

Electoral Systems and Inequalities in Government Interventions

Garance Genicot*  Laurent Bouton*  Micael Castanheira[§]

March 6, 2019


ABSTRACT

This paper studies the political determinants of inequalities in government interventions under the majoritarian and proportional representation systems. Using a model of electoral competition with targetable government interventions and heterogeneous localities, we uncover a novel *relative electoral sensitivity effect* in majoritarian systems. This effect, which depends on the geographic distribution of voters, can incentivize parties to allocate resources more equally under majoritarian than proportional representation systems. This contrasts with the conventional wisdom that government interventions are more unequal in majoritarian systems. Our results also inform the debate about reforming the US Electoral College.

JEL Classification Numbers: D72, H00

Keywords: Distributive Politics, Electoral Systems, Electoral College, Public Good, Inequality.

Acknowledgements: We benefited from comments by Timm Betz, Patrick François, Leyla Karakas, Alessandro Lizzeri, Dilip Mookherjee, Massimo Morelli, Nicola Persico, Amy Pond, and Debraj Ray, as well as seminar and conference participants at Princeton, Georgetown, MSU, Texas A&M, UBC, Bocconi, Collegio Carlo Alberto, ECARES, EUI, TSE, ETH Zurich, Copenhagen, CERGE-EI, Essex, King's College, Cornell PE Conference 2018, ThReD Conference 2018, Quebec PE Conference 2018, POLECON UK 2nd Annual Workshop, CORE Belgian-Japanese Public Finance Workshop, and the 13th Conference on Economic Growth and Development. We are grateful to Dario Sansone for excellent research assistance.

*Georgetown University, CEPR and NBER; [§]Université Libre de Bruxelles and CEPR.  Names are in random order following Ray  Robson (2018). Shortened citation should be Genicot  *al.* (2018). Correspondence: bouton11j@gmail.com, micael.casta@gmail.com, and garance.genicot@georgetown.edu.

1 Introduction

Government interventions are fraught with inequalities. Substantial geographic inequalities have been documented both in terms of quantity and quality of public goods and services (Alesina et al. (1999), Banerjee et al. (2008), World Bank (2004)) and taxation (Albouy (2009), Troiano (2017)). A large empirical literature on distributive politics highlights the importance of political factors. These factors include the apportionment of constituencies (Ansolabehere et al. (2002)); their electoral contestability (Strömberg (2008)); and voters' characteristics such as turnout (Martin (2003), Strömberg (2004)), information (Besley and Burgess (2002), Strömberg (2004)), the presence of core supporters (or co-ethnics) of the candidate/party (Schady (2000), Hodler and Raschky (2014)), and their responsiveness to electoral promises (Johansson (2003), Strömberg (2008)). Overall, the political distortions of government interventions appear sizable: For instance, using Brazilian data, Finan and Mazzocco (2016) estimate that 25 percent of public funds allocated by legislators are distorted relative to the socially optimal allocation.

This paper studies the effect of electoral systems on inequalities in government interventions. We focus on two classes of systems: the *Majoritarian* (MAJ) and the *Proportional Representation* (PR) systems.¹ These two systems are ubiquitous² and debates about whether to use one or the other are common, both at the time of democratic transitions and in more established democracies. For the latter, these debates lead relatively frequently to actual electoral reforms.³ Moreover, our analysis is also directly relevant to the debates surrounding the reforms of the U.S. electoral college⁴ and the system to select the

¹In MAJ systems, there are a multitude of electoral districts that each select a limited number of representatives using some winner-take-all method. The archetype of those systems is the version with single-member districts and first-past-the-post voting. In PR systems, there are fewer electoral districts that each select at least two representatives, more or less in proportion to the vote shares of each party. The epitome of PR systems is the version with a single nationwide electoral district.

²For instance, 82 percent of legislative elections held in the 2000s employed either MAJ or PR (Bormann and Golder (2013)).

³Colomer (2004) (p. 55) counts “82 major electoral system changes for assemblies [...] in 41 countries.” between the early nineteenth century and 2002. Of those reforms, 40 are cases in which MAJ has been replaced by PR (or a mixed systems), and 13 are case of reforms from PR to MAJ (or mixed). The case of Italy is particularly striking: the electoral system was reformed three times between 1993 and 2015.

⁴As discussed in Whitaker and Neale (2004), this is a recurrent debate: “Since the adoption of the Constitution, [...] in almost every session of Congress, resolutions have been introduced proposing electoral college reform. Indeed, more proposed constitutional amendments have been introduced in Congress regarding electoral college reform than on any other subject.” (p. CRS-17) A current initiative among

President of the European Commission.⁵

The conventional wisdom is that MAJ systems are more conducive to inequality, because they provide steeper incentives for targeting government interventions onto specific groups (Persson and Tabellini (1999, 2000); Persson (2002); Lizzeri and Persico (2001); Milesi-Ferretti et al. (2002); Grossman and Helpman (2005)). This view is based on multiple theoretical arguments, which we detail in Section 2. One of the most powerful is that of Lizzeri and Persico (2001): In MAJ systems, parties only need *fifty* percent of the votes in *fifty* percent of the electoral districts to win a majority of seats in the national assembly. By contrast, they need fifty percent of *all* votes in PR systems, doubling the number of votes required to hold a majority of seats. Similar arguments lie at the core of a literature in trade, arguing that countries with MAJ systems are more likely to impose targeted trade barriers to favor specific regions. Instead, countries with PR systems are more likely to have free trade (Rogowski (1987) and Grossman and Helpman (2005)).

This view overlooks another important difference: the geographic distribution of voters matters differently in the two systems. In MAJ systems, parties must win in different electoral districts in order to win multiple seats. Paraphrasing Lizzeri and Persico (2001), they need to win fifty percent of the votes in *at least* fifty percent of the districts. This geographical constraint is largely absent in PR systems: Additional votes from any location help a party win more seats in the national assembly.

To take account of the geographical distribution of voters, we develop a model of electoral competition in which two parties compete by targeting governmental resources (cash transfers, goods, or services) to group of voters, called *localities*. The first key feature is that localities are heterogeneous along several dimensions: they may differ in population size, turnout rate, and swingness.⁶ Together these characteristics define what we call the

states to award all their electoral votes to the candidate that wins the popular vote, the National Popular Vote Interstate Compact (www.nationalpopularvote.com), is gaining momentum.

⁵Many are advocating for a reform of the current indirect selection process to replace it by a direct Europe-wide election (see, e.g., goo.gl/9Akj4T and goo.gl/UQa9hq). If European Union members were to decide in favor of a such a direct election, the immediate key question would be: what electoral system should be used?

⁶In practice, there can be substantial variations across localities in each of these dimensions. For instance, Stashko (2018) shows that, in 2012, U.S. counties had a mean population size of 99,970, with a standard deviation of 319,922. She also identifies substantial variation in turnout: in 2010, the mean turnout was 0.38 with a standard deviation of 0.15.

electoral sensitivity of a locality: a measure of the electoral responsiveness of that group of voters to government intervention.

The second key feature is that interventions are targetable at a level potentially finer than the electoral district. Such targetability is relevant for a wide array of government structures and interventions. It can be the central/federal government providing goods, services, or transfers at the (sub-)regional/(sub-)state level. It can also be a lower level of government (region or state) allocating resources at the district or sub-district level. Taking the case of the US, a substantial fraction of federal spending appears to be discretionary and targetable at the congressional district and/or the county level (see Berry et al. (2010)). And state governments target transfers to counties through a variety of programs, including transportation infrastructure, social transfers, and education (see Ansolabehere et al. (2002) and Stashko (2018)).

With these two features, our model uncovers a *relative electoral sensitivity effect* present only in MAJ systems. In those systems, because parties must accumulate victories in different districts to increase their seat share, they have incentives to allocate more resources to localities that are more electorally sensitive as compared to the other localities *in the same district*. These localities have a high relative electoral sensitivity. By contrast, in PR systems, the geographic distribution of the votes collected by each party is irrelevant. This creates an incentive to allocate more resources to localities that are more electorally sensitive, independently of their location. Stashko (2018) proposes an empirical test of our relative electoral sensitivity effect. Using US state governments and legislative elections, she finds that the effect is statistically and economically significant. That is, the amount received by a county depends not only on the electoral sensitivity of that county, but also on the electoral sensitivity of the other counties in the same district.

We then explore the consequence of the relative electoral sensitivity effect for the comparative level of inequality in government interventions under MAJ and PR systems. We find that this effect may imply a higher concentration of resources in PR. Consider for instance an hypothetical country in which low turnout localities are grouped in one district and high turnout localities are located in another district. Therefore, within-district heterogeneity is relatively small, whereas between-district heterogeneity is relatively large.

Under MAJ systems, competition for resources only happens among localities within a district. Turnout would then have no effect on the allocation of resources. A switch to PR would imply that low turnout localities are exposed to more direct competition from the high turnout localities. This would induce the parties to concentrate resources on the latter, which would generate inequalities.

More generally, we are able to identify conditions under which the relative sensitivity effect induces parties to “sprinkle” resources across districts, resulting in lower inequalities in government intervention in MAJ systems than in PR ones. These conditions depend on the heterogeneity of within- and between-district levels of electoral sensitivities and on district contestabilities (*i.e.* how likely the district is to be tied for victory).

Our findings shed a new light on the mixed empirical evidence in the literature that compares government interventions under those systems (see Section 2).

2 Related Literature

Our paper contributes to the literature on distributive politics, which studies the allocation of governmental resources to various subsets of the population. Within that literature, our work more closely relates to studies of the effects of electoral systems. As already mentioned, a recurrent theme in this literature is that parties want to target a smaller fraction of the population in MAJ systems than in PR ones. There are various mechanisms that produce this outcome. We already mentioned the fifty-of-fifty percent mechanism (Lizzeri and Persico (2001)), which becomes fifty-of-*at-least*-fifty percent in the presence of heterogeneity at the locality level.

Another mechanism highlights the importance of district contestability (the likelihood that electoral promises change which party wins a district) in MAJ systems. Persson and Tabellini (1999, 2000) and Persson (2002) show that parties target the most contestable districts. As long as there are voters to be swung all over the country, such incentives do not exist in PR systems. Building up on that mechanism, Strömberg (2008) highlights the importance of the pivotability of a district in the national assembly – *i.e.*, the likelihood that the identity of the party holding a majority of seats will change. Parties have incen-

tives to target districts that are both contestable and pivotal. Our model captures those different incentives.

Grossman and Helpman (2005) highlights the importance of bargaining between party leaders (who care about national welfare) and legislators (who care about the welfare of their constituents).⁷ In PR systems, legislators have a national constituency. Their incentives are therefore aligned with those of party leaders, which means they do not geographically target policies. By contrast, in MAJ systems, legislators' constituencies are geographically determined, hence the tension with party leaders. As soon as legislators have bargaining power, this leads to more geographically targeted policies than under PR systems. Our approach is complementary. While our model does not capture the tension between party leaders and legislators, we show that the party leaders' preferred allocation of resources varies with the electoral system. Due to the relative sensitivity effect, there are situations in which party leaders have stronger incentives to target policies under PR than under MAJ.

Rogowski and Kayser (2002) points to the seats-votes elasticity as a key factor influencing the targeting of government interventions. When elasticity is higher, parties have stronger incentives to target groups that can deliver many votes at the margin. Given that MAJ systems have a higher seats-votes elasticity than PR systems (Taagepera and Shugart (1989)), there should be more targeting toward electorally responsive groups under MAJ systems. Our results somehow refine this prediction. We microfound the electoral sensitivity of localities and show how it differentially affects government interventions under MAJ and PR systems.

There is a large empirical literature comparing MAJ and PR systems. This literature can be divided into two strands.⁸ The first relies on government accounts and has to make assumptions about which items can reasonably be thought to represent broad public goods

⁷Party discipline gives more power to the former, and tends to be higher in Parliamentary than in Presidential systems (see *e.g.* Tsebelis (1995); Carey (2007); Dewan and Spirling (2011)) – a dimension not explicitly considered in our model. Even in presidential systems, however, the importance of the party label means that the platform of the party as a whole may end up being more important than that of individual politicians (see *e.g.* Snyder and Ting (2002); Krasa and Polborn (2018)).

⁸We are aware of two studies that do not fit that nomenclature, because they focus on the behavior of individual politicians instead of the behavior of the parties controlling the government budget. These are (*i*) Gagliarducci et al. (2011), which uses Italian data, and (*ii*) Stratmann and Baur (2002), which uses German data.

as opposed to targeted transfers. Using cross-country regressions, most of these studies find that PR systems are associated with lower levels of spending classified as targeted and higher levels of spending classified as universal (Persson and Tabellini (1999, 2000); Milesi-Ferretti et al. (2002); Blume et al. (2009); Funk and Gathmann (2013)). One exception is Aidt et al. (2006), which studies the *changes* from MAJ to PR rules that took place in 10 European countries between 1830 and 1938. It finds that these led to a *decrease* in spending classified as universal.

The second strand focuses on trade policies. These studies compare trade barriers in MAJ and PR systems. The interpretation is that trade barriers are targeted transfers. The empirical evidence is mixed: Using cross-country regressions, a number of studies find that MAJ countries are more protectionist (Evans (2009); Hatfield and Hauk (2014); Rickard (2012)), while others find more protectionism in PR countries (Mansfield and Busch (1995); Rogowski and Kayser (2002); Chang et al. (2008); Betz (2017)). The difference seems to originate in the type of trade barriers considered: Non-tariff barriers tend to be used more often in PR systems, while tariffs tend to be used more heavily in MAJ systems.

There are a number of methodological challenges that remain unaddressed in these studies. In particular, Keefer (2004) and Golden and Min (2013) have criticized the arbitrary nature of the classification of expenditures as broad public goods or targeted transfers. Moreover, these classifications happen to vary across studies, and a given classification of government expenditures by type is unlikely to fit all countries.⁹ Another issue is that such classifications rest on the assumption that there exists such a thing as a “universal public good.” Instead (with some exceptions, such as nuclear deterrence), one is bound to admit that “public goods” are targetable, geographically or otherwise. The key question, then, is to identify when governments exploit their margin of action to target them in practice.¹⁰

To understand how these issues affect empirical findings, we have revisited Persson and

⁹For instance, road expenditures are typically seen as targeted expenditures (“pork-barrel”) in the US, but could be envisioned as broad public goods in small and/or developing countries.

¹⁰Large-sample cross-country or panel analyses (see, e.g., Persson and Tabellini (2003), Iversen and Soskice (2006), or Blume et al. (2009)) have typically avoided this problem. Only a few recent analyses have looked at a much more granular level to measure how public goods are supplied locally – e.g., between municipalities of a similar district (see, e.g., Azzimonti (2015); Blakeslee (2015); Funk and Gathmann (2013); Gagliarducci et al. (2011); Min (2015); Strömberg (2008), and Golden and Min (2013) for a survey).

Tabellini (1999) and Blume et al. (2009) using a new measure of how *encompassing*, as opposed to targeted, governmental spending is. This measure is based on the assessment of local experts. While obviously imperfect, it has the advantage of (i) not relying on the choices of the econometrician and (ii) potentially taking into account the specificities of the different countries in the sample. We find no significant differences in how targeted government expenditures are between MAJ and PR countries. The details of the analysis can be found in Appendix 2.

As mentioned in the Introduction, our relative electoral sensitivity effect sheds a new light on this mixed empirical evidence. The theoretical literature provides at least one other reason why PR systems may lead to more targeting of government interventions: There are usually more parties in PR systems. And, as shown by Cox (1990) and Myerson (1993), an increase in the number of parties should be associated with the targeting of government interventions toward a narrower subset of the electorate. The incentive to provide global public goods should thus also decrease (see Lizzeri and Persico (2005)). Using Indian data, Chhibber and Nooruddin (2004) finds that the provision of public good decreases when the number of parties increases, and conversely for the provision of club goods. Similar results emerge in multi-country panel analyses such as Park and Jensen (2007), which focuses on agricultural subsidies in OECD countries, or Castanheira et al. (2012), which focuses on tax reforms in EU countries.

3 The Model

3.1 The Economy

Consider a country with a continuum of individuals of total mass 1. The population is partitioned into *localities* $l \in \{1, 2, \dots, L\}$ of size n_l , *s.t.* $\sum_l n_l = 1$. Each locality belongs to an *electoral district* $d \in \{1, 2, \dots, D\}$.

An elected government has to allocate a total budget y among the different localities. We denote by q_l the amount of government intervention *per capita* in locality l . The intervention of the government is then summarized by $\mathbf{q} = \{q_1, \dots, q_L\}$.

If $L > D$, our model implies that governmental resources can be targeted at a finer level than the electoral district, while if $L = D$, then localities corresponds to district.¹¹ In Section 6.1, we extend the model to a setting where the politician allocates a budget to both a global public good and to transfers targetable at the locality level.

We cover a variety of government interventions that range from pure local public goods to pure transfers. The central difference between the types of government interventions is the extent of the economies of scale with respect to population size. With pure public goods, costs are independent of the number of individuals who benefit from the intervention. In contrast, with pure transfers, costs are proportional to the number of individuals who benefit. To also capture intermediate situations, we assume that the cost of providing q_l to the n_l individuals in locality l is: $n_l^\alpha q_l$, with $\alpha \in [0, 1]$. The government's aggregate budget constraint is thus:

$$\sum_l n_l^\alpha q_l \leq y, \quad (1)$$

When $\alpha = 1$, the government intervention is a pure transfer, and the budget constraint becomes: $\sum_l n_l q_l \leq y$. When $\alpha = 0$, q_l is a pure local public good, and the budget constraint becomes $\sum_l q_l \leq y$.

Individuals of locality l have preferences $u_l(\mathbf{q})$ for the government intervention, with $\partial^2 u_l(\mathbf{q})/\partial q_l^2 < 0 < \partial u_l(\mathbf{q})/\partial q_l$ – the function is strictly increasing and strictly concave in q_l . Moreover, we assume that $u_l(\mathbf{q}) = u(q_l)$, meaning that government interventions do not produce spillover across localities.¹²

3.2 Normative Benchmark

Before introducing the structure of electoral competition, we establish the politics-free benchmark. In this context and as argued in Becker (1958), the outcomes of electoral

¹¹The model can be readily extended to situations where localities spans multiple districts. In order to test our results on US data, Stashko (2018) considers the case of US counties split across districts.

¹²With externalities, the value of a public good in a locality l would be the sum of the local returns weighted by the intensity of the externality. Anticipating on equation (3) for instance, it would become: $\frac{\partial u_l(\mathbf{q})}{\partial q_l} = \lambda^{\text{SW}} \sum_j e_{l,j} n_j^{\alpha-1} \forall l$, with $e_{l,j}$ denoting the externality from l to j , with $e_{l,l} := 1$. This would require defining an entire network matrix between each locality and substantially longer sensitivity coefficients, without much new insight for our purpose here.

competition are always Pareto efficient. However, all allocations are not equivalent when using a Benthamite welfare function. A utilitarian social planner would:

$$\max_{\mathbf{q}} \mathcal{W}(\mathbf{q}) = \sum_l n_l u_l(\mathbf{q}), \quad s.t. \quad \sum_l n_l^\alpha q_l = y. \quad (2)$$

The socially optimal allocation according to this criteria must satisfy the standard Samuelsonian conditions:

$$\frac{\partial u_l(\mathbf{q})}{\partial q_l} = \lambda^{\text{SW}} n_l^{\alpha-1} \quad \forall l, \quad (3)$$

where λ^{SW} is the Lagrange multiplier associated with the budget constraint.

It is important to note that, except for the limit case of pure transfers (that is as long as $\alpha < 1$) the socially optimal allocation responds positively to population size in the locality. This implies that the social optimum tolerates “vertical inequalities”: Individuals in localities of different population sizes should benefit from different levels of government intervention. Electoral competition may, however, generate incentives that lead to different patterns of government interventions.

3.3 Electoral Competition

We consider an election with two parties, A and B , that compete for seats in the national assembly. We assume for now that their objective is to *maximize their expected number of seats*. There is debate in political economy about the objective of parties, in particular on whether they maximize the expected number of seats or the probability of winning a majority of seats. Section 6.2 extends our results to the case in which parties maximize their probability of winning a majority of seats.

We contrast two different electoral systems: the *proportional representation system* (PR henceforth), where seats are attributed in proportion to the *fraction of national votes* garnered by each party, and the *majoritarian system* (MAJ henceforth), where seats are proportional to the *fraction of districts* won by each party. In line with the literature, we consider a “pure” majoritarian system, in which each electoral district sends a single seat to the national assembly and the party with the most votes in a district wins its seat. In Sections 6.3 and 6.4, we show how our results extend to other commonly used versions of

the majoritarian and proportional representation systems.

To maximize their expected seat share, both parties simultaneously make a binding budget proposal, \mathbf{q}^A and \mathbf{q}^B , that details the allocation of resources across localities.¹³ These proposals must satisfy the government budget constraint (1).

Beyond their population size, localities are heterogeneous in various dimensions. They may differ in turnout rates, and/or the distribution of voter preferences, and they may belong to different electoral districts. We could easily include other dimensions of heterogeneity, such as information about electoral promises, and partisanship (see Section 6.5).

Consider individual i in locality l . Because of eligibility constraints (*e.g.*, age) or other reasons (*e.g.*, excessive cost of voting) she may, or may not, cast a valid vote on election day. We denote by t_l – for turnout – the exogenous probability with which a randomly sampled individual in locality l actually casts a valid vote. Hence, out of a local population size n_l , the number of *active voters* is $t_l n_l$.¹⁴

Conditional on casting a valid ballot, and given the parties' proposals, we assume that individual i votes for party A if and only if:

$$\Delta u_l(\mathbf{q}) \geq \nu_{i,l} + \delta_d, \quad (4)$$

where $\Delta u_l(\mathbf{q}) := u_l(\mathbf{q}_A) - u_l(\mathbf{q}_B)$ is the policy component of the preferences, and the shocks $\nu_{i,l}$ and δ_d capture all the political dimensions that do not belong to the budget constraint (party-related scandals, foreign policy shocks, etc.) and/or political preferences that are *ex ante* unknown to the parties, in the probabilistic voting tradition. For simplicity and in line with Persson and Tabellini (1999, 2000), we assume these shocks are uniformly distributed:¹⁵

$$\nu_{i,l} \sim U \left[\frac{-1}{2\phi_l}, \frac{1}{2\phi_l} \right] \quad \text{and} \quad \delta_d \sim U \left[\beta_d - \frac{1}{2\gamma_d}, \beta_d + \frac{1}{2\gamma_d} \right].$$

¹³As explained in Golden and Min (2013, p.84) many studies find that politicians are rewarded by voters for distributive allocations.

¹⁴We could also endogenize turnout, as in Lindbeck and Weibull (1987). As in their model, the equilibrium allocations would not be substantially affected.

¹⁵We conjecture that our main results would hold if the shocks were normally distributed. This would nonetheless come at the cost of a stark decrease in the model's tractability.

Hence, from the parties' standpoint, each individual in a locality l has political preferences that are the result of two random shocks.¹⁶ The first, $\nu_{i,l}$, captures differences in individual-specific preferences. These shocks are independent and identically distributed draws from a locality-specific distribution. The parameter $\phi_l(> 0)$, which identifies the density of this distribution, is what is called the *swingness* of locality l . The second shock, δ_d , captures district-level shifts in preferences. These shifts represent the ex ante uncertainty faced by parties regarding their overall support in the district. From an ex ante standpoint, they only know the district's deterministic bias β_d in favor of (respectively against) B when positive (respectively negative) and the density $\gamma_d(> 0)$ of the distribution. We call γ_d the *contestability* of district d because the probability that district d is within $\varepsilon/2$ of a tie is equal to $\gamma_d\varepsilon$.

4 Equilibrium Analysis

This section characterizes the unique pure strategy equilibrium under both PR and MAJ, and discusses properties of the equilibrium allocations under the two systems.

4.1 Preliminaries

For any given district shock δ_d , we can use equation (4) to identify the swing voter in locality l :

$$\nu_l(\mathbf{q}, \delta_d) \equiv \Delta u_l(\mathbf{q}) - \delta_d.$$

Voters to the left (resp. right) of that swing voter, *i.e.* with $\nu_{i,l} < (>) \nu_l(\mathbf{q}, \delta_d)$, strictly prefer to vote A (B).

Throughout the paper, we assume that there are voters to be swung in all localities:

Assumption 1 (Interior) For all \mathbf{q} and δ_d , $\nu_l(\mathbf{q}, \delta_d) \in \left(-\frac{1}{2\phi_l}, \frac{1}{2\phi_l}\right)$ in all localities.

With this assumption (which is discussed in Appendix 1), locality-level vote shares can be

¹⁶For our purposes, adding locality-specific biases and/or a national shock would only complicate the notation without adding insight.

computed as:

$$\pi_l(\mathbf{q}; \delta_d) = \frac{1}{2} + \phi_l (\Delta u_l(\mathbf{q}) - \delta_d), \quad (5)$$

and the vote share of party B is therefore $1 - \pi_l(\mathbf{q}; \delta_d)$.

4.2 Proportional Representation System

Under PR, maximizing the expected share of seats in the national assembly is equivalent to maximizing the country-wide expected vote share. This translates into the following objective function for party A (see Appendix 1):

$$\begin{aligned} \max_{\mathbf{q}^A | \sum_l n_l^\alpha q_l = y} \pi_{PR}(\mathbf{q}) &= \sum_l \frac{t_l n_l}{T} \pi_l(\mathbf{q}; \delta_d) \\ &= \frac{1}{2} + \sum_l \frac{s_l}{T} (\Delta u_l(\mathbf{q}) - \beta_{d(l)}), \end{aligned} \quad (6)$$

where $T := \sum_k t_k n_k$ is the total number of votes (total turnout), $d(l)$ is the district to which locality l belongs, and s_l is the *electoral sensitivity* of locality l :

$$s_l := t_l n_l \phi_l. \quad (7)$$

The first order conditions are thus:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{T n_l^\alpha}{s_l} \lambda^{PR}, \quad \forall l, \quad (8)$$

where λ^{PR} is the Lagrange multiplier of the budget constraint under PR. Following the same steps for party B shows that $\mathbf{q}^A = \mathbf{q}^B$ in equilibrium. We prove that this equilibrium exists and is unique in Appendix 1.

It follows that localities with higher electoral sensitivity benefit from more government interventions. That is, they combine a large population, a high turnout, and a more ideologically homogeneous population. Comparing these conditions with the Samuelsonian conditions (3), we see that only the effect of population size n_l is identical. Any other component of electoral sensitivity introduces deviations from the social optimum.

4.3 Majoritarian System

Under MAJ, maximizing the expected share of seats in the national assembly requires the parties to maximize the number of districts won.

We first characterize a party's vote share at the district level:

$$\begin{aligned}\pi_d(\mathbf{q}; \delta_d) &= \sum_{l \in d} \frac{t_l n_l}{\sum_{k \in d} t_k n_k} \pi_l(\mathbf{q}; \delta_d) \\ &= \frac{1}{2} + \sum_{l \in d} \frac{s_l}{\sum_{k \in d} t_k n_k} (\Delta u_l(\mathbf{q}) - \delta_d),\end{aligned}\tag{9}$$

which is thus a weighted average of the locality vote shares in that district, where each locality is weighted by its number of valid ballots.

The probability that the party wins the district seat is the probability that this share is at least fifty percent. We denote it by $p_d(\mathbf{q}) := \Pr(\pi_d(\mathbf{q}; \delta_d) \geq 1/2)$. Using (9), this becomes:

$$p_d(\mathbf{q}) = \Pr\left(\delta_d \leq \sum_{l \in d} \frac{s_l}{\sum_{j \in d} s_j} \Delta u_l(\mathbf{q})\right).\tag{10}$$

To avoid corner solutions and ensure that payoffs are differentiable everywhere, throughout the paper we assume that this probability is non-degenerate for any allocation. In other words, we assume that all districts are *contestable*:¹⁷

Assumption 2 (Contestability) $p_d(\mathbf{q}) \in (0, 1), \forall d, \mathbf{q}$.

Appendix 1 identifies the parameter conditions needed for Assumption 2 to hold and shows that party A 's objective function under MAJ can then be written as:

$$\max_{\mathbf{q}^A | \sum_l n_l^\alpha q_l = y} \pi_{MAJ}(\mathbf{q}) = \frac{1}{2} + \frac{1}{D} \sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l}{\sum_{j \in d} s_j} \Delta u_l(\mathbf{q}) - \beta_d \right].\tag{11}$$

¹⁷Following Persson and Tabellini (1999) and Galasso and Nunnari (2018), we could also consider some *non-contestable* districts. Non-contestable districts are such that, for any allocation, one of the parties has a zero probability of winning – that is, $p_d(\mathbf{q}) = 0$ or $p_d(\mathbf{q}) = 1 \forall \mathbf{q}$. By definition, non-contestable districts cannot be swung, and therefore parties would not spend any of their budget on localities belonging to such districts.

The first order conditions are thus:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\sum_{j \in d(l)} s_j}{s_l} \frac{n_l^\alpha}{\gamma_{d(l)}} \lambda^{MAJ} \quad \forall l, \quad (12)$$

where λ^{MAJ} is the Lagrange multiplier associated with the budget constraint under MAJ. As in PR, there is a unique pure strategy equilibrium with $\mathbf{q}^A = \mathbf{q}^B$ (see Appendix 1).

The key difference with PR is that the localities receiving a larger share of the government budget are now the ones with a higher *relative* electoral sensitivity. In other words, parties only compare their electoral sensitivity s_l to the average sensitivity *in the same district*: $\sum_{l \in d(l)} s_j$.

4.4 Comparing the Systems

In this section, we compare government interventions under MAJ and PR systems. The following Theorem, which follows directly from Sections 4.2 and 4.3, forms the basis of the comparison:

Theorem 1 *In PR, $q_l \geq q_{l'}$ iff $s_l n_l^{-\alpha} \geq s_{l'} n_{l'}^{-\alpha}$. In MAJ, $q_l \geq q_{l'}$ iff $\gamma_{d(l)} \frac{s_l n_l^{-\alpha}}{\sum_{k \in d(l)} s_k} \geq \gamma_{d(l')} \frac{s_{l'} n_{l'}^{-\alpha}}{\sum_{k \in d(l')} s_k}$.*

Theorem 1 identifies key differences between the two systems. First, localities that are electorally more sensitive are systematically better treated in PR. In contrast, in MAJ, the electoral sensitivity of each locality is only assessed in comparison to that of the other localities in its district: It is the *relative electoral sensitivity* that matters. Second, as already emphasized by the literature, MAJ introduces the distortion that localities belonging to more contestable districts (high $\gamma_{d(l)}$) receive a disproportionately large share of the resources. Last, in the limit case of pure transfers, the effects of local population size disappear, but the effects of (relative) turnout and (relative) swingness remain.

Which of the two systems produces the highest level of inequality is thus far from clear. To better identify the comparative properties of the two systems, we analyze two dimensions of inequality in turn. First, we consider *horizontal* inequality: how government interventions differ across localities with *identical* characteristics. Second, we turn to *vertical*

inequality: how government interventions differ across localities with *different* characteristics. Throughout, we focus on $\alpha = 0$ (pure public good), since it captures the essence of the results for all $\alpha < 1$.

4.4.1 Horizontal Inequality

The most straightforward implication of Theorem 1 concerns the possible discrimination between two localities with the same characteristics. While they necessarily receive the same allocation under PR, they may receive substantially different government interventions under MAJ. This can either be the result of different district contestabilities or – and this is the novel effect identified here – because they are surrounded by different localities in their respective districts.

We illustrate the latter effect with an example that builds on the case of CRRA utility functions developed in Appendix 1. Table 1 considers utility functions $u(q_l) = \sqrt{q_l}$, and four localities grouped in two districts. To isolate the relative electoral sensitivity effect, we assume that the districts have the same contestability: $\gamma_1 = \gamma_2$.

Consider localities 2 and 3 in Table 1: they have the same electoral sensitivity s_l , but they belong to two different districts. As shown in Theorem 1, they must receive the same allocation under PR: 5.5% of the total budget in Table 1.

District	locality	Sensitivity (s_l)	q_l^{PR}	q_l^{MAJ}
1	1	0.5	1%	9%
1	2	1	5.5%	36%
2	3	1	5.5%	3%
2	4	4	88%	52%

Table 1: equilibrium allocations under PR and MAJ
 $(u(q_l) = \sqrt{q_l}, \alpha = 0)$

The allocation is noticeably different under MAJ: It is strongly skewed towards locality 2, which ends up receiving about 12 times more resources than locality 3, only because it is surrounded by *other* localities with different characteristics. Locality 2 is the most sensitive in district 1. In district 2, however, locality 3 is electorally less sensitive than

locality 4. Following the adage that “in the land of the blind, the one-eyed is king,” in MAJ, more governmental resources flow to locality 2 than to locality 3.

Theorem 1 implies that there will generically be more *horizontal* inequality under MAJ than PR. While this corroborates traditional results (see, *e.g.*, Persson and Tabellini (2000), Strömberg (2008)), the mechanism is different (relative sensitivities, instead of district contestabilities). Moreover, focusing on horizontal inequality overshadows the differences in treatment between localities with different characteristics.

4.4.2 Vertical Inequality

To isolate how MAJ and PR affect vertical inequality, we focus again on the case of equally contestable districts ($\gamma_d = \gamma, \forall d$). The differences between these localities can only be their electoral sensitivity (s_l) and their *relative* electoral sensitivity ($s_l / \sum_{k \in d(l)} s_k$). Theorem 1 tells us that, in PR, a locality with a higher s_l always receives a larger fraction of the government resources. In contrast, if districts consist of localities that are electorally more homogeneous, then MAJ will tend to produce *less* inequality.

Let us return to the example in Table 1 to illustrate: Under PR, locality 1 “competes” directly with locality 4. Since it is electorally the least sensitive of the four localities, it receives 88 times less than locality 4. Under MAJ, districts act as a fence that insulate some localities from one another. Locality 1 only competes with locality 2, and locality 3 only against locality 4. While this comes at a cost of some resources for localities 3 and 4, it substantially benefits locality 1. Under MAJ, locality 1 receives only about 6 times less than locality 4. In this example, the Gini coefficient of inequality is actually lower under MAJ than under PR.

To go beyond the example, consider the case in which each locality would form a district by itself. There are then L districts, and all localities have a *relative* electoral sensitivity equal to 1. As a consequence, they all receive the same level of government intervention as long as districts have the same contestability. Under PR, this would only be true if $s_l = s_{l'} \forall l, l'$. In other words, with a fine level of districting or large homogeneous districts, inequalities in government intervention may well disappear completely under MAJ: It induces parties

to *sprinkle public interventions all over the country*.

5 Normative Analysis

We can conclude from the previous section that either system may create the largest level of inequality, depending on the circumstances. However, looking at inequality alone does not give us the full picture: not all inequalities are socially undesirable.

To measure the social cost of politically motivated distortions, we propose to build on Atkinson (1970, 1983), who introduce a welfare-based measure of inequality. We adapt his approach to derive a measure of inequality in government interventions for the utilitarian social welfare function defined in equation (2).

Following Atkinson, we work under the assumption of CRRA preferences, with $\rho (> 0)$ denoting individual risk aversion:

$$u_l(\mathbf{q}) = \begin{cases} \ln(q_l) & \text{if } \rho = 1 \\ \frac{q_l^{1-\rho}}{1-\rho} & \text{if } \rho \neq 1. \end{cases}$$

Under these preferences, maximizing the social welfare function (2) implies that a locality l *should* receive a share $\sigma_l^{SW} = n_l^{\frac{1}{\rho}} / (\sum_j n_j^{\frac{1}{\rho}})$ of the budget y (see Appendix 1). Denoting by $\widetilde{W}(y)$ the indirect utility function that represents the result of the planner's maximization problem under budget y , we have:

$$\widetilde{W}(y) = \begin{cases} \sum_l n_l \ln(n_l) + \ln(y) & \text{if } \rho = 1 \\ \left(\sum_l n_l^{\frac{1}{\rho}} \right)^{\rho} \frac{y^{1-\rho}}{1-\rho} & \text{if } \rho \neq 1. \end{cases}$$

We then contrast the level of welfare with the one that results from the *actual* allocation of resources across localities, \mathbf{q} . We denote that level by $\mathcal{W}(\mathbf{q})$. Generically, the budget actually needed to reach that level of welfare can be reduced by reoptimizing the allocation \mathbf{q} . This allows us to define y^E as the smallest budget needed to reach the level of social welfare $\mathcal{W}(\mathbf{q})$:

$$y^E(\mathbf{q}) = \widetilde{W}^{-1}(\mathcal{W}(\mathbf{q}))$$

Following Atkinson’s approach, we use the comparison between y^E and y to measure inequality in government interventions:

$$A(\mathbf{q}) \equiv 1 - \frac{y^E(\mathbf{q})}{y} = \begin{cases} 1 - \frac{1}{y} \Pi_l (q_l/n_l)^{n_l} & \text{if } \rho = 1 \\ 1 - \left[\frac{\sum_l n_l (q_l/y)^{1-\rho}}{\left(\sum_j n_j^{\frac{1}{\rho}} \right)^\rho} \right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \quad (13)$$

This proposed measure of inequality of governmental allocation captures the social cost of politically motivated distortions: the fraction of the budget that could be saved by improving the allocation of government interventions while maintaining welfare at a constant level. At the extremes, $A(\mathbf{q})$ is 0 when the allocation is fully efficient, and 1 when it is pure waste.

Using this measure, we say that PR *Atkinson-dominates* MAJ when $A(\mathbf{q}^{PR}) < A(\mathbf{q}^{MAJ})$ and vice versa. We show in Appendix 1 that:¹⁸

Lemma 1 *PR Atkinson-dominates MAJ if and only if:*

$$\frac{\sum_l n_l (s_l)^{\frac{1-\rho}{\rho}}}{\left(\sum_k (s_k)^{\frac{1}{\rho}} \right)^{1-\rho}} \leq \frac{\sum_l n_l \left(\frac{\gamma_{d(l)} s_l}{\sum_{k \in d(l)} s_k} \right)^{\frac{1-\rho}{\rho}}}{\left(\sum_k \left(\frac{\gamma_{d(k)} s_k}{\sum_{j \in d(k)} s_j} \right)^{\frac{1}{\rho}} \right)^{1-\rho}} \text{ for } \rho \geq 1, \quad (14)$$

We can interpret each side of this inequality as the “score” of an electoral system on the Atkinson scale. The higher the score, the lower the distortion. The left-hand side is the score of PR, which only depends on the absolute sensitivity of each locality. On the right-hand side, the score of MAJ depends on district contestability and relative sensitivities within each district.

To gain further understanding, it is useful to consider specific scenarios. First, consider the scenario in which all localities have the same turnout and swingness ($s_l = k n_l, \forall l$). In that case, electoral sensitivity only varies with population size, and PR produces the socially optimal allocation: PR Atkinson-dominates MAJ. Second, consider the opposite

¹⁸In Appendix 1, we also solve for the case where $\rho = 1$.

scenario: Localities are identical in terms of population size, and districts have the same contestability ($n_l = 1/L$, $\forall l$ and $\gamma_d = \gamma$, $\forall d$), but they differ in electoral sensitivity s_l . Let us also assume that there is one locality per district, such that all districts/localities have a relative electoral sensitivity of one ($s_l / \sum_{k \in d(l)} s_k = 1$). In this case, MAJ leads to the socially optimal allocation: All localities should be and are treated equally. By contrast, due to the heterogeneity in electoral sensitivity, PR does not lead to a socially optimal allocation.

Moving to a more general comparison is complex. We have seen that the Atkinson measure of inequality increases as political forces distort the allocation further away from the social optimum. The nature of these political forces differ between MAJ and PR: under MAJ deviations are driven by differences in contestability across districts, while under PR deviations are driven by differences in electoral sensitivity, both between and within districts. Proposition 1 shows that in the case of log utility and well-apportioned districts, the comparison between the two systems boils down to comparing the spread in contestability to the spread in electoral sensitivities:

Proposition 1 *For log utility ($\rho = 1$) and well-apportioned districts ($\sum_{l \in d} n_l = 1/D \forall d$), we have that PR Atkinson-dominates MAJ if $\gamma_d / \sum_{d=1}^D \gamma_d$ is a mean preserving-spread of $s_l / \sum_{k=1}^L s_k$, and conversely.*

This difference between MAJ and PR is useful to interpret findings in the empirical literature. For instance, let us consider the analysis in Strömberg (2008). He finds that replacing the Electoral College with the National Popular Vote (essentially a switch from MAJ to PR, as we discuss in Section 6.3) would lead to a decrease in cross-states inequalities in the allocation of campaign resources, and in policies implemented by Presidents while in office. This is exactly what our model predicts. Indeed, for the elections he analyses, cross-state differences in electoral contestability in the US are substantially larger than cross-state differences in electoral sensitivity.¹⁹ But, our model also suggests that the effect of such a reform could vary over time if, for instance, differences in electoral

¹⁹As we discuss below, he works under the assumption that campaign resources are targetable at the district level. He thus abstracts from the effect of within-state heterogeneity in electoral sensitivity.

contestability were to become less stark.²⁰

6 Extensions

This section explores different extensions of our model. The first extension allows for two instruments of government intervention: one that is targetable at the local level and a global public good. The second modifies the objective of parties. The third considers the implications of our model for the debate about the US Electoral College. The fourth studies a variant of PR systems. The last extension considers other dimensions of heterogeneity among voters.

6.1 Targeted versus Universal Spending

This section models politicians choosing the allocation of two types of policy instruments: targeted transfers and a global public good. Our purpose here is to highlight the role of the sprinkling effect and district contestability in the choice between those two instruments, in line with the questions raised by Lizzeri and Persico (2001) and Persson and Tabellini (2000).

Following Persson and Tabellini (2000), we assume that individuals in locality l have quasi-linear preferences in a transfer q_l (corresponding to $\alpha = 1$ in the previous setup) and a global public good that benefits the entire population:

$$w_l(\mathbf{q}, G) = q_l + u(G),$$

with $u(\cdot)$ strictly increasing and strictly concave in G . The budget is exogenously given as y so that the budget constraint becomes $\sum_l n_l q_l + G = y$.

As shown in detail in Appendix 1, in general only one locality receives transfers. In the unique equilibrium under PR, this is the locality with the highest s_l/n_l . If some transfers

²⁰There are substantial variations over time of the differences in electoral contestability across states (see, e.g., [goo.gl/f3nU7Q](https://www.google.com/search?q=contestability+variations+over+time)).

are given, then:

$$u'(G) = \max_l \frac{s_l}{n_l} \frac{1}{\sum_{k=1}^L s_k}. \quad (15)$$

Under MAJ, the equivalent FOC characterizing the equilibrium if some transfers are given is:

$$\max_l \frac{\gamma_{d(l)}}{\sum_{d \in D} \gamma_d} \frac{1}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} = u'(G). \quad (16)$$

By comparing (33) and (35), we can identify whether PR or MAJ leads to the largest provision of the global public good. To simplify the comparison, assume that localities corresponds to electoral districts, $L = D$ and that districts are well apportioned $n_l = 1/D$ for all l .

There are two effects pushing in opposite directions. On the one hand, there is the effect of district contestability in MAJ, as identified by Lizzeri and Persico (2001) and Persson and Tabellini (2000). Heterogeneous contestabilities increase transfers and decrease the provision of the national public good in MAJ versus PR. If all localities are identical in s_l then there is no transfer under PR and heterogeneous district contestabilities make transfers more attractive in MAJ.

On the other hand, there is the relative electoral sensitivity effect. When targetability of transfer is at the district level, i.e., there is one locality per district, the incentive to sprinkle resources across the country in MAJ is maximal (all relative electoral sensitivities are equal to one). If $\gamma_d = \gamma$ for all d , it is easy to see that a sufficient condition under which $G^{PR} \leq G^{MAJ}$ in equilibrium – with a strict inequality when $G^{MAJ} > 0$ – is that the highest level of electoral sensitivity is larger than the average. It follows that, in this case, *any* heterogeneity in electoral sensitivities leads to higher provision of the global public good in MAJ] than in PR, a reversal of the standard result in the literature.

The following proposition summarizes the comparison between the two systems:

Proposition 2 *If targetability is at the district level ($L = D$) and districts are well-apportioned ($n_l = \frac{1}{D}$ for all l) then $\max_d \frac{s_d}{\sum_{d' \in D} s_{d'}} > (<) \max_d \frac{\gamma_d}{\sum_{d' \in D} \gamma_{d'}}$ implies $G^{MAJ} \geq (<) G^{PR}$.*

6.2 The Objective of Parties

So far, we have worked under the assumption that parties maximize their expected seat share. We now discuss the validity of this assumption and then show that our main results are robust to parties maximizing their probability of winning a majority of seats instead.

Some political economy models assume that parties maximize their probability of obtaining a majority of seats in MAJ and their expected vote share in PR (see, *e.g.*, Lizzeri and Persico (2001), Strömberg (2008)). The main motivation for using system-specific utility functions is the perception that the party winning a majority of seats obtains an extra payoff under MAJ as compared to PR. As discussed in Snyder (1989), modeling MAJ in this way highlights the pivotability of a seat/district in the national assembly. However, just because a party has a one-seat majority in the legislative assembly does not automatically mean it can pass all the legislation it wants (one case in point is the current situation in the US Senate). Passing legislation is typically much easier when it has a comfortable super-majority. Hence, even in MAJ, parties benefit from earning extra seats beyond simple majority, a benefit that we are trying to capture with our objective function. Finally, there is empirical evidence in support of our assumption that parties maximize the number of seats in the national assembly (see Jacobson (1985) and Incerti (2015)).

To address any further potential concerns about this assumption, we study the case of parties that maximize their probability of winning a majority of seats in the national assembly, both for the PR and the MAJ systems. Under PR and MAJ respectively, the parties' objective functions (6) and (11) become:

$$\text{In PR: } \max_{\mathbf{q}} \frac{1}{2} + \Pr \left[\sum_l \frac{s_l}{T} (\Delta u_l(\mathbf{q}) - \beta_{d(l)}) \geq 0 \right], \quad (17)$$

$$\text{In MAJ: } \max_{\mathbf{q}} \Pr \left[\sum_d \mathbb{1}_d \geq \frac{D}{2} \right], \quad (18)$$

where $\mathbb{1}_d$ takes value 1 if $\pi_d(\mathbf{q}; \delta_d) \geq 1/2$, and 0 otherwise.

The objective function (17) under PR is just a monotone transformation of the original objective function (6). For this reason, it produces the same first order conditions, and

therefore the same equilibrium allocations as in Section 4.

The differences are more consequential under MAJ, where achieving a majority in each district separately no longer matters. Winning a given district only matters insofar as it helps reach the threshold of 50% of all districts.

As explained in Lindbeck and Weibull (1987) and Strömberg (2008), this problem is technically intractable. However, we can focus on its approximate solution, which exploits Lyapunov’s central limit theorem. In Appendix 1, we detail how this can be applied to our model in the case of a large enough number of districts. Defining:

$$\sigma_E^2(\mathbf{q}) := \sum_d p_d(\mathbf{q}) [1 - p_d(\mathbf{q})]$$

to be the variance of the distribution of seat shares, and letting λ' be the Lagrange multiplier, we find that the equilibrium allocation must satisfy:

$$\lambda' = \gamma_{d(l)} \frac{s_l n_l^{-\alpha}}{\sum_{j \in d(l)} s_j} u'(q_l) \left[1 + \frac{\sum_d \gamma_d \beta_d}{\sigma_E^2(\mathbf{q})} \gamma_{d(l)} \beta_{d(l)} \right], \quad (19)$$

which directly compares to (12), the FOC under MAJ. We see that the two are identical except for the second term inside the square bracket. This implies that both the relative electoral sensitivity of localities and the contestability of districts are still key in explaining government interventions.

The second term in the square bracket has a natural interpretation. The fraction denotes the average, national, bias in favor of B : If positive, B is more likely to win than A , and vice versa. Let us assume it is positive for the sake of discussion. In this case, the localities benefiting from more government interventions are those belonging to districts that are more contestable *and* also biased toward B ($\gamma_{d(l)} \beta_{d(l)}$ large). This is the same “pivotality effect” as the one identified in Lindbeck and Weibull (1987, pp288-289): “[District d] is more likely to be a pivot [district] the stronger is [its] bias in favour of the more popular party, since the exclusion of such [a district] from the electorate leaves the remaining electorate as little biased as possible, and hence also as likely as possible to produce a tie.”

6.3 Electoral College *vs.* National Popular Vote

The election of the U.S. president is organized through the Electoral College: general elections are organized at the State level, and each State is awarded a number of Electors equal to its number of representatives in Congress (the sum of the number of Senators and of House Representatives). The rule to select the Electors is at the discretion of the individual States but, today, it is Winner-takes-all in most States (the exceptions are Maine and Nebraska).

There are two main reasons why this indirect electoral system creates frustrations today: (i) a candidate who loses the Popular Vote can still win the election and, (ii) the United States eventually became divided into “safe” and “swing” states, which induces candidates to devote too much resources on the latter.

One oft-proposed reform is to replace the Electoral College by the direct election of the President under the National Popular Vote rule. Our model allows us to compare the effectiveness of that proposal with the current system, and a broader set of reforms, which include a re-weighting of the Electoral College and making the allocation of Electors proportional.

The current form of the US Electoral College can be represented as a weighted majoritarian system, in which each state is represented by a district d , and its share of the 538 electors is represented by ω_d . The objective function of party A looks like:

$$\max_{\mathbf{q}^A | \sum_l n_l^\alpha q_l = y} \pi^{\text{College}}(\mathbf{q}) = \frac{1}{2} + \frac{1}{|C|} \sum_d \gamma_d \omega_d \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q}) + \zeta_d}{\mathbf{s}_d} - \beta_d \right],$$

where swing (resp. safe) districts are associated with high (resp. low) values of γ_d .

It is straightforward to see that the Electoral College tilts the allocation of government resources towards districts with a higher ω_d . Indeed, the FOCs are:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\lambda^{\text{College}}}{\omega_{d(l)}} \frac{\mathbf{s}_{d(l)}/\gamma_{d(l)}}{s_l/n_l^\alpha}, \quad \forall l, \quad (20)$$

where the only difference with MAJ is the presence of the weight $\omega_{d(l)}$.

Reforming the electoral college to replace it with the National Popular Vote rule would instead amount to merging all the existing districts, into a single one. In terms of the resulting allocation of government interventions, we find that it is outcome-equivalent to a reform towards a the PR system analyzed in Section 3.3.

Using these first order conditions allows us to identify how the reform may affect government interventions. The effect varies with the nature of within- and between-state heterogeneity. If between-state differences among localities are severe, and differences in contestability (as measured by γ_d) are mild, switching to the National Popular Vote may actually amplify distortions: it is easy to build examples in which the Electoral College Atkinson-dominates the National Popular Vote. By contrast, if between-state differences among localities are mild (the s_l are similar), and differences in contestability are severe (the γ_d are very different), then a move to the Popular Vote should reduce distortions, and thus be socially preferred.

These results shed a new light on the findings of Strömberg (2008) that replacing the Electoral College with the National Popular Vote would lead to an overall decrease in cross-states inequalities in the allocation of campaign resources. He works under the assumption that campaign resources are only targetable at the state level. This implies that his analysis abstracts from within-state variations in electoral sensitivity, a source of socially undesirable inequalities under the National Popular Vote but not the Electoral College. This introduces a risk of underestimating the inequalities that would result from pooling localities from different states in a single electoral district under the National Popular Vote.

Beyond this pre-existing proposal, our model allows us to assess a broader set of reforms and identify the socially optimal electoral system:

Proposition 3 *Starting from the Electoral College system, there exists a combination of redistricting and reapportionment that implements the social optimum.*

The underlying rationale for this result is (i) that an arbitrary reweighting of the Electoral College can always tilt the politicians' incentives towards any district, and (ii) that a division into more districts increases the number of weights available. Hence, there must

exist some districting that offers sufficiently many instruments (district weights) to reach the social optimum. In contrast, the National Popular Vote restricts the number of districts to a single one.

In practice, delivering such a reform may run into first-order obstacles. First, a modification of both the districts and the distribution of Electors in the Electoral College requires a constitutional change. Second, the optimal weights may substantially differ from “one (wo)man, one vote”. Last but not least, the optimal weights are likely to change over time (when the electoral sensitivity of localities change), which would require constant reform.

We can also show (see Appendix 1) that restoring proportionality in each state election for the Electoral College may deliver the same, and potentially better, outcomes than the National Popular Vote. Such a reform also has the advantage that it need not require any constitutional amendment.

6.4 Other Proportional Representation Systems

In line with most of the literature, thus far we have assumed that, in PR, the number of seats that a party obtains is proportional to its total number of votes in the population. While this is a good representation of, for instance, the Dutch electoral system, some countries instead use a district-specific proportional election system. In Belgium or Brazil, for instance, each electoral district is entitled to a pre-determined seat share that is proportional to each district’s total population.

We can extend our model to these *district-specific* PR systems by allowing each district to receive some arbitrary fraction μ_d of the seats, with $\sum_d \mu_d = 1$. The objective function (6) then becomes:

$$\max_{\mathbf{q}} \pi_{PR}^{\text{districts}}(\mathbf{q}) = \frac{1}{2} + \sum_d \mu_d \sum_{l \in d} \frac{s_l}{m_d} [\Delta u_l(\mathbf{q}) - \mathbf{E}[\delta_d]],$$

where m_d is the total number of active voters in the district. Defining the average turnout rate in a district as $t_d := \sum_{l \in d} t_l \frac{n_l}{n_d}$, with $n_d := \sum_{l \in d} n_l$, we obtain $m_d = t_d n_d$. Taking first order conditions and letting λ^{DPR} denote the multiplier on the budget constraint, we

have:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \left[\frac{\mu_d}{n_d} \frac{s_l}{t_d} \right]^{-1} n_l^\alpha \lambda^{DPR} \forall l, \quad (21)$$

where $\frac{\mu_d}{n_d}$ is equal to 1 when seat shares are perfectly apportioned, and above/below 1 when the district is over/under-apportioned. The second fraction, $\frac{s_l}{t_d}$, is the electoral sensitivity of the locality relative to the district turnout.

We now see how the nationwide and district-specific versions of PR differ in terms of government interventions. In the nationwide version of PR, μ_d is implicitly made equal to the number of voters: $n_d t_d$. For this reason, district borders become immaterial to parties' platforms. In the district-specific version of PR, each locality's turnout is compared to the turnout of the other localities in the same district. Moreover, *ceteris paribus*, localities in over-apportioned districts will receive more than those in under-apportioned districts.

District-specific PR systems thus share features with both systems in Section 4. Like in MAJ, a high-turnout locality will receive less if it is located in a higher, as opposed to lower, turnout district. The other results remain identical: District contestability and the relative swingness of the locality are immaterial to the eventual allocation of governmental resources in PR.²¹

6.5 Other Dimensions of Heterogeneity.

In our baseline model, we considered only two sources of heterogeneity among voters of different localities: their swingness and their turnout rate. This was done for the sake of expositional clarity. We could easily consider other sources of heterogeneity, such as information and partisanship.

Information. Following Strömberg (2004), we could assume that some voters do not observe the parties' proposals by the time of the vote. For each locality l , parties would then assign a probability χ_l that a voter knows the parties' proposals. In that version of the model, the electoral sensitivity of locality l , s_l , would include the parameter χ_l . The level of information voters have would then influence the allocation of governmental

²¹The same holds for other possible dimensions of heterogeneity, such as information and partisanship, discussed in Section 6.5.

resources in the same way swingness does under both PR and MAJ.

Partisan Voters. A reality that we did not incorporate in our baseline model is the coexistence of swing and partisan voters (i.e., the fact that different population groups can have different party affinities (*e.g.*, they are from the same ethnic or cultural group), see the discussion in Dixit and Londregan (1996)). Partisan voters are those who have such a strong preference for or against a party that they would never modify their ballot based on government interventions in their locality.

Formally, we could assume that each voter in locality l has a probability μ_l of being partisan, and that a fraction f_l^μ of these partisans vote for A . The others vote for B . In that version of the model, the electoral sensitivity of locality l , s_l , would include the parameter $1 - \mu_l$. The fraction of partisan voters in a locality would then influence the allocation of governmental resources in the opposite way swingness does under both PR and MAJ. In particular, in MAJ, the *higher* the relative fraction of partisan voters in a locality (compared to other localities in the same district), the *lower* the allocation of governmental resources to that localities. By contrast, under PR the fraction of partisan voters in other localities in the same district does not have any influence.

7 Conclusions

In this paper, we studied the impact of the electoral system on inequalities in government interventions. We compared majoritarian (MAJ) and proportional representation (PR) systems. The main novelty of our approach is that we take account of the fact that the geographic distribution of votes matters more in MAJ systems. We uncovered a novel *relative electoral sensitivity effect* in MAJ systems that can induce parties to “sprinkle” resources across districts, thus reducing inequality. When this effect is strong enough, inequalities in government intervention end up being *lower* in MAJ than PR systems. This result runs against a recurrent theme in the literature that parties target a smaller fraction of the population under MAJ systems. The same effect can result in some cases in lower incentives to provide a broad public good, as opposed to targeted transfers, in PR systems.

Our approach yields useful insights into on-going debates about the US Electoral College, and whether the president should be elected by National Popular Vote (NPV). Strömberg (2008) shows that states with high weights and high pivotability benefit from the Electoral College system compared to the NPV. Our model predicts that, in addition to these effects, localities that have a high relative electoral sensitivity compared to their absolute electoral sensitivity benefit from the status quo. This could generate additional inequalities in government intervention under NPV. Quantifying these effects would help informing the current debate.

The relative sensitivity effect has implications for the existing empirical literature on distributive politics (see, e.g., the literature reviews in Berry et al. (2010) and Golden and Min (2013)). It implies that there is a risk of omitted variable bias in studies of the allocations of governmental resources at the sub-district level in MAJ systems that do not control for the electoral sensitivity of other groups of voters in the same district.

Finally, our tractable framework will allow us to explore the role of the relative sensitivity effect on other important issues such as the gerrymandering process or the effect of redistricting on the allocation of governmental resources. In both cases, the composition of the electoral districts, which directly affects the relative electoral sensitivity, is indeed at the heart of the problem.

References

- Aidt, T.S., Jayasri Dutta, and Elena Loukoianova (2006) “Democracy comes to Europe: Franchise extension and fiscal outcomes 1830-1938,” *European Economic Review*, Vol. 50, No. 2, pp. 249–283.
- Albouy, David (2009) “The Unequal Geographic Burden of Federal Taxation,” *Journal of Political Economy*, Vol. 117, No. 4, pp. 635–667.
- Alesina, Alberto, Reza Baqir, and William Easterly (1999) “Public Goods and Ethnic Divisions,” *The Quarterly Journal of Economics*, Vol. 114, No. 4, pp. 1243–1284.
- Ansolabehere, Stephen, Alan Gerber, and James Snyder (2002) “Equal Votes, Equal Money: Court-Ordered Redistricting and Public Expenditures in the American States,” *The American Political Science Review*, Vol. 96, No. 4, pp. 767–777.
- Atkinson, A.B. (1970) “On the Measurement of Inequality,” *Journal of Economic Theory*, pp. 244–263.
- (1983) *The Economics of Inequality*: Clarendon Press, II ed. Oxford, UK.
- Azzimonti, Marina (2015) “The dynamics of public investment under persistent electoral advantage,” *Review of Economic Dynamics*, Vol. 18, No. 3, pp. 653–678.
- Banerjee, Abhijit, Lakshmi Iyer, and Rohini Somanathan (2008) “Public Action for Public Goods,” in Schultz, T. Paul and John A. Strauss eds. *Handbook of Development Economics*, Vol. 4: Elsevier, Chap. 49, pp. 3117–3154.
- Banks, Jeffrey S. and John Duggan (1999) “The Theory of Probabilistic Voting in the Spatial Model of Elections,” working paper, University of Rochester.
- Berry, Christopher R., Barry C. Burden, and William G. Howell (2010) “The President and the Distribution of Federal Spending,” *American Political Science Review*, Vol. 104, No. 4, p. 783799.
- Besley, Timothy and Robin Burgess (2002) “The Political Economy of Government Responsiveness: Theory and Evidence from India,” *The Quarterly Journal of Economics*, Vol. 117, No. 4, pp. 1415–1451.

- Betz, Timm (2017) “Trading Interests: Domestic Institutions, International Negotiations, and the Politics of Trade,” *The Journal of Politics*, Vol. 79, No. 4, pp. 1237–1252.
- Blakeslee, David S. (2015) “Politics and Public Goods in Developing Countries: Evidence from the Assassination of Rajiv Gandhi,” working papers, New York University.
- Blume, Lorenz, Jens Möller, Stefan Voigt, and Carsten Wolf (2009) “The economic effects of constitutions: replicating and extending Persson and Tabellini,” *Public Choice*, Vol. 139, No. 1, pp. 197–225.
- Bormann, Nils-Christian and Matt Golder (2013) “Democratic Electoral Systems around the world, 1946-2011,” *Electoral Studies*, Vol. 32, No. 2, pp. 360 – 369.
- Carey, John M (2007) “Competing principals, political institutions, and party unity in legislative voting,” *American Journal of Political Science*, Vol. 51, No. 1, pp. 92–107.
- Castanheira, Micael, Gaëtan Nicodème, and Paola Profeta (2012) “On the political economics of tax reforms: survey and empirical assessment,” *International Tax and Public Finance*, Vol. 19, No. 4, pp. 598–624.
- Chang, Eric C. C., Mark Andreas Kayser, and Ronald Rogowski (2008) “Electoral Systems and Real Prices: Panel Evidence for the OECD Countries, 1970 2000,” *British Journal of Political Science*, Vol. 38, No. 04, pp. 739–751.
- Chhibber, Pradeep and Irfan Nooruddin (2004) “Do Party Systems Matter? The Number of Parties and Government Performance in the Indian States,” *Comparative Political Studies*, Vol. 37, No. 2, pp. 152–187.
- Colomer, Josep M. (2004) *The Strategy and History of Electoral System Choice*, pp. 3–78, London: Palgrave Macmillan UK.
- Cox, Gary W. (1990) “Centripetal and Centrifugal Incentives in Electoral Systems,” *American Journal of Political Science*, Vol. 34, No. 4, pp. 903–935.
- Dewan, Torun and Arthur Spirling (2011) “Strategic opposition and government cohesion in Westminster democracies,” *American Political Science Review*, Vol. 105, No. 2, pp. 337–358.

- Dixit, Avinash and John Londregan (1996) “The Determinants of Success of Special Interests in Redistributive Politics,” *The Journal of Politics*, Vol. 58, No. 4, pp. 1132–1155.
- Evans, Carolyn L. (2009) “A Protectionist Bias In Majoritarian Politics: An Empirical Investigation,” *Economics & Politics*, Vol. 21, No. 2, pp. 278–307.
- Finan, Frederico and Maurizio Mazzocco (2016) “Electoral Incentives and the Allocation of Public Funds,” Working Paper 21859, National Bureau of Economic Research.
- Funk, Patricia and Christina Gathmann (2013) “How Do Electoral Systems Affect Fiscal Policy? Evidence From Cantonal Parliaments, 1890-2000,” *Journal of the European Economic Association*, Vol. 11, No. 5, pp. 1178–1203.
- Gagliarducci, Stefano, Tommaso Nannicini, and Paolo Naticchioni (2011) “Electoral Rules and Politicians’ Behavior: A Micro Test,” *American Economic Journal: Economic Policy*, Vol. 3, No. 3, pp. 144–74.
- Galasso, Vincenzo and Salvatore Nunnari (2018) “The Economic Effects of Electoral Rules: Evidence from Unemployment Benefits,” CEPR Discussion Papers 13081, C.E.P.R. Discussion Papers.
- Golden, Miriam A. and Brian K. Min (2013) “Distributive Politics Around the World,” *Annual Review of Political Science*, Vol. 16, pp. 73–99.
- Grossman, Gene M. and Elhanan Helpman (2005) “A Protectionist Bias in Majoritarian Politics,” *The Quarterly Journal of Economics*, Vol. 120, No. 4, pp. 1239–1282.
- Hatfield, John William and William R. Hauk (2014) “Electoral regime and trade policy,” *Journal of Comparative Economics*, Vol. 42, No. 3, pp. 518–534.
- Hodler, Roland and Paul A. Raschky (2014) “Regional Favoritism*,” *The Quarterly Journal of Economics*.
- Incerti, D. (2015) “The Optimal Allocation of Campaign Funds in House Elections.,” working paper.
- Iversen, Torben and David Soskice (2006) “Electoral Institutions and the Politics of Coalitions: Why Some Democracies Redistribute More Than Others,” *American Political Science Review*, Vol. 100, No. 2.

- Jacobson, Gary C. (1985) “Party Organization and Distribution of Campaign Resources: Republicans and Democrats in 1982,” *Political Science Quarterly*, Vol. 100, No. 4, pp. 603–625.
- Johansson, Eva (2003) “Intergovernmental grants as a tactical instrument: empirical evidence from Swedish municipalities,” *Journal of Public Economics*, Vol. 87, No. 5, pp. 883 – 915.
- Keefer, Philip (2004) “What does political economy tell us about economic development and vice versa?” *Annual Review of Political Science*, Vol. 7, pp. 247–272.
- Krasa, Stefan and Mattias K Polborn (2018) “Political competition in legislative elections,” *American Political Science Review*, Vol. 112, No. 4, pp. 809–825.
- Lindbeck, Assar and Jörgen Weibull (1987) “Balanced-budget redistribution as the outcome of political competition,” *Public Choice*, Vol. 52, No. 3, pp. 273–297.
- Lizzeri, Alessandro and Nicola Persico (2001) “The Provision of Public Goods under Alternative Electoral Incentives,” *American Economic Review*, Vol. 91, No. 1, pp. 225–239.
- (2005) “A Drawback Of Electoral Competition,” *Journal of the European Economic Association*, Vol. 3, No. 6, pp. 1318–1348.
- Mansfield, Edward D. and Marc L. Busch (1995) “The political economy of nontariff barriers: a cross-national analysis,” *International Organization*, Vol. 49, No. 04, pp. 723–749.
- Martin, Paul S. (2003) “Voting’s Rewards: Voter Turnout, Attentive Publics, and Congressional Allocation of Federal Money,” *American Journal of Political Science*, Vol. 47, No. 1, pp. 110–127.
- Milesi-Ferretti, Gian Maria, Roberto Perotti, and Massimo Rostagno (2002) “Electoral Systems and Public Spending,” *The Quarterly Journal of Economics*, Vol. 117, No. 2, pp. 609–657.
- Min, Brian (2015) “Electrifying the Poor: Power and Politics in India,” Technical report.
- Myerson, Roger B. (1993) “Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems,” *The American Political Science Review*, Vol. 87, No. 4, pp. 856–869.

- Park, Jong Hee and Nathan Jensen (2007) “Electoral Competition and Agricultural Support in OECD Countries,” *American Journal of Political Science*, Vol. 51, No. 2, pp. 314–329.
- Persson, Torsten (2002) “Do Political Institutions Shape Economic Policy?,” *Econometrica*, Vol. 70, No. 3, pp. 883–905.
- Persson, Torsten and Guido Tabellini (1999) “The size and scope of government:: Comparative politics with rational politicians,” *European Economic Review*, Vol. 43, No. 4-6, pp. 699–735.
- (2000) *Political Economics: Explaining Economic Policy*, MIT Press Books: The MIT Press.
- (2003) *The Economic Effects of Constitutions*, Vol. 1 of MIT Press Books: The MIT Press.
- Ray, Debraj & Arthur Robson (2018) “Certified Random: A New Order for Coauthorship,” *American Economic Review*, Vol. 108, No. 2, pp. 489–520.
- Rickard, Stephanie (2012) “A Non-Tariff Protectionist Bias in Majoritarian Politics: Government Subsidies and Electoral Institutions,” Vol. 56, pp. 777–785.
- Rogowski, Ron and Mark Kayser (2002) “Majoritarian electoral systems and consumer power: price-level evidence from the OECD countries,” *American Journal of Political Science*, Vol. 46, No. 3, pp. 526–539.
- Rogowski, Ronald (1987) “Trade and the variety of democratic institutions,” *International Organization*, Vol. 41, No. 02, pp. 203–223.
- Schady, Norbert R. (2000) “The Political Economy of Expenditures by the Peruvian Social Fund (FONCODES), 1991-95,” *The American Political Science Review*, Vol. 94, No. 2, pp. 289–304.
- Snyder, James M (1989) “Election Goals and the Allocation of Campaign Resources,” *Econometrica*, Vol. 57, No. 3, pp. 637–660.
- Snyder, James M and Michael M Ting (2002) “An informational rationale for political parties,” *American Journal of Political Science*, Vol. 46, No. 1, pp. 90–110.

- Stashko, Allison (2018) “Crossing the District Line: Border Mismatch and Targeted Redistribution,” *Georgetown mimeo*.
- Stratmann, Thomas and Martin Baur (2002) “Plurality Rule, Proportional Representation, and the German Bundestag: How Incentives to Pork-Barrel Differ across Electoral Systems,” *American Journal of Political Science*, Vol. 46, No. 3, pp. 506–514.
- Strömberg, David (2004) “Radio’s Impact on Public Spending,” *The Quarterly Journal of Economics*, Vol. 119, No. 1, pp. 189–221.
- (2008) “How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida,” *American Economic Review*, Vol. 98, No. 3, pp. 769–807.
- Taagepera, Rein and Matthew Soberg Shugart (1989) *Seats and Votes: The effects and determinants of electoral systems*: Yale University Press.
- Troiano, Ugo (2017) “Do Taxes Increase Economic Inequality? A Comparative Study Based on the State Personal Income Tax,” Working Paper 24175, National Bureau of Economic Research.
- Tsebelis, George (1995) “Decision Making in Political Systems: Veto Players in Presidentialism, Parliamentarism, Multicameralism and Multipartyism,” *British Journal of Political Science*, Vol. 25, No. 3, p. 289325.
- Whitaker, Paige and Thomas H. Neale (2004) “The Electoral College: An Overview and Analysis of Reform Proposals,” CRS Report for Congress RL30804, Congressional Research Service.
- World Bank (2004) *World Development Report 2004: Making Services Work for Poor People*: World Bank and Oxford University Press.

Appendix 1: Theory

Assumption 1

Assumption 1 posits that there are swingable voters in any localities, that is:

$$\tilde{v}_l(\mathbf{q}, \delta) \equiv \Delta u_l(\mathbf{q}) - \delta_d \in \left(-\frac{1}{2\phi_l}, \frac{1}{2\phi_l}\right)$$

for all \mathbf{q} and δ . Let $\bar{\Delta} = u(y) - u(0)$ be the largest possible utility difference coming from the allocation of public goods. There are always some swing voters in l if

$$-\bar{\Delta} - \beta_d - \frac{1}{2\gamma_d} > -\frac{1}{2\phi_l} \quad \& \quad \bar{\Delta} - \beta_d + \frac{1}{2\gamma_d} < \frac{1}{2\phi_l}.$$

Notice that the first (second) inequality is more likely to bind if β_d is positive (negative). The assumption is satisfied if

$$|\beta_d| < -\bar{\Delta} - \frac{1}{2\gamma_d} + \frac{1}{2\phi_l}.$$

Assumption 1 requires the variance in the individual preference to be large enough compared to the bias.

Objective in PR

Parties maximize their expected nationwide vote share. The vote share of party A is the weighted average of its locality vote shares: $\sum_l \frac{t_l n_l}{T} \pi_l(\mathbf{q}; \delta_{d(l)})$, where $T := \sum_k t_k n_k$ is the total number of votes in the country.

Recall that $s_l = n_l t_l \phi_l$ is the *electoral sensitivity* of locality l . By (5), party A 's objective function is thus:

$$\begin{aligned} \max_{\mathbf{q} | \sum_l n_l^\alpha q_l \leq y} \pi_{PR}(\mathbf{q}) &:= \mathbb{E}_\delta \left[\sum_l \frac{t_l n_l}{T} \left[\frac{1}{2} + \phi_l (\Delta u_l(\mathbf{q}) - \delta_{d(l)}) \right] \right] \\ &= \frac{1}{2} + \frac{1}{T} \sum_l s_l (\Delta u_l(\mathbf{q}) - \mathbb{E}[\delta_{d(l)}]), \\ &= \frac{1}{2} + \frac{1}{T} \sum_l s_l (\Delta u_l(\mathbf{q}) - \beta_{d(l)}), \end{aligned} \tag{22}$$

where $d(l)$ is the district to which locality l belongs.²²

Assumption 2

The set of contestable districts is $C \equiv \{d | p_d^A(\mathbf{q}) \in]0, 1[\forall \mathbf{q}\}$. Therefore, a district is contestable if and only if:

$$\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\mathbf{s}_d} \in \left[\beta_d - \frac{1}{2\gamma_d}, \beta_d + \frac{1}{2\gamma_d}\right].$$

²²For this last equality, note that $\sum_d \beta_d \sum_{l \in d} s_l$ can be rewritten as $\sum_l s_l \beta_{d(l)}$.

Let $\overline{\Delta U}_d = \max_{\mathbf{q}^A | \sum_l q_l^A = y} \sum_{l \in d} \frac{s_l}{s_d} (u_l(\mathbf{q}^A) - u_l(\mathbf{0}))$ be the largest possible utility gain in the district coming from the allocation of public goods. The district is contestable if

$$-\overline{\Delta U}_d \geq \beta_d - \frac{1}{2\gamma_d} \quad \& \quad \overline{\Delta U}_d \leq \beta_d + \frac{1}{2\gamma_d}.$$

Notice that the first (second) inequality is more likely to bind if β_d is positive (negative). Hence, the assumption is satisfied iff: $\overline{\Delta U}_d + |\beta_d| \leq \frac{1}{2\gamma_d}$. That is, to be contestable, the variance of the district shock must be large enough compared to the bias.

Objective in MAJ

Under MAJ, seats are proportional to the number of districts won by each party. From (9), the probability that A wins at least 50% of the votes in district d is:

$$\begin{aligned} p_d(\mathbf{q}) &= \Pr \left[\sum_{l \in d} \frac{s_l}{\sum_{k \in d} t_k n_k} (\Delta u_l(\mathbf{q}) - \delta_d) \geq 0 \right] \\ &= \Pr \left[\delta_d \leq \frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} \right], \end{aligned} \quad (23)$$

where the second line is obtained by multiplying both sides of the inequality by $\sum_{k \in d} t_k n_k$, and isolating δ_d .

Under Assumption 2, this probability is always strictly between 0 and 1, and can be directly derived from the CDF of a uniform distribution:

$$\begin{aligned} F_{\delta_d} \left[\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} \right] &= \gamma_d \times \left[\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} + \frac{1}{2\gamma_d} - \beta_d \right] \\ &= \frac{1}{2} + \gamma_d \times \left[\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right] \end{aligned} \quad (24)$$

Aggregating these probabilities across districts yields A 's expected seat share:

$$\pi_{MAJ}(\mathbf{q}) = \frac{1}{2} + \frac{\sum_d \gamma_d \left[\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]}{D}$$

Equilibrium Existence and Uniqueness

The set of feasible allocations $Q = \{\mathbf{q} | \sum_l n_l^\alpha q_l \leq y\}$ is compact and convex. Let's define the expected plurality shares a la Banks and Duggan (1999) $P_l^A(\mathbf{q}) = 2s_l (\Delta u_l(\mathbf{q}) - \mathbb{E}[\delta_{d(l)}]) - n_l t_l$ and $P_l^B(\mathbf{q}) = n_l t_l - 2s_l (\Delta u_l(\mathbf{q}) - \mathbb{E}[\delta_{d(l)}])$. Since $P_l^A(\mathbf{q})$ and $P_l^B(\mathbf{q})$ are jointly continuous in \mathbf{q} , $P_l^j(\mathbf{q})$ is strictly concave in \mathbf{q}^j for $j \in \{A, B\}$ and $P_l^A(\mathbf{q}) + P_l^B(\mathbf{q})$ is constant for all \mathbf{q} then Theorems 2 and 3 of Banks and Duggan (1999) guarantee existence and uniqueness of the equilibrium. The argument for existence and uniqueness of the equilibrium is the same for PR.

Example: CRRA utility function

This appendix constructs the explicit solution of the welfare optimum and of the equilibrium under each electoral system for a particular utility function: the Constant Relative Risk Aversion (CRRA) utility functions:

$$\begin{aligned} u(q_l) &= \frac{q_l^{1-\rho}}{1-\rho}, \text{ if } \rho \neq 1 \\ &= \log q_l, \text{ if } \rho = 1. \end{aligned}$$

For simplicity, we focus on the case of pure local public goods ($\alpha = 0$) so that the budget constraint is: $\sum_l q_l \leq y$.

In this case, (3) tells us that the socially optimal allocation of public goods is:

$$q_l^* = y \left(\frac{n_l^{\frac{1}{\rho}}}{\sum_k n_k^{\frac{1}{\rho}}} \right). \quad (25)$$

Under PR, the FOCs in (8) and some straightforward manipulations produces the following allocation of public goods:

$$q_l^{PR} = y \left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^L (s_k)^{1/\rho}} \right), \quad (26)$$

where $s_l = n_l t_l \phi_l$ is the *electoral sensitivity* of locality l .

Finally, the FOCs for MAJ (12) yield:

$$q_l^{MAJ} = y \left(\frac{(\gamma_{d(l)} s_l / \mathbf{s}_{d(l)})^{1/\rho}}{\sum_{k=1}^L (\gamma_{d(k)} s_k / \mathbf{s}_{d(k)})^{1/\rho}} \right), \quad (27)$$

where $\mathbf{s}_d = \sum_{k \in d} s_k$.

CRRA utility implies that the budget shares of each locality are independent of the budget size y . We see that while the socially optimal allocation only depends on local population size, the shares under PR weighs localities by their electoral sensitivity and the shares under MAJ weighs the locality by their relative sensitivity in the district and the contestability of their district.

Proof of Lemma 1

Clearly $A(\mathbf{q}^{PR}) < A(\mathbf{q}^{MAJ})$ iff $y^E(\mathbf{q}^{PR}) > y_\rho^E(\mathbf{q}^{MAJ})$ where

$$y^E(\mathbf{q}) = \begin{cases} \prod_l (q_l / n_l)^{n_l} & \text{if } \rho = 1; \\ \left[\frac{\sum_l n_l (q_l)^{1-\rho}}{\left(\sum_j n_j^{\frac{1}{\rho}} \right)^\rho} \right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \quad (28)$$

Consider first the logarithmic case ($\rho = 1$). Plugging the values \mathbf{q}^{PR} and \mathbf{q}^{MAJ} into y^E/y tells us

that $A(\mathbf{q}^{PR}) < A(\mathbf{q}^{MAJ})$ iff:

$$\begin{aligned} \Pi_l \left(\frac{t_l \phi_l}{\sum_{k=1}^L s_k} \right)^{n_l} &> \Pi_l \left(\frac{t_l \phi_l \gamma_{d(l)} / \mathbf{s}_{d(l)}}{\sum_{k=1}^L s_k \gamma_{d(k)} / \mathbf{s}_{d(k)}} \right)^{n_l} \\ \Pi_l \left(\frac{1}{\sum_{k=1}^L s_k} \right)^{n_l} &> \Pi_l \left(\frac{\gamma_{d(l)} / \mathbf{s}_{d(l)}}{\sum_{k=1}^L s_k \gamma_{d(k)} / \mathbf{s}_{d(k)}} \right)^{n_l}. \end{aligned} \quad (29)$$

Note that the denominator on the RHS of (29) is equivalent to $\sum_d \frac{\gamma_d}{\mathbf{s}_d} \sum_{k \in d} s_k = \sum_d \gamma_d$. Similarly, we can re-write the denominator on the LHS of (29) as $\sum_d \mathbf{s}_d$.

Substituting for these into (29), we get

$$\Pi_l \left(\frac{1}{\sum_d \mathbf{s}_d} \right)^{n_l} > \Pi_l \left(\frac{\gamma_{d(l)} / \mathbf{s}_{d(l)}}{\sum_d \gamma_d} \right)^{n_l}.$$

Taking logarithms, and noting that $\sum n_l = 1$, yields:

$$\begin{aligned} -\log \left[\sum_{d'} s_{d'} \right] &> \sum_l n_l \log \left[\frac{\gamma_{d(l)}}{\mathbf{s}_{d(l)}} \right] - \log \left[\sum_{d'} \gamma_{d'} \right] \text{ or} \\ -\log \left[\sum_{d'} s_{d'} \right] &> \sum_d \mathbf{n}_d \log \gamma_d - \sum_d \mathbf{n}_d \log \mathbf{s}_d - \log \left[\sum_{d'} \gamma_{d'} \right] \text{ or} \\ \sum_d \mathbf{n}_d \log \left[\frac{s_d}{\sum_{d'} s_{d'}} \right] &> \sum_d \mathbf{n}_d \log \left[\frac{\gamma_d}{\sum_{d'} \gamma_{d'}} \right], \end{aligned} \quad (30)$$

where $\mathbf{n}_d = \sum_{l \in d} n_l$.

This proves Lemma 1 for the logarithmic case.

Similarly for $\rho \neq 1$, we substitute the equilibrium values of the allocation under each electoral system into y^E/y and multiply by $\left(\sum_j n_j^\rho \right)^{\rho/(1-\rho)}$. This tells us that $A(\mathbf{q}^{PR}) < A(\mathbf{q}^{MAJ})$ iff:

$$\left[\sum_l n_l \left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^L (s_k)^{1/\rho}} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}} > \left[\sum_l n_l \left(\frac{(s_l \gamma_{d(l)} / \mathbf{s}_{d(l)})^{1/\rho}}{\sum_{k=1}^L (s_k \gamma_{d(k)} / \mathbf{s}_{d(k)})^{1/\rho}} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}} \quad (31)$$

which directly leads to Lemma 1.

Proof of Proposition 1

We normalize $y = 1$ without loss of generality since, with CRRA utility functions, equilibrium budget shares are budget invariant.

Using Lemma 1 makes proving Proposition 1 straightforward.

Consider the case $\rho = 1$. When all districts are well apportioned ($\mathbf{n}_d = 1/D \forall d$), inequality (30) becomes

$$\frac{1}{D} \sum_d \log \left[\frac{s_d}{\sum_{d'} s_{d'}} \right] > \frac{1}{D} \sum_d \log \left[\frac{\gamma_d}{\sum_{d'} \gamma_{d'}} \right]$$

Atkinson (1983) shows the strict concavity of the log implies that this inequality holds if $\frac{\gamma_d}{\sum_{d'} \gamma_{d'}}$

is a mean preserving spread of $\frac{s_d}{\sum_{d'} s_{d'}}$ (and vice versa).

Targeted versus Universal Spending

Equilibrium Allocations

With transfers and a global public good, the objectives in MAJ and PR are similar to (6) for PR and (11) for MAJ adjusting for the utility $\Delta w_l(\mathbf{q}, G) =$, The budget constraint becomes

$$\sum_l n_l q_l + G = y.$$

In PR, the first order conditions are thus:

$$\begin{cases} \sum_l s_l u'(G) = T\lambda^{PR}, \\ \frac{s_l}{n_l} = T\lambda^{PR} \quad \forall l \text{ with } q_l > 0 \end{cases} \quad (32)$$

where λ^{PR} is the Lagrange multiplier of the budget constraint under PR.

In a symmetric equilibrium, only individuals in the locality with the highest s_l/n_l receives a transfer. The following inequality holds:

$$u'(G) \leq \max_l \frac{s_l}{n_l} \frac{1}{\sum_{k=1}^L s_k}, \quad (33)$$

with equality if some positive transfer is given.

In MAJ, the first order conditions become:

$$\begin{cases} \sum_{d \in D} \gamma_d u'(G) = D\lambda^{MAJ}, \\ \frac{\gamma_{d(l)}}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} = D\lambda^{MAJ} \quad \forall l \text{ with } q_l > 0 \end{cases} \quad (34)$$

where λ^{MAJ} is the Lagrange multiplier of the budget constraint under MAJ.

In a symmetric equilibrium, only individuals in the locality with the highest $\frac{\gamma_{d(l)}}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k}$ receive a transfer. It follows from (34) that

$$\max_l \frac{\gamma_{d(l)}}{\sum_{d \in D} \gamma_d} \frac{1}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} \geq u'(G), \quad (35)$$

with equality if some transfers arise in equilibrium.

Proof of Proposition 2

If $L = D$ and $dn_l = \frac{1}{D} \forall l$ then (33) and (35) tell us that

$$\begin{cases} u'(G^{PR}) \leq \max_l \frac{s_l}{\sum_{k=1}^L s_k} D, \\ u'(G^{MAJ}) \leq \max_l \frac{\gamma_l}{\sum_k \gamma_k} D \end{cases} \quad (36)$$

where the respective equalities hold whenever $G < y$.

It follows directly that if $\max_l \frac{s_l}{\sum_k s_k} > (<) \max_l \frac{\gamma_l}{\sum_k \gamma_k}$ then $G^{PR} \leq (geq) G^{MAJ}$.

Parties' Objective

Let

$$\mu(\mathbf{q}) := \sum_d p_d(\mathbf{q}) = \frac{D}{2} + \sum_d \gamma_d \times \left[\frac{\sum_{l \in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]$$

be A 's expected seat share, and define:

$$\sigma_E^2(\mathbf{q}) := \sum_d p_d(\mathbf{q}) [1 - p_d(\mathbf{q})].$$

Since, the individual $p_d(\mathbf{q})$ are statistically independent from one another, the CLT of Liapounov tells us that:

$$\frac{\sum_d \mathbb{1}_d - \mu(\mathbf{q})}{\sigma_E(\mathbf{q})},$$

is asymptotically distributed as a standard normal.

The probability that A wins a majority of the seats given policy platforms \mathbf{q} is therefore:

$$\pi_A(\mathbf{q}) = \Pr \left(\frac{\sum_d \mathbb{1}_d - \mu(\mathbf{q})}{\sigma_E(\mathbf{q})} \geq \frac{D/2 - \mu(\mathbf{q})}{\sigma_E(\mathbf{q})} \right)$$

Using the asymptotic distribution in this, the probability that A wins is:

$$\pi_A(\mathbf{q}) \approx 1 - \Phi[S(\mathbf{q})],$$

where $S(\mathbf{q}) = \frac{D/2 - \mu(\mathbf{q})}{\sigma_E(\mathbf{q})}$ and $\Phi[\cdot]$ is the standard normal cumulative density function.

Note that:

$$\sigma_E^2(\mathbf{q}) = \frac{D}{4} - \sum_{d \in C} \gamma_d^2 \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]^2 \quad (37)$$

which implies:

$$S(\mathbf{q}) = \frac{-\sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]}{\left(\frac{D}{4} - \sum_d \gamma_d^2 \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]^2 \right)^{1/2}} \quad (38)$$

Assuming that parties maximize their approximate probability of winning, the problem of party A becomes:

$$\begin{aligned} \max_{\mathbf{q}^A} & 1 - \Phi \left[\frac{-\sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]}{\left(\frac{D}{4} - \sum_d \gamma_d^2 \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]^2 \right)^{1/2}} \right] \\ \text{s.t.} & \sum_l n_l^\alpha q_l = y, \end{aligned}$$

which leads to the first order conditions:

$$\begin{aligned} n_l^\alpha \lambda^A &= -\phi(S(\mathbf{q})) S(\mathbf{q}) \times \left[\frac{-\frac{\partial \mu(\mathbf{q})}{\partial q_l}}{\frac{D}{2} - \mu(\mathbf{q})} - \frac{\frac{\partial \sigma_E^2(\mathbf{q})}{\partial q_l}}{\sigma_E^2(\mathbf{q})} \right] \\ &= -\phi(S(\mathbf{q})) S(\mathbf{q}) \times \left[\frac{\gamma_{d(l)} \frac{s_l}{\sum_{j \in d(l)} s_j} u'(q_l)}{\sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d \right]} - \frac{\gamma_{d(l)}^2 \frac{s_l}{\sum_{j \in d(l)} s_j} u'(q_l) \times \left[p_{d(l)}(\mathbf{q}) - \frac{1}{2} \right]}{\sigma_E^2(\mathbf{q})} \right] \end{aligned}$$

As explained by Strömberg (2008), the first term captures the incentive of the candidate to influence the expected number of electoral votes won, the mean of the distribution, while the second term arises from the incentive to influence the variance in the number of electoral votes.

It is easy to show that in equilibrium, $\mathbf{q}_A = \mathbf{q}_B$, which allows us to simplify the FOC into:

$$\frac{\lambda^A \times \sum_d \gamma_d \beta_d}{\phi(S(\mathbf{q})) S(\mathbf{q})} = \gamma_{d(l)} \frac{s_l n_l^{-\alpha}}{\sum_{j \in d(l)} s_j} u'(q_l) \left[1 + \frac{\gamma_{d(l)} \beta_{d(l)} \times \sum_d \gamma_d \beta_d}{\sigma_E^2(\mathbf{q})} \right]$$

where the left-hand side of the equation is independent of l . We can thus label it as λ' , which leads to (19).

Electoral College vs. National Popular Vote

Proof of Proposition 3

The proof is straightforward. Let us consider what happens when we divide districts into finer and finer partitions of localities with similar characteristics (so that $\frac{s_{d(l)}}{n_{d(l)}} = \frac{s_l}{n_l}$, $\forall l$).²³ This ensures that the FOC under the Electoral College becomes:

$$\frac{\partial u_l(\mathbf{q}^{\text{College}})}{\partial q_l^A} \propto \frac{n_{d(l)}/\gamma_{d(l)}}{\omega_{d(l)}} n_l^{\alpha-1}.$$

Then, reapportioning the electoral college so that $\omega_{d(l)} = n_{d(l)}/\gamma_{d(l)}$ ensures that the resulting allocation of public goods converges towards the social welfare optimum. In the limit, *i.e.* redistricting up to the point where $D \rightarrow L$, ensures that the first condition can be attained.

Proportional Allocation of Electors

We want to show that organizing PR elections in each district under the Electoral College can deliver an outcome that is welfare superior to that under the Popular Vote. In particular, apportioning district evenly by setting ω_d equal to the population share of the district produces strictly higher welfare than the Popular Vote if (i) turnout and information rates differ across districts ($E(t_l \chi_l | d) \neq E(t_l \chi_l | d')$); (ii) they are positively correlated ($\rho(t_l, \chi_l) > 0$) and (iii) information and turnout are orthogonal to swingness ($\rho(t_l, \phi_l) = 0$).

²³To hold everything else constant, the distribution of district-specific shocks δ_d is the same as in the original, larger, district. At the limit $D = L$, however, the distribution of these shocks becomes irrelevant, since they can be corrected by reweighting the Electoral College.

To prove this, note first that, as shown in equation (9) in the main text, the vote share of A in district d is:

$$\pi_d(\mathbf{q}; \delta_d) = \frac{1}{2} + \sum_{l \in d} \frac{t_l n_l}{m_d} (\chi_l \phi_l [\Delta u_l(\mathbf{q}) - \delta_d] + (1 - \chi_l) \zeta_l),$$

where $m_d := \sum_{k \in d} t_k n_k$ is the expected number of votes in district d .

When there are D local elections under PR and each district receives a fraction ω_d of the Electors, the seat share of party A in the Electoral College becomes: $\sum_d \omega_d \pi_d(\mathbf{q}; \delta_d)$. Taking expectations, the objective function of party A is then:

$$\pi_{PR}^{\text{College}}(\mathbf{q}) = \frac{1}{2} + \zeta_{PR}^{\text{College}} + \sum_d \frac{\omega_d}{m_d} \sum_{l \in d} s_l (\Delta u_l(\mathbf{q}) - \beta_{d(l)}),$$

where $\zeta_{PR}^{\text{College}} := \sum_k \frac{\omega_{d(k)} t_k n_k}{m_{d(k)}} (1 - \chi_k) \zeta_k$.

To compare, maximizing the probability of winning under the Popular Vote would produce the same allocation of public goods if $\omega_d = m_d$ in the electoral college under PR.

If instead $\omega_d = \sum_{k \in d} n_k$, *i.e.* if the electoral college were well apportioned, then lower turnout and information rates in a given district would not translate into a lower provision of public goods in that district. This brings the allocation closer to the social optimum.

Appendix 2: Empirical

As discussed in the introduction, the empirical literature on the economic effect of constitution suffer from a certain arbitrariness in deciding that some government expenditure categories are more easily geographically targetable than others. To avoid this issue, a possibility is to remove the decision of whether a government policy is targetable or broad from the hand of the researchers. The variable *encompassing* from the VDem dataset²⁴ allows us to do just that. It aggregates national expert assessments about whether national government policy is more “particularistic” (score of 1 or 2) or more “encompassing” (score of 3 or 4). Knowing that such an assessment is bound to be influenced by the comparison with neighboring countries, we compare the regression results with and without regional fixed effects in the cross-sectional analysis.

Using cross-country data for 68 democratic countries from 1992 to 2013 (we keep only country-years with *polity2* > 0), we estimate the following:

$$E_{it} = \beta_0 + \beta_1 Pr_{it} + \beta_2 X_{it} + \delta_t + \beta_4 R_i + \epsilon_{it} \quad (39)$$

²⁴Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, with David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Kelly McMann, Pamela Paxton, Daniel Pemstein, Jeffrey Staton, Brigitte Zimmerman, Rachel Sigman, Frida Andersson, Valeriya Mechkova, and Farhad Miri. 2015. “V-Dem Codebook v5.” Varieties of Democracy (VDem) Project.

where E_{it} is our measure of encompassiveness ; Pr_i is one of our measure of proportional representation (PR or District Magnitude), X_{it} includes a set of country controls (see the list of variables below) including whether the country is considered presidential²⁵, δ_t are year fixed effects and R_i are either regional dummies for OECD, AFRICA, ASIAE and LAAM or country fixed effects. Standard errors are clustered at the country level.

The results are in Table 1. The first three columns use PR as a measure of proportional representation while the last three use District magnitude. Column (1) and (5) use Persson and Tabellini (1999)'s list of controls without regional dummies. Column (2) and (6) use Persson and Tabellini (1999)'s list of controls with regional dummies. Finally, column (5) and (7) use Blume et al. (2009)'s full list of controls with regional fixed effects. Column (4) and (8) use Blume et al. (2009)'s full list of controls with country fixed effects.²⁶

Overall, we do not find that proportional representation correlate with more encompassing policies. Only with the most parsimonious controls and no regional fixed effect do we find a weakly significant correlation between district magnitude and encompassing policies. This correlation has the opposite sign to the usual prediction in the literature that PR systems are more conducive to encompassing policies and not robust to the addition of further controls. With country fixed effects, we find a negative correlation between PR and encompassiveness in the few countries that have changed electoral system.

Variables

- *PR* which take value 1 if, according to the *HOUSESYS* variable from Keefer (2012)'s "Database of Political Institutions", a minority of the Lower House (parliament or congress) seats are elected in SMD.
- *District Magnitude* is the the ratio of Mean District Magnitude (MDMH) and Total Seats. MDMH is the weighted average of the number of representatives in the lower house elected by each constituency size (if available). This variable has been derived from the Database of Political Institutions, 2012 version (DPI2012). Total Seats is the number of total seats in the legislature, or in the case of bicameral legislatures, the total seats in the lower house.²⁷

Following Persson and Tabellini (1999), we include the following controls:

²⁵The interactions between presidential and the electoral system being insignificant we omitted them in the results presented here.

²⁶The variation in district magnitude within countries over time represents a large share of the overall variation. More surprisingly maybe is that a number of countries in our sample change electoral system over the period: Armenia, Bolivia, Cambodia, Ecuador, El Salvador, Guinea-Bissau, Iraq, Italy, Kyrgyzstan, Liberia, Macedonia, Mongolia, Nepal, Sierra Leone and Ukraine.

²⁷Using MDMH instead of MDMH/Total Seats in the regressions above does not affect the results.

Table 1: Proportional Representation and Encompassing 1992-2013

Equation No.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PR	-0.122 (-0.78)	-0.001 (-0.01)	0.015 (0.11)	-0.362** (-2.13)				
Presidential	-0.025 (-0.13)	0.075 (0.46)	0.044 (0.33)	0.127* (1.87)	-0.057 (-0.29)	0.073 (0.46)	0.057 (0.45)	0.127* (1.86)
District Magn					-0.396* (-1.92)	-0.198 (-0.73)	-0.095 (-0.48)	0.013 (0.12)
Regional FE	No	Yes	Yes	No	No	Yes	Yes	No
Country FE	No	No	No	Yes	No	No	No	Yes
R-squared	0.309	0.501	0.638	0.116	0.317	0.505	0.640	0.076
N	533	533	393	393	535	535	395	395

Notes: Year dummies included. Standard errors clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- GDP: log of per capita income; is obtained from the World Economic Output of the IMF. In particular, we have taken the Purchasing Power Parity (PPP) per capita GDP (divided by 1000). To avoid simultaneity issues, we have taken the lagged value (1 year lagged).
- TRADE: log of openness, measured as the log of the sum of exports plus imports in gdp;
- PROP65: the share of the population above 65;
- FEDERAL: centralization of government spending (measured as expenditures of central government divided by expenditures of general government).

In addition, following Blume et al. (2009) we also include the following controls: ²⁸

- POLITY2, a measure of democracy,
- NYYSDEM: number of democratic years since 45 from Acemoglu,
- PROP1564: the share of the population between 15 and 64,
- Variable capturing a history of colonization by the UK, Spain and other colonial powers.

²⁸Blume et al. (2009) use the Gastil measure instead of Polity2 and the age of a democracy instead of NYYEAR. The results are not sensitive to the measure of democracy used.