and, thus, increasing in  $\alpha$ . In any interior extreme value of the welfare function, the second term is constant in  $\rho$ . Thus,  $h$  is decreasing in  $\rho$  at each interior root and so is the term in brackets. As  $h(\rho)$  is continuous in  $\rho$ , this implies that the welfare function has at most one interior maximum and no interior minimum, i.e., it is strictly quasi-concave.

## Proof of Proposition 5

In the next step, we show how the optimal level of  $\rho$  shifts with changes in  $\theta$ . The optimal level of  $\rho$  is interior whenever  $\lim_{\rho \rightarrow \infty} h(\rho) < 0$ , since  $\lim_{\rho \rightarrow 0} h(\rho) > 0$  for all  $\theta$ . For  $\theta \rightarrow 0$ , we get  $\alpha = c$  from the equilibrium condition. The limit of  $h(\rho)$  at  $\theta = 0$  is given by  $\lim_{\rho \rightarrow \infty} \Phi(c)W(\rho)$ . This is strictly positive for all  $\rho < \infty$ . Hence, the optimal institution embodies full concentration of power for  $\theta \rightarrow 0$ . Due to continuity, this is also true for an interval around 0. In contrast, for large  $\theta$ ,  $\alpha$  is close to zero if power concentration is high and  $\lim_{\rho \rightarrow \infty} h(\rho) < 0$ . Hence, the optimal level of power concentration is interior. Finally, we show that the optimal  $\rho$  decreases monotonically in  $\theta$  at any interior maximum.

Implicit differentiation gives

$$\frac{d\rho^*}{d\theta} = -\frac{\frac{dh(\rho)}{d\theta}}{\left.\frac{dh(\rho)}{d\rho}\right|_{\rho=\rho^*}}.$$

As argued before, the term in the denominator is negative. With respect to the numerator, note that the equilibrium cutoff  $\alpha$  is decreasing in  $\theta$ ,  $\frac{d\alpha}{d\theta} = -\frac{g(\pi,\rho)}{D} < 0$ . Consequently, the same is true for welfare,  $\frac{dW}{d\theta} = \frac{\partial W}{\partial \alpha} \frac{d\alpha}{d\theta} < 0$ . Hence,  $h$  is monotonically decreasing in  $\theta$ . In total, we conclude that  $\frac{d\rho^*}{d\theta} < 0$ . Overall, this implies that there is a cutoff  $\bar{\theta}$  such that if and only if  $\theta > \bar{\theta}$ , the optimal level of  $\rho$  is smaller than infinity and strictly falling in  $\theta$ .

## Appendix B: Proofs for modifications and extensions

### Proof of Proposition 6

The cutoff  $\tilde{\theta}(\rho)$  is defined by  $\tilde{\theta} = \frac{c}{2g(\pi,\rho)}$ . At this point, a candidate with ability equal to zero is indifferent between proposing a reform or the status quo if all other types propose a reform.  $\tilde{\theta}$  is decreasing in  $\rho$ . For  $\rho \rightarrow 0$ , we get that  $g(\pi,\rho) \rightarrow 0$  implying  $\tilde{\theta} \rightarrow \infty$ . Hence, for any given  $\theta$ , the monotonicity of  $\tilde{\theta}(\rho)$  implies the following: There is exactly one cutoff  $\tilde{\rho}(\theta)$  such that for all  $\rho < \tilde{\rho}(\theta)$  the unique D1 equilibrium must be informative.

### Proof of Proposition 8

For the case of limited commitment, the proofs of Proposition 1-5 need to be considered one by one. We shorten the proof whenever it is analogous or very similar to the case with full commitment. The proof of Proposition 1 does not rely on full commitment and thus carries over to the new setting.

### Proof of Lemma 2 with limited commitment

We only need to deal with the case of a politician with ability lower than  $c$ , since candidates with ability above  $c$  always choose to reform. Limited commitment changes the payoff from entering office from  $\theta + a - c$  to  $\theta + \lambda(a - c)$ . The rest of the proof is analogous to the case with full commitment.

### Proof of Proposition 2 with limited commitment

We just need to prove symmetry of cutoffs. The proof with regard to the classification of equilibria is identical to the case with full commitment. In equilibrium, the reform

incentive with limited commitment simplifies to

$$\begin{aligned}
R_1(\alpha_1, \alpha_1, \alpha_2) &= \\
&\theta [\Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi_2) [g(\pi_1 - \pi_2, \rho) + g(\pi_2, \rho)]] + \\
&\lambda(\alpha_1 - c) \left[ \frac{1}{2} + \Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2))g(\pi_1 - \pi_2) \right] = 0 \\
&\Leftrightarrow \frac{\theta [\Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2)) [g(\pi_1 - \pi_2, \rho) + g(\pi_2, \rho)]]}{\lambda(c - \alpha_1)} = \\
&\frac{1}{2} + \Phi(\alpha_2)g(\pi_1, \rho) + (1 - \Phi(\alpha_2))g(\pi_1 - \pi_2).
\end{aligned}$$

Subtracting the corresponding equation for the second player and proceeding as in the proof with full commitment, we obtain

$$\begin{aligned}
&\Leftrightarrow \left[ \frac{\theta\Phi(\alpha_2)}{\lambda(c - \alpha_1)} - \frac{\theta(1 - \Phi(\alpha_1))}{\lambda(c - \alpha_2)} - \Phi(\alpha_2) \right] g(\pi_1, \rho) + \\
&\underbrace{\left[ (1 - \Phi(\alpha_2)) \left( \frac{\theta}{\lambda(c - \alpha_1)} - 1 \right) + (1 - \Phi(\alpha_1)) \left( \frac{\theta}{\lambda(c - \alpha_2)} - 1 \right) \right]}_{>0} g(\pi_1 - \pi_2, \rho) = \\
&\left[ \frac{\theta\Phi(\alpha_1)}{\lambda(c - \alpha_2)} - \frac{\theta(1 - \Phi(\alpha_2))}{\lambda(c - \alpha_1)} - \Phi(\alpha_1) \right] g(\pi_2, \rho).
\end{aligned}$$

If  $\alpha_1 = \alpha_2$ , this condition is trivially fulfilled. Assuming wlog  $\alpha_1 > \alpha_2$ , the equality above implies that

$$\left[ \frac{\theta\Phi(\alpha_2)}{\lambda(c - \alpha_1)} - \frac{\theta(1 - \Phi(\alpha_1))}{\lambda(c - \alpha_2)} - \Phi(\alpha_2) \right] g(\pi_1, \rho) < \left[ \frac{\theta\Phi(\alpha_1)}{\lambda(c - \alpha_2)} - \frac{\theta(1 - \Phi(\alpha_2))}{\lambda(c - \alpha_1)} - \Phi(\alpha_1) \right] g(\pi_2, \rho).$$

However, we have  $\pi_1 > \pi_2$ . Moreover, we can show that

$$\begin{aligned}
\frac{\theta}{\lambda(c - \alpha_1)}\Phi(\alpha_2) - \frac{\theta}{\lambda(c - \alpha_2)}(1 - \Phi(\alpha_1)) - \Phi(\alpha_2) &> \frac{\theta}{\lambda(c - \alpha_2)}\Phi(\alpha_1) - \frac{\theta}{\lambda(c - \alpha_1)}(1 - \Phi(\alpha_2)) - \Phi(\alpha_1) \\
&\Leftrightarrow \frac{\theta}{\lambda(c - \alpha_1)} + \Phi(\alpha_1) > \frac{\theta}{\lambda(c - \alpha_2)} + \Phi(\alpha_2).
\end{aligned}$$

Thus, the reform incentive functions  $R_1$  and  $R_2$  can never simultaneously attain zero for  $\alpha_1 > \alpha_2$ . Thus, there can only be symmetric equilibria.

## Existence

The reform incentive function simplifies to

$$R(\alpha, \rho) = \left[ \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right] \lambda(\alpha - c) + \theta g(\pi, \rho) = 0.$$

Note that it is always positive if  $\alpha = 1$ . If  $R(0, \rho) < 0$ , the reform incentive is equal to zero at least once, due to the continuity and there exists an interior equilibrium. If

$R(0, \rho) \geq 0$ , it is an equilibrium that all candidates choose to reform. Hence, there is at least one equilibrium.

## Proof of Proposition 3 with limited commitment

Next, we establish uniqueness. The derivative with respect to  $\alpha$  is

$$\frac{\partial R}{\partial \alpha} = \underbrace{(\theta + (\alpha - c)\lambda\Phi(\alpha))g_{\pi}(\pi, \rho)}_A \frac{\partial \pi}{\partial \alpha} + \lambda \underbrace{\left(\frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho)\right)}_B.$$

The remainder of the proof is analogous to the case with perfect commitment.

## Proof of Proposition 4 with limited commitment

We use implicit differentiation to prove the proposition:

$$\frac{d\alpha}{d\rho} = -\frac{\frac{\partial R}{\partial \rho}}{\frac{\partial R}{\partial \alpha}} = -\frac{(\theta + \lambda(\alpha - c)\Phi(\alpha))g_{\rho}(\pi, \rho)}{(\theta + (\alpha - c)\lambda\Phi(\alpha))g_{\pi}(\pi, \rho)\frac{\partial \pi}{\partial \alpha} + \lambda\left(\frac{1}{2} + (\Phi(\alpha) + (\alpha - c)\phi(\alpha))g(\pi, \rho)\right)} < 0.$$

While the numerator is unambiguously positive, the positive sign of the denominator follows from Assumption 1.

## Proof of Proposition 5 with limited commitment

Inserting equilibrium strategies, the welfare function can be simplified to

$$\frac{W(\rho)}{2} = \underbrace{\left(\lambda \int_{\alpha}^1 \phi(a)(a - c)da + (1 - \lambda) \int_c^1 \phi(a)(a - c)da\right)}_{z(\alpha)} \left(\frac{1}{2} + \Phi(\alpha)g(\pi, \rho)\right).$$

The total derivative of the welfare function can be simplified along the same lines as with full commitment and yields the same necessary and sufficient condition for extreme values:

$$h(\rho) = \frac{\Phi(\alpha)}{\phi(\alpha)(c - \alpha)} - \left(1 + \frac{\theta}{W(\rho)}\right) = 0.$$

Thus, the rest of the proof is equivalent.

For the second step, we have to show how the unique maximum changes with  $\theta$ . For  $\theta \rightarrow 0$ , we get from the reform incentive  $\alpha = c$  and

$$\frac{dW(\rho)}{d\rho} \Big|_{\theta=0} = \frac{g_{\rho}(\pi, \rho)}{D} \Phi(\alpha) \lambda W(\rho).$$

This is positive. For  $\theta \rightarrow 0$ , the optimal institution hence fully concentrates power. Due to continuity, we get that this is also true for an interval around 0. Since  $h(\rho, \alpha)$  does

not change with limited commitment, we again refer the reader to the proofs for full commitment to see that the optimal  $\rho$  is monotonically decreasing in  $\theta$ .

## Proof of Proposition 9

In all proofs of Appendix A, we only use one important feature of the vote share  $v_i(x_1, x_2, \sigma)$ . This is, the vote share is weakly increasing in the expected abilities of the candidates and thus in the difference in welfare contributions between a reform and a status quo proposal. In the following, we show that this still holds for the case of heterogeneous policy preferences. All other proofs do not change. In the new setting, voter  $i$  votes for candidate 1 if

$$x_1(\mu_{1i}\hat{a}_1 - c) \geq x_2(\mu_{2i}\hat{a}_2 - c),$$

where  $\hat{a}_j$  denotes the expected ability of candidate  $j \in \{1, 2\}$ .

If both candidates propose a reform, the vote share of candidate 1 is

$$v_1(x_1 = 1, x_2 = 1, \sigma) = \int_k^l \int_k^{\frac{\mu_1 \hat{a}_1}{\hat{a}_2}} \xi(\mu_2) \xi(\mu_1).$$

The derivative with respect to  $\hat{a}_1$  is positive, since we assume the mean of the preferences to be larger than zero. If candidate 1 faces a status quo proposing opponent, his expected vote share is

$$v_1(x_1 = 1, x_2 = 0, \sigma) = \int_{\frac{c}{\hat{a}_1}}^l \xi(\mu_1).$$

The derivative with respect to  $\hat{a}_1$  is again positive. In the third case, where candidate 1 proposes the status quo, the vote share does not depend on the expected ability, since the payoff is independent of it. Hence, the expected overall vote share of candidate  $i$  is weakly increasing in his expected competence and, thus, his welfare contribution.

## Proof of Lemma 4

In an informative equilibrium, the expected utility of voter  $k$  with reform preference  $\mu_k$  is given by

$$V(\mu_k, \rho) = 2 \int_{\alpha}^1 \phi(a)(\mu_k a - c) da \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right).$$

It is strictly increasing in  $\mu_k$ , and negative for any  $\mu_k \leq c$ . Its derivative with respect to power concentration follows as

$$\frac{dV(\mu_k, \rho)}{d\rho} = 2\Phi(\alpha) \frac{dg}{d\rho} \int_{\alpha}^1 \phi(a)(\mu_k a - c) da$$

$$\begin{aligned}
& +2 \left[ \left( \phi(\alpha)g + \Phi(\alpha) \frac{dg}{d\rho} \right) \int_{\alpha}^1 \phi(a)(\mu_k a - c) da \right. \\
& \quad \left. - \phi(\alpha)(c - \mu\alpha) \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) \right] \frac{d\alpha}{d\rho} \\
& = 2 \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) (\Phi(\alpha) + \phi(\alpha - c)) \frac{dg}{d\rho} \int_{\alpha}^1 \phi(a)(\mu_k a - c) da \\
& \quad + 2 \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) \phi(\alpha)(c - \mu\alpha) \frac{d\alpha}{d\rho} \\
& = \frac{g\rho}{D} \left\{ [\Phi(\alpha) - \phi(\alpha)(c - \alpha)] \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) \int_{\alpha}^1 \phi(a)(\mu_k a - c) da \right. \\
& \quad \left. - \phi(\alpha)(c - \mu\alpha) \frac{\theta}{2} \right\}.
\end{aligned}$$

For any  $\rho \leq \rho^*(\theta)$ , the term  $\Phi(\alpha) - \phi(c - \alpha)$  is positive by Proposition 5 and Assumption 2. Thus, the expected utility of every voter with  $\mu_k < c$  is strictly decreasing in  $\rho$  on this interval. By a similar argument as used in Lemma 3, it can be shown that  $V(\mu_k, \rho)$  has at most one minimizer. For the limit case  $\rho \rightarrow \infty$ , however, we find that  $\frac{dV}{d\rho} \leq 0$ . In this limit, we have  $g(\pi, \rho) = 1$ , which implies  $\theta \geq [1 + \Phi(\alpha)](c - \alpha)$  and a negative sign of the bracket in the last line above. Thus,  $V(\mu_k, \rho)$  is monotonically decreasing in  $\rho$ .

## Proof of Proposition 10

The proof consists of two steps. First, we show that there exists at least one maximum for some  $\rho < \rho^*$ . Second, we ensure that there can never be a maximum for any  $\rho \geq \rho^*$ . Note that the expected utility  $V(\mu_k, \rho)$  is increasing in the reform preference  $\mu_k$ . Due to the strict concavity of  $w$ , this directly implies that  $w'(V(\mu_k, \rho)) > w'(V(\mu'_k, \rho))$  for any  $\mu_k < \mu'_k$ . Moreover, the cross derivative of expected utility with respect to  $\rho$  and  $\mu_k$  is

$$\frac{d^2 V(\mu_k, \rho)}{d\rho d\mu_k} = \frac{2g\rho(\pi, \rho)}{D} \left[ \left( \frac{1}{2} + \Phi(\alpha)g(\pi, \rho) \right) (\Phi(\alpha) + \phi(\alpha)(\alpha - c)) \int_{\alpha}^1 a\phi(a) da + \frac{\theta}{2} \alpha\phi(\alpha) \right]. \quad (5)$$

Since this term does not depend on  $\mu_k$ , the marginal effect of  $\rho$  on expected utility is monotonic in  $\mu_k$ . Take any welfare function of an inequality averse society:

$$W_{IA}(\rho) = \int_{\underline{\mu}}^{\bar{\mu}} w(V(\mu_k, \rho)) \xi(\mu_k) d\mu_k.$$

Its derivative with respect to  $\rho$  is

$$\frac{dW_{IA}(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} w'(V(\mu_k, \rho)) \frac{dV(\mu_k, \rho)}{d\rho} \xi(\mu_k) d\mu_k.$$

For the case of  $\rho' < \rho^*$ , the cross derivative  $\frac{d^2V(\mu_k, \rho)}{d\rho d\mu_k}$  is larger than zero. All terms of it are always positive except for  $(\Phi(\alpha) + \phi(\alpha - c))$ . This, however, is positive for all  $\rho' \leq \rho^*$  (see Proposition 5 and Assumption 2). The positive cross derivative yields

$$\frac{dW_{IA}(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} w'(V(\mu_k, \rho)) \frac{dV(\mu_k, \rho)}{d\rho} \xi(\mu_k) d\mu_k < \frac{dW(\rho)}{d\rho} = \int_{\underline{\mu}}^{\bar{\mu}} \frac{dV(\mu_k, \rho)}{d\rho} \xi(\mu_k) d\mu_k.$$

The derivatives of the expected utility are smaller for voters with smaller  $\mu_k$ . Exactly these utilities are weighted more strongly in the case of inequality aversion, since  $w'(V(\mu_k, \rho)) > w'(V(\mu'_k, \rho))$ . Hence, the derivative of the welfare function at  $\rho^*$  is negative, and there exists at least one local maximum for some level  $\rho' < \rho^*$ .

Now consider the case of  $\rho' > \rho^*$ , where  $\frac{dW}{d\rho} < 0$ . From Equation (5), we know that the cross derivative is throughout either positive or negative. Suppose the cross derivative is positive. Then, we have that  $\left. \frac{dW_{IA}}{d\rho} \right|_{\rho'} < \left. \frac{dW}{d\rho} \right|_{\rho'} < 0$ , and there cannot be a maximum at  $\rho'$ . Suppose that the cross derivative is negative at  $\rho'$ . From Lemma 4, we know that the marginal effect of  $\rho$  is throughout negative for voters with  $\mu < c$ . As a consequence of the negative cross derivative, the marginal effect is negative for all agents. Thus, the derivative of  $W_{IA}$  is certainly negative. Overall, there cannot be any maximum in the range  $[\rho^*, \infty)$ .

## Proof for Proposition ??

Let  $1 - \mu$  be the fraction of low office-motivation types with  $\theta_l = 0$  and  $\mu$  the fraction of high office motivation types with  $\theta_h = \theta > 0$ . And define by  $K(a)$  the weighted distribution function  $(1 - \mu)\Phi(c) + \mu\Phi(a)$  with derivative  $k(a) = \mu\phi(a)$ . First note that Assumption 1 implies that the same assumption also holds for  $K(a)$  instead of  $\Phi(a)$ . The proofs of propositions 2 to 4 then convey to the new setting with the new distribution function  $K(a)$ . The reform incentive function with these definitions is given by:

$$R(\alpha) = (\alpha - c) \left[ \frac{1}{2} + g(\pi, \rho)K(\alpha) \right] + \theta g(\pi, \rho)$$

## Proof of Proposition 12

With the property  $\frac{\partial R}{\partial \mu} > 0$ . Using the symmetry in equilibrium, welfare can be simplified considerably.

$$\frac{W(\rho)}{2} = \underbrace{(1 - \mu) \int_c^1 \phi(a)(a - c) da + \mu \int_\alpha^1 \phi(a)(a - c) da}_{z(\alpha)} \left[ (1 - \mu) \left( \frac{1}{2} + K(\alpha)g(\pi, \rho) \right) \right].$$

Note that there is a direct effect on welfare, since the function  $g(\pi, \rho)$  depends on  $\rho$ , and an indirect effect, since  $\rho$  changes the strategies of the politicians. Hence, we evaluate the



total derivative of  $W(\rho)$ :

$$\frac{dW}{d\rho} = \frac{\partial W}{\partial \rho} + \frac{\partial W}{\partial \alpha} \frac{d\alpha}{d\rho}.$$

In the following, we denote by  $D > 0$  the denominator of the derivative of  $\alpha$  with respect to  $\rho$ . Clearly, the welfare function has the same structure as before with  $K(\alpha)$  instead of  $\Phi(\alpha)$  and a slightly changed  $z(\alpha)$ . The following goes through without further changes.

$$\begin{aligned} \frac{dW}{d\rho} &= K(\alpha)z(\alpha)g_\rho(\pi, \rho) + \\ &\quad + \left\{ (c - \alpha)k(\alpha) \left( \frac{1}{2} + K(\alpha)g(\pi, \rho) \right) + z(\alpha) \left( k(\alpha)g(\pi, \rho) + K(\alpha)g_\pi(\pi, \rho) \frac{\partial \pi}{\partial \alpha} \right) \right\} \frac{d\alpha}{d\rho} \\ &= \left\{ K(\alpha)z(\alpha) [\theta + (\alpha - c)K(\alpha)] g_\pi \frac{d\pi}{d\alpha} + K(\alpha)z(\alpha) \left[ \frac{1}{2} + (K(\alpha) + (\alpha - c)k(\alpha))g(\pi, \rho) \right] \right. \\ &\quad \left. - \left[ (c - \alpha)k(\alpha) \left( \frac{1}{2} + K(\alpha)g(\pi, \rho) \right) + z(\alpha) \left( k(\alpha)g(\pi, \rho) + K(\alpha)g_\pi(\pi, \rho) \frac{\partial \pi}{\partial \alpha} \right) \right] \right. \\ &\quad \left. [\theta + K(\alpha)(\alpha - c)] \right\} \frac{g_\rho(\pi, \rho)}{D} \\ &= \frac{g_\rho(\pi, \rho)}{D} \left\{ K(\alpha)z(\alpha) \left[ \frac{1}{2} + K(\alpha)g(\pi, \rho) + (\alpha - c)k(\alpha)g(\pi, \rho) \right] \right. \\ &\quad \left. - [k(\alpha)\theta g(\pi, \rho) + z(\alpha)k(\alpha)g(\pi, \rho)] \frac{c - \alpha}{2g(\pi, \rho)} \right\} \\ &= \frac{g_\rho(\pi, \rho)}{D} \left\{ K(\alpha)z(\alpha) \left( \frac{1}{2} + K(\alpha)g(\pi, \rho) \right) - k(\alpha)(c - \alpha) \left[ \frac{\theta}{2} + z(\alpha) \left( \frac{1}{2} + K(\alpha) \right) \right] \right\} \\ &= \frac{g_\rho(\pi, \rho)}{D} \left\{ K(\alpha) \frac{W(\rho)}{2} - k(\alpha)(c - \alpha) \left[ \frac{\theta}{2} + \frac{W(\rho)}{2} \right] \right\} \\ &= \frac{g_\rho(\pi, \rho)}{2D} \{ K(\alpha)W(\rho) - k(\alpha)(c - \alpha) (\theta + W(\rho)) \} \end{aligned}$$

In any interior extreme value of the welfare function, the term in brackets has to equal zero, since its factor is always positive. Rearranging, we get the following necessary and sufficient condition for interior extreme values of the welfare function:

$$h(\rho) \equiv \frac{K(\alpha)}{k(\alpha)(c - \alpha)} - \left( 1 + \frac{\theta}{W(\rho)} \right) = 0.$$

Next, we prove that function  $h$  has at most one root in  $\rho$ , i.e., the welfare function attains at most one interior maximum. Assumption 2 is a sufficient condition for the first term to be decreasing in  $\rho$  and, thus, increasing in  $\alpha$ . In any interior extreme value of the welfare function, the second term is constant in  $\rho$ . Thus,  $h$  is decreasing in  $\rho$  at each root and so is the term in brackets. As  $h(\rho)$  is continuous in  $\rho$ , this implies that the welfare function has at most one interior maximum and no interior minimum, i.e., it is strictly quasi-concave.

In the next step, we show how the optimal level of  $\rho$  shifts with changes in  $\mu$ . For  $\mu = 0$ , we get the first best solution  $\alpha = c$  and full power concentration is the optimal institution. Due to continuity, this is also true for an interval around 0. In contrast, we

know from the proposition in the main part that for  $\theta > \bar{\theta}$  and  $\mu = 1$ , the optimal level of power concentration is interior. Moreover, the optimal  $\rho$  decreases monotonically in  $\mu$  at any interior maximum. Implicit differentiation gives

$$\frac{d\rho^*}{d\mu} = -\frac{\frac{dh(\rho)}{d\mu}}{\left.\frac{dh(\rho)}{d\rho}\right|_{\rho=\rho^*}}.$$

As argued before, the term in the denominator is negative. With respect to the numerator, note that the equilibrium cutoff  $\alpha$  is decreasing in  $\mu$ ,

$$\frac{d\alpha}{d\mu} = -\frac{(\alpha - c)g(\pi, \rho)(\Phi(\alpha) - \Phi(c))}{D} < 0$$

Consequently, the same is true for welfare,  $\frac{dW}{d\mu} = \frac{\partial W}{\partial \mu} + \frac{\partial W}{\partial \alpha} \frac{d\alpha}{d\mu} < 0$ . Hence,  $h$  is monotonically decreasing in  $\mu$ . In total, we conclude that  $\frac{d\rho^*}{d\mu} > 0$  in any interior equilibrium. This implies that there exists a  $\bar{\mu}$  such that for all  $\mu > \bar{\mu}$  the optimal institution exhibits some dispersion of power and is decreasing in  $\mu$ .

# Appendix C: Data

## Description and sources of variables

### Main variables

Growth in real GDP per capita	Average growth rate. Computed based on per capita GDP in constant 2000 US\$. World Bank (2012).
GDP per capita	Denominated in constant 2000 TUS\$. World Bank (2012).
Office motivation	International Social Survey Programme 2004: Citizenship I. ISSP Research Group (2012).
Power dispersion	Lijphart's index for executive-parties dimension. Lijphart (1999).

### Variables for robustness checks

Trust in political parties	World Values Survey, third wave. WVS (2009).
Power dispersion	Time-variant proxy for Lijphart's executive-parties dimension, year 2004. Armingeon et al. (2011).
Power dispersion, institutional	Time-variant proxy for Lijphart's executive-parties dimension, institutional factors, year 2004. Armingeon et al. (2011).
Checks and balances	Number of veto players. Keefer & Stasavage (2003).
Plurality electoral system	Dummy variable. Beck et al. (2001).
Electoral system	Type of electoral system, 9 minor categories. IDEA (2004).

### Country list

Australia	Austria	Canada	Denmark
Finland	France	Germany	Ireland
Israel	Japan	Netherlands	Norway
Portugal	Spain	Sweden	Switzerland
United Kingdom	United States		

## Summary of variables

	Mean	Std. dev.	Min	Max	Poss. values
Power dispersion	0.31	0.98	-1.21	1.77	[-2,2]
Office motivation	3.37	0.37	2.61	4.20	[1,5]
GDP p.c.	26.98	7.69	11.55	39.83	
GDP p.c. growth (2004-2011)	0.68	0.74	-0.61	2.40	
GDP p.c. growth (1991-2004)	2.08	1.07	0.56	5.59	

For the regression analysis, the variables power dispersion and office motivation are rescaled to range between 0 and 1.

## Correlation table

	Power dispersion	Office motivation	GDP p.c.	GDP p.c. growth (2004-2011)
Office motivation	-0.20 (0.43)	1		
GDP p.c.	0.20 (0.43)	-0.58 (0.01)	1	
GDP p.c. growth (2004-2011)	0.44 (0.07)	0.072 (0.78)	-0.15 (0.56)	1
GDP p.c. growth (1991-2004)	-0.27 (0.28)	-0.10 (0.69)	-0.021 (0.93)	-0.48 (0.05)

Pearson's correlation coefficient,  $p$ -values in parentheses.