Electoral Systems and Inequalities in Government Interventions

Garance Genicot* ⓡ Laurent Bouton* ⓡ Micael Castanheira§

October 16, 2018

Abstract

This paper studies the political determinants of inequality in government interventions under the majoritarian and proportional representation systems. A model of electoral competition with targetable government intervention and heterogeneous localities allows us to 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 under proportional representation systems. This contrasts with the conventional wisdom that government interventions are more unequal in majoritarian systems.

JEL Classification Numbers: D72, H00

Key Words: Distributive Politics, Electoral Systems, Public Good, Inequality.

Acknowledgements: We benefited from comments by Patrick François, Leyla Karakas, Alessandro Lizzeri, Dilip Mookherjee, Massimo Morelli, Nicola Persico, Debraj Ray as well as seminar and conference participants at Princeton, Georgetown, MSU, Texas A&M, UBC, Bocconi, Collegio Carlo Alberto, EUI, TSE, ETH Zurich, Copenhagen, CERGE-EI, Cornell PE Conference 2018, ThReD Conference 2018, Quebec PE Conference 2018, POLECON UK 2nd Annual Workshop, CORE Belgian-Japanese Public Finance Workshop and 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: boutonllj@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, Baqir, and Easterly (1999), Banerjee, Iyer, and Somanathan (2008), World Bank (2004)) and in terms of 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, Gerber, and Snyder (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 substantial: for instance, using Brazilian data, Finan and Mazzocco (2016) estimate that 25 percent of the public funds allocated by legislators are distorted relative to the socially optimal allocation.

This paper studies the political determinants of inequalities in government interventions under *Majoritarian* (MAJ) and *Proportional Representation* (PR) systems.[1] These systems are ubiquitous: 82 percent of legislative elections held in the 2000s employed either MAJ or PR (Bormann and Golder (2013)).

The conventional wisdom is that MAJ systems are more conducive to inequality, because they provide steeper incentives to target government interventions onto specific groups (Persson and Tabellini (1999, 2000); Persson (2002); Lizzeri and Persico (2001); Milesi-Ferretti, Perotti, and Rostagno (2002)). Multiple arguments

[1] In MAJ systems, there is a multitude of electoral districts that each select a limited number of representatives using some winner-take-all method. The epitome of those systems is the version with single-member districts and first-past-the-post. 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.

underlie this view (see Section 2 for details). A powerful one is Lizzeri and Persico (2001)'s: in MAJ, 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 are 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)).

However, this view overlooks another important difference: the geographic distribution of voters matters differently in the two systems. In MAJ systems, parties ought to 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 helps winning additional seats in the national assembly.

Taking account of the geographical distribution of voters uncovers a *relative electoral sensitivity effect* that is only present in MAJ systems. We identify conditions under which this effect induces parties to "sprinkle" resources across districts, resulting in lower inequalities in government intervention in MAJ than in PR systems. These findings are in line with the mixed empirical evidence in the literature that compares targeting under those systems (see Section 2).

In our model of electoral competition, two parties compete by targeting governmental resources (cash transfers, goods or services) to heterogeneous localities, potentially smaller than electoral districts. Heterogeneity is multidimensional: localities may differ in population size, turnout rate, and swingness.[2] Together, these dimensions determine the *electoral sensitivity* of a locality. The objective of parties

[2]In 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.

is identical across systems, implying that the difference between MAJ and PR only stems from how votes translate into seats.[3] In PR, the number of seats won by a party is proportional to its nationwide vote share. In MAJ, there are several single-member districts. To win a district's seat, a party needs to obtain at least 50% of the votes from the localities that compose that district.

We first characterize the allocation of government resources under each system. We show that PR creates an incentive to allocate more resources to localities that are electorally more sensitive (*e.g.* with higher turnout), independently of their location. In MAJ instead, parties must accumulate victories in different districts to increase their seat share. This has several implications. First, parties have incentives to discriminate against localities in non (or little) contestable districts. This mechanism is the one driving the standard result in the existing literature. Second, for a given level of district contestability, parties have incentives to allocate more resources to localities with higher *relative* electorally sensitivity. The comparison is with the other localities *in the same district*. We find that, as a consequence, parties may "sprinkle" governmental resources over more localities than in PR.

How do these differences determine the eventual level of inequality in government interventions? We consider two different perspectives. First, we consider the implications in terms of *horizontal inequality*: how government intervention differs across localities with *identical* characteristics. We find that the relative electoral sensitivity effect reinforces horizontal inequality under MAJ. This feature is in line with the existing literature. Second, we focus on *vertical inequality*: how government intervention differs across localities with *different* characteristics. Here, we find that the standard result in the literature can actually be reversed: MAJ may end up producing less vertical inequality than PR. This typically happens when contestability effects are weak, and localities are more similar within a district than between districts.

[3]The main model focuses on the case in which parties maximize their share of seats in the national assembly. In the extensions, we consider parties that maximize their probability of winning at least fifty percent of the seats.

However, all inequalities need not be socially undesirable. The Samuelson condition requires for instance that more populated localities receive higher levels of public goods. To determine which system produces the least politically distorted allocations, we develop a welfare-based measure of inequality in government interventions *à la Atkinson*. Using that measure, we identify conditions under which either PR or MAJ produces more socially undesirable inequalities. These conditions depend on the heterogeneity of within- and between-district levels of electoral sensitivities, and in district contestabilities.

Section 6 explores various extensions of the model. The first one studies the effect of the electoral system on the *composition* of government spending (pure targeted transfers *vs.* global public good). This allows us to revisit Lizzeri and Persico (2001) and show that the relative electoral sensitivity effect can overturn its conclusion. The second extension shows that our results are robust to a modification in the parties' objective. Parties can either maximize their share of seats in the national assembly, or their probability of winning a majority of seats. The third extension studies PR systems in which each electoral district carries a pre-determined seat share. Except for the influence of turnout, our results for PR with a nationwide district extend to those alternative systems. The last extension shows how other dimensions of heterogeneity among voters would enter the electoral sensitivity of localities.

Finally, Section 7 concludes with a discussion of the implications of the relative electoral sensitivity effect for the empirical literature on distributive politics.

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, it is more closely related to study of the effects of electoral systems.

As already mentioned, a recurrent theme in that literature is that parties want to target a smaller fraction of the population in MAJ than in PR systems. There are various mechanisms that produce that outcome. We already mentioned the fifty-of-fifty percent mechanism (Lizzeri and Persico (2001)), that 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 changes 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 to change the identity of the party holding a majority of seats. Parties have incentives 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). In PR systems, legislators have a national constituency. Their incentives are therefore aligned with those of party leaders, which are to 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. This mechanism, which is related to Rogowski (1987), is absent from our model.

Rogowski and Kayser (2002) points to the seats-votes elasticity as a key factor influencing the targeting of government interventions. When elasticity is higher, incentives to target groups that can deliver many votes are higher. Given that MAJ systems have a higher seats-votes elasticity than PR systems (Taagepera and Shugart (1989)), there should be more targeting towards electorally responsive groups under

MAJ systems. We could view our results as refining 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.[4] The first one relies on government accounts and has to make assumptions about which items can reasonably be thought to represent broad public goods 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, Perotti, and Rostagno (2002); Blume et al. (2009); Funk and Gathmann (2013)). One exception is Aidt, Dutta, and Loukoianova (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, Kayser, and Rogowski (2008); Betz (2017)). The difference seems to originate in the type of trade barriers considered: non-tariff barriers tend to be used more in PR systems, while tariffs tend to be used more 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

[4]At least two studies 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, Nannicini, and Naticchioni (2011) that use Italian data, and (*ii*) Stratmann and Baur (2002) that use German data.

targeted transfers. Moreover, these classifications happen to vary across studies and a given classification of government expenditures by types is unlikely to fit all countries.[5] 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 is then to identify when governments exploit their margin of action to target them in practice or not.[6]

To understand how these issues affect empirical findings, we have revisited Persson and Tabellini (1999) and Blume et al. (2009) using a new measure of how *encompassing* as opposed to targeted governmental spending are. 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) or Myerson (1993), an increase in the number of parties should be associated with the targeting of government interventions towards a narrower subset of the electorate. The incentives 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

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

[6]Large-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, Nannicini, and Naticchioni (2011); Min (2015); Strömberg (2008), and Golden and Min (2013) for a survey).

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, Nicodème, and Profeta (2012), which focuses on tax reforms in EU countries.

Last but not least, Stashko (2018) provides evidence of the relevance of the relative electoral sensitivity effect. After generalizing this effect to a setup in which localities ("counties" in her model) can span multiple districts, she tests this effect using data on U.S. state governments and legislative elections. She finds that the effect is statistically and economically significant: the amount received by a county depends both on the electoral sensitivity of that county and the electoral sensitivity of other counties in the same districts.

3 The Model

In this section, we lay out our model of electoral competition. It is a standard probabilistic voting model, in the tradition of Enelow and Hinich (1982); Lindbeck and Weibull (1987); Dixit and Londregan (1996); Grossman and Helpman (1996); Persson (2002); Strömberg (2004)). Our model has two key features: (i) we allow for targeting of government interventions at the sub-district level, and (ii) we consider multidimensional voter heterogeneity.

3.1 The Economy

Consider a country with a continuum of individuals of total mass 1. The population is partitioned into *localities* $l \in \{1, 2....L\}$ of size n_l, *s.t.* $\sum_l n_l = 1$. Each locality belongs to an *electoral district* $d \in \{1, 2....D\}$.[7]

[7]We use the word "locality" to describe any sub-electoral district population group that can be targeted within a district. We picked that term because governments often use geography to target their interventions. But, when feasible, they may also target additional sub-groups, using

An elected government has to allocate a budget y to 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 $\mathbf{q} = \{q_1, ..., q_L\}$. This implies that governmental resources can be targeted at a finer level than the electoral district (except for the special case $L = D$, when there is exactly one locality per district).

We cover a variety of government interventions that range from pure local public goods to pure transfers. The central difference between the different 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. With pure transfers instead, 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:

$$\sum_l n_l^\alpha q_l \leq y, \tag{1}$$

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 spillovers across localities.

for instance gender, ethnic origin, or age criteria. Stashko (2018) considers the case of localities split across districts.

3.2 Normative Benchmark

Before introducing the structure of electoral competition, we establish the politics-free benchmark. To maximize utility in the population, a social planner would:

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

The socially optimal allocation must satisfy the standard Samuelsonian conditions:

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

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, *i.e.* $\forall \alpha < 1$, 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.* Section 6.2 further discusses this objective function and covers the polar case in which they maximize their probability of winning.

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.

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.[8] 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, in the distribution of voter preferences, and may belong to different electoral districts. In Section 6.4, we show that the model could include other dimensions of heterogeneity such as information, intensity of preferences and partisanship.

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 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 $t_l n_l$.[9]

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

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

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...) and/or political

[8] As explained in Golden and Min (2013, p.84): "Studies overwhelmingly find that incumbent politicians are rewarded by voters for distributive allocations, and in particular for those that are clientelistic and from which recipients can be excluded."

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

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:

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

Hence, from the parties' standpoint, each individual in a locality l has political preferences that are the results of two random shocks.[10] 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 (resp. against) B when positive (resp. when 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.

[10] For our purpose, adding locality-specific biases and/or a national shock would only complicate the notation without adding insights.

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 computed as:

$$\pi_l(\mathbf{q}; \delta_d) = \tfrac{1}{2} + \phi_l \left(\Delta u_l(\mathbf{q}) - \delta_d\right), \tag{5}$$

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):

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

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

$$s_l := t_l n_l \phi_l$$

is the *electoral sensitivity* of locality l.

The first order conditions are thus:

$$\frac{\partial u_l\left(\mathbf{q}^A\right)}{\partial q_l^A} = \frac{T n_l^\alpha}{s_l} \lambda^{PR}, \ \forall l, \tag{7}$$

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 there exists a unique equilibrium in Appendix 1.

It follows that localities with higher electoral sensitivity benefit from more government interventions. That is, they combine a large population, with a high turnout, and with an ideologically more 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:

$$\begin{aligned} \pi_d\left(\mathbf{q}; \delta_d\right) &= \sum_{l \in d} \frac{t_l n_l}{\sum_{k \in d} t_k n_k} \pi_l\left(\mathbf{q}; \delta_d\right) \\ &= \tfrac{1}{2} + \sum_{l \in d} \frac{s_l}{\sum_{k \in d} t_k n_k} \left(\Delta u_l\left(\mathbf{q}\right) - \delta_d\right), \end{aligned} \tag{8}$$

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 (8), this becomes:

$$p_d(\mathbf{q}) = \Pr\left(\delta_d \leq \sum_{l \in d} \frac{s_l}{\sum_{j \in d} s_j} \Delta u_l(\mathbf{q})\right). \tag{9}$$

To avoid corner solutions and ensure that payoffs are everywhere differentiable, throughout the paper, we assume that this probability is non-degenerate, for any allocation. In other words, we assume that all districts are *contestable*:[11]

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

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

$$\max_{\mathbf{q}^A | \sum_l n_l^\alpha q_l = y} \pi_{\text{MAJ}}(\mathbf{q}) = \frac{1}{2} + \frac{1}{D} \sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l}{\sum_{j \in d} s_j} \Delta u_l(\mathbf{q}) - \beta_d\right]. \tag{10}$$

The first order conditions are thus:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\sum_{j \in d(l)} s_j}{s_l} \frac{n_l^\alpha}{\gamma_{d(l)}} \lambda^{\text{MAJ}} \; \forall l, \tag{11}$$

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 $\mathbf{q}^A = \mathbf{q}^B$.

[11]Following 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 key difference with PR is that the localities receiving a larger share of the government budget are now the ones with a higher *relative* electorally sensitivity: 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, underlies the comparison:

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

Theorem 1 identifies key differences between the two systems. First, localities that are electorally more sensitive are systematically better treated in PR. In MAJ instead, the electoral sensitivity of each locality is only assessed in comparison with 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 they are surrounded by *other* localities with different characteristics. Locality 2 is the most sensitive in district 1. In district 2 instead, 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 in MAJ than in PR. While this corroborates traditional results (see *e.g.* Persson and Tabellini (2000), Strömberg (2008)), the mechanism is different (relative sensititivites, 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 ($\gamma_d = \gamma, \ \forall d$). Hence, the differences between 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 regroup 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 this: 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 insulates some localities from one another. Locality 1 only competes with locality 2, and locality 3 only against locality 4. While this costs some resources to localities 3 and 4, it substantially benefits locality 1. Under MAJ, locality 1 receives only about 6 times less than to locality 4. The Gini coefficient of inequality is actually lower under MAJ than under PR in this example.

To go beyond the example, consider the case in which each locality would form a district by itself. There is then L different 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, instead, 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

A conclusion that emerges from the previous section is that either system may create the largest level of inequality. However, looking at inequality does not tell the full story: 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:

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

With 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 $\mathcal{W}(\mathbf{q})$:

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

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

$$A(\mathbf{q}) \equiv 1 - \frac{y^E(\mathbf{q})}{y} = \begin{cases} 1 - \frac{1}{y}\Pi_l \left(q_l/n_l\right)^{n_l} & \text{if } \rho = 1 \\ 1 - \left[\frac{\sum_l n_l (q_l/y)^{1-\rho}}{\left(\sum_j n_j^{\frac{1}{\rho}}\right)^{\rho}}\right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \tag{12}$$

This is a measure of 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. $A(\mathbf{q})$ 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 $A\left(\mathbf{q}^{PR}\right) < A\left(\mathbf{q}^{MAJ}\right)$ and vice versa. We show in Appendix 1 that:[12]

Lemma 1 PR *Atkinson-dominates* MAJ *iff:*

$$\frac{\sum_l n_l \left(s_l\right)^{\frac{1-\rho}{\rho}}}{\left(\sum_k \left(s_k\right)^{\frac{1}{\rho}}\right)^{1-\rho}} \lessgtr \frac{\sum_l n_l \left(\frac{\gamma_{d(l)} s_l}{\sum_{k\in d(l)} s_k}\right)^{\frac{1-\rho}{\rho}}}{\left(\sum_k \left(\frac{\gamma_{d(k)} s_k}{\sum_{j\in d(k)} s_j}\right)^{\frac{1}{\rho}}\right)^{1-\rho}} \text{ for } \rho \gtrless 1, \tag{13}$$

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-

[12]In Appendix 1, we also solve for the case $\rho = 1$.

hand side is the score of PR, which only depends on the absolute sensitivity of each locality. On the right-hand side, that of MAJ depends on district contestability and on relative sensitivities within each district.

To gain further intuition, 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 generically Atkinson-dominates MAJ . Second, consider the polar 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 and are treated equally. MAJ induces maximal sprinkling of government resources.

Moving to a more general comparison, when we shut down the population size effect, we find that:

Proposition 1 PR *Atkinson-dominates* MAJ *if* $\gamma_{d(l)} / \sum_d (\gamma_d)^{1/\rho}$ *is a mean preserving-spread of* $s_l / \sum_{k=1}^{L} (s_k)^{1/\rho}$ *(and conversely) when either:*
(1) $\rho = 1$*, and* $\sum_{l \in d} n_l = 1/D\ \forall d$ *, or*
(2) $\rho \neq 1$*, there is one locality per district, and* $n_l = 1/D\ \forall d$*.*

The intuition behind this result is rather straightforward: we have seen that the Atkinson measure of inequality increases when the equilibrium allocation further differs from the social optimum. Such deviations can either be due to deviations in electoral sensitivities caused by heterogeneous turnout or political preferences, or to unequal district contestabilities. Proposition 1 shows that this intuition is formally correct when districts are well-apportioned (for $\rho = 1$), and capture all the targetability of government resources (for $\rho \neq 1$). Well-apportioned districts imply that government interventions should be similar across districts, in which case any further heterogeneity in s_l or in γ_d becomes socially harmful.

6 Extensions

This section explores different extensions of our model. The first extension allows for two instruments of government intervention: pure targeted transfers and a global public good. The second one modifies the objective of parties. The third one studies a variant of PR systems. The last one considers other dimensions of heterogeneity among voters.

6.1 Targeted versus Universal Spending

The model in this section includes two instruments that the government can use: targeted transfers and a global public good. Our purpose here is to highlight the role of the sprinkling effect and of 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 at y so that the budget constraint becomes $\sum_l n_l q_l + G = y$.

As shown in Appendix 1, only one locality may receive transfers. In the unique equilibrium under PR, this is the locality with the highest s_l/n_l. If some transfers are given, then:

$$u'(G) = \max_{l} \frac{s_l}{n_l} \frac{1}{\sum_{k=1}^{L} s_k}. \tag{14}$$

Under MAJ, the equivalent FOC characterizing the equilibrium is:

$$\max_{l} \frac{\gamma_{d(l)}}{\sum_{d \in D} \gamma_d} \frac{1}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} = u'(G). \tag{15}$$

Comparing (14) and (15) identifies whether PR or MAJ leads to the largest provision of the global public good. First, 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. Suppose that all localities are identical in s_l and n_l and that all electoral districts have the same number of localities. Then, there is no transfer under PR. Heterogeneous district contestabilities instead make transfers more attractive in MAJ.

Second, there is the electoral sensitivity effect. To isolate its influence, let us switch off the former channel by assuming that contestability is the same across districts ($\gamma_d = \gamma$ for all d) and focus on a situation of maximal incentive to sprinkle resources across the country. As seen in Section 4.4, this would for instance happen when there is one locality per district (such that all relative electoral sensitivities are equal to one) and these localities all have the same population size.

For this situation, we find a simple sufficient condition under which $G^{PR} \leq G^{MAJ}$ in equilibrium – with a strict inequality when $G^{MAJ} > 0$. This condition is that the largest level of electoral sensitivity ($max_l \ s_l$) is larger than the average. That is, as soon as there is *some* heterogeneity in electoral sensitivities, the standard result in the literature, *i.e.* that the level of broad public good should be higher in PR than in MAJ, gets reversed.

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 instead maximizing their probability of winning a majority of seats.

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, a party cannot easily pass all the legislation it wants as soon as it has a one-seat majority in the legislative assembly (the current situation in the US Senate is a case in point). This 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)).

One may still be concerned with this assumption. To address this potential concern, 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 respectively PR and MAJ, the parties' objective functions (6) and (10) become:

$$\text{In PR:} \quad \max_{\mathbf{q}} \ \tfrac{1}{2} + \Pr\left[\sum_l \frac{s_l}{T}\left(\Delta u_l(\mathbf{q}) - \beta_{d(l)}\right) \geq 0\right], \tag{16}$$

$$\text{In MAJ:} \quad \max_{\mathbf{q}} \ \Pr\left[\sum_d \mathbb{1}_d \geq \frac{D}{2}\right], \tag{17}$$

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

The objective function (16) 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 what matters is no longer to achieve a majority in each district separately. Winning a given district only matters insofar as it helps reaching the threshold of 50% of all the 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 for the case of a large enough number of districts. Defining:

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

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:

$$\lambda' = \gamma_{d(l)} \frac{s_l n_l^{-\alpha}}{\sum_{j \in d(l)} s_j} u'(q_l) \left[1 + \frac{\sum_d \gamma_d \beta_d}{\sigma_E^2(\mathbf{q})} \gamma_{d(l)} \beta_{d(l)}\right], \tag{18}$$

which directly compares to (11), 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 the discussion. Then, the localities benefiting from more government interventions than in the baseline case are those belonging to districts that are more contestable *and* also biased towards 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 Other Proportional Representation Systems

In line with most of the literature, so 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:

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

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 $n_d := \sum_{l \in d} n_l$, obtains that $m_d = t_d n_d$. Taking first order conditions and letting λ^{DPR} denote the multiplier on the budget constraint, we have:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \left[\frac{\mu_d}{n_d}\,\frac{s_l}{t_d}\right]^{-1} n_l^{\alpha}\,\lambda^{DPR}\ \forall l, \tag{19}$$

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 the 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: $n_d \cdot t_d$. For this reason, district borders become immaterial to the parties' platforms. In the district-specific version of PR instead, 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 of Section 4. Like in MAJ, a high-turnout locality will receive less if it is located in a higher, rather than a 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.[13]

6.4 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 is for the sake of expositional clarity. We could easily consider other sources of heterogeneity such as information, partisanship, and preference intensity.

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 of voters would then influence the allocation of governmental resources in the same way as swingness under both PR and MAJ.

Preference Intensity. Again, following Strömberg (2004), we could assume that

[13] The same holds for other possible dimensions of heterogeneity such as information and partisanship discussed in Section 6.4.

some voters benefit more from government interventions than others. Denoting by η_l the preference intensity of voters in locality l, we would then have that the electoral sensitivity of locality l, s_l, would include the parameter η_l. The same conclusion as for information follows.

Partisanship and Core Voters. While our baseline model focuses on the "swing voter" theory of elections, Dixit and Londregan (1996) shows how it could be extended to study the impact of "core voters". This alternate approach captures the fact that different population groups can have different party affinities (*e.g.*, they are from the same ethnic or cultural group). When each party is more effective in delivering favors to its core supporters, either the parties' costs of providing public goods, or the voters' valuation of each party promises, become party-locality specific. As a result, each locality's electoral sensitivity parameter s_l must be replaced by party-specific values s_l^P. The upshot, as shown by Dixit and Londregan (1996), is that each party favors their core voters. The key difference with the swing voter approach is that the two parties make different electoral promises, *i.e.* $\mathbf{q}^A \neq \mathbf{q}^B$.

Yet, the effect of partisanship would differ across systems. Indeed, in MAJ, it is the relative partisanship of localities within districts that would be key. Consider for instance a locality l with a large fraction of, say, A-core voters. That locality would receive more from party A (and less from party B) if it belongs to a district with many B-core voters than if it belongs to a district with many A-core voters.

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. This effect can induce parties to "sprinkle" resources across districts, and hence reduce inequality. We found that this effect can be strong enough that inequalities in government intervention end up being lower in MAJ than PR systems. This result runs against a recurrent theme in the literature, which argues that parties target a smaller fraction of the population under MAJ systems. We also explored the implications of the relative electoral sensitivity effect for the composition of government expenditures (broad public good *vs.* targeted transfers).

The relative electoral sensitivity effect has important implications for the large empirical literature on distributive politics (see, *e.g.*, the literature reviews in Berry, Burden, and Howell (2010) and Golden and Min (2013)). This effect 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. These studies indeed ought to control for the electoral sensitivity of other groups of voters in the same district. In a companion paper, Bouton ⓡ al. (2018), we revisit this empirical literature with the relative electoral sensitivity effect in mind.

References

Aidt, T.S., Jayasri Dutta, and Elena Loukoianova. 2006. "Democracy comes to Europe: Franchise extension and fiscal outcomes 1830-1938." *European Economic Review* 50 (2):249–283.

Albouy, David. 2009. "The Unequal Geographic Burden of Federal Taxation." *Journal of Political Economy* 117 (4):635–667.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* 114 (4):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* 96 (4):767–777.

Atkinson, A.B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* :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* 18 (3):653–678.

Banerjee, Abhijit, Lakshmi Iyer, and Rohini Somanathan. 2008. "Public Action for Public Goods." In *Handbook of Development Economics*, vol. 4, edited by T. Paul Schultz and John A. Strauss, chap. 49. Elsevier, 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* 104 (4):783799.

Besley, Timothy and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *The Quarterly Journal of Economics* 117 (4):1415–1451.

Betz, Timm. 2017. "Trading Interests: Domestic Institutions, International Negotiations, and the Politics of Trade." *The Journal of Politics* 79 (4):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* 139 (1):197–225.

Bormann, Nils-Christian and Matt Golder. 2013. "Democratic Electoral Systems around the world, 1946-2011." *Electoral Studies* 32 (2):360 – 369.

Castanheira, Micael, Gaëtan Nicodème, and Paola Profeta. 2012. "On the political economics of tax reforms: survey and empirical assessment." *International Tax and Public Finance* 19 (4):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* 38 (04):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* 37 (2):152–187.

Cox, Gary W. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34 (4):903–935.

Dixit, Avinash and John Londregan. 1996. "The Determinants of Success of Special Interests in Redistributive Politics." *The Journal of Politics* 58 (4):1132–1155.

Enelow, James and Melvin Hinich. 1982. “Nonspatial Candidate Characteristics and Electoral Competition.” *The Journal of Politics* 44 (1):115–130.

Evans, Carolyn L. 2009. “A Protectionist Bias In Majoritarian Politics: An Empirical Investigation.” *Economics & Politics* 21 (2):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* 11 (5):1178–1203.

Gagliarducci, Stefano, Tommaso Nannicini, and Paolo Naticchioni. 2011. “Electoral Rules and Politicians’ Behavior: A Micro Test.” *American Economic Journal: Economic Policy* 3 (3):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. URL `https://ideas.repec.org/p/cpr/ceprdp/13081.html`.

Golden, Miriam A. and Brian K. Min. 2013. “Distributive Politics Around the World.” *Annual Review of Political Science* 16:73–99.

Grossman, Gene M. and Elhanan Helpman. 1996. “Electoral Competition and Special Interest Politics.” *The Review of Economic Studies* 63 (2):265–286.

———. 2005. “A Protectionist Bias in Majoritarian Politics.” *The Quarterly Journal of Economics* 120 (4):1239–1282.

Hatfield, John William and William R. Hauk. 2014. “Electoral regime and trade policy.” *Journal of Comparative Economics* 42 (3):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* 100 (2).

Jacobson, Gary C. 1985. "Party Organization and Distribution of Campaign Resources: Republicans and Democrats in 1982." *Political Science Quarterly* 100 (4):603–625.

Johansson, Eva. 2003. "Intergovernmental grants as a tactical instrument: empirical evidence from Swedish municipalities." *Journal of Public Economics* 87 (5):883 – 915.

Keefer, Philip. 2004. "What does political economy tell us about economic development and vice versa?" *Annual Review of Political Science* 7:247–272.

Lindbeck, Assar and Jörgen Weibull. 1987. "Balanced-budget redistribution as the outcome of political competition." *Public Choice* 52 (3):273–297.

Lizzeri, Alessandro and Nicola Persico. 2001. "The Provision of Public Goods under Alternative Electoral Incentives." *American Economic Review* 91 (1):225–239.

———. 2005. "A Drawback Of Electoral Competition." *Journal of the European Economic Association* 3 (6):1318–1348.

Mansfield, Edward D. and Marc L. Busch. 1995. "The political economy of nontariff barriers: a cross-national analysis." *International Organization* 49 (04):723–749.

Martin, Paul S. 2003. "Voting's Rewards: Voter Turnout, Attentive Publics, and Congressional Allocation of Federal Money." *American Journal of Political Science* 47 (1):110–127.

Milesi-Ferretti, Gian Maria, Roberto Perotti, and Massimo Rostagno. 2002. "Electoral Systems and Public Spending." *The Quarterly Journal of Economics* 117 (2):609–657.

Min, Brian. 2015. "Electrifying the Poor: Power and Politics in India." Tech. rep.

Myerson, Roger B. 1993. "Incentives to Cultivate Favored Minorities Under Alternative Electoral Systems." *The American Political Science Review* 87 (4):856–869.

Park, Jong Hee and Nathan Jensen. 2007. "Electoral Competition and Agricultural Support in OECD Countries." *American Journal of Political Science* 51 (2):314–329.

Persson, Torsten. 2002. "Do Political Institutions Shape Economic Policy?" *Econometrica* 70 (3):883–905.

Persson, Torsten and Guido Tabellini. 1999. "The size and scope of government:: Comparative politics with rational politicians." *European Economic Review* 43 (4-6):699–735.

———. 2000. *Political Economics: Explaining Economic Policy*. MIT Press Books. The MIT Press.

———. 2003. *The Economic Effects of Constitutions*, *MIT Press Books*, vol. 1. The MIT Press.

Rickard, Stephanie. 2012. "A Non-Tariff Protectionist Bias in Majoritarian Politics: Government Subsidies and Electoral Institutions." 56:777–785.

Rogowski, Ron and Mark Kayser. 2002. "Majoritarian electoral systems and consumer power: price-level evidence from the OECD countries." *American Journal of Political Science* 46 (3):526–539.

Rogowski, Ronald. 1987. "Trade and the variety of democratic institutions." *International Organization* 41 (02):203–223.

Schady, Norbert R. 2000. "The Political Economy of Expenditures by the Peruvian Social Fund (FONCODES), 1991-95." *The American Political Science Review* 94 (2):289–304.

Snyder, James M. 1989. "Election Goals and the Allocation of Campaign Resources." *Econometrica* 57 (3):637–660.

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* 46 (3):506–514.

Strömberg, David. 2004. "Radio's Impact on Public Spending." *The Quarterly Journal of Economics* 119 (1):189–221.

———. 2008. "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida." *American Economic Review* 98 (3):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.

World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People.* World Bank and Oxford University Press.

Ray, D. ⓡ A. Robson, 2018, Certified random: A new order for co-authorship, *American Economic Review* 108, 489–520.

8 Appendix 1: Theory

Assumption 1

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

$$\widetilde{\nu}_l(\mathbf{q}, \delta) \equiv \Delta u_l(\mathbf{q}) - \delta_d \in (-\frac{1}{2\phi_l}, \frac{1}{2\phi_l})$$

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

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

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: $\sum_l \frac{t_l n_l}{T} \pi_l\left(\mathbf{q}; \delta_{d(l)}\right)$, 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:

$$\begin{aligned}
\max_{\mathbf{q}|\sum_l n_l^{\alpha} q_l \leq y} \pi_{\text{PR}}(\mathbf{q}) &:= \mathbb{E}_{\boldsymbol{\delta}}\left[\sum_l \frac{t_l n_l}{T} \left[\tfrac{1}{2} + \phi_l \left(\Delta u_l(\mathbf{q}) - \delta_{d(l)}\right)\right]\right] \\
&= \tfrac{1}{2} + \tfrac{1}{T} \sum_l s_l \left(\Delta u_l(\mathbf{q}) - \mathbb{E}\left[\delta_{d(l)}\right]\right), \\
&= \tfrac{1}{2} + \tfrac{1}{T} \sum_l s_l \left(\Delta u_l(\mathbf{q}) - \beta_{d(l)}\right), \qquad (20)
\end{aligned}$$

where $d(l)$ is the district to which locality l belongs.[14]

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

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) $P_l^A(\mathbf{q}) = 2s_l\left(\Delta u_l(\mathbf{q}) - \mathbb{E}\left[\delta_{d(l)}\right]\right) - n_l t_l$ and $P_l^B(\mathbf{q}) = n_l t_l - 2s_l\left(\Delta u_l(\mathbf{q}) - \mathbb{E}\left[\delta_{d(l)}\right]\right)$. 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 as for PR.

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:

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

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

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

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 (8), the probability that A wins at least 50% of the votes in district d is:

$$\begin{aligned} p_d(\mathbf{q}) &= \Pr\left[\sum_{l\in d} \frac{s_l}{\sum_{k\in d} t_k n_k}\left(\Delta u_l(\mathbf{q}) - \delta_d\right) \geq 0\right] \\ &= \Pr\left[\delta_d \leq \frac{\sum_{l\in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j\in d} s_j}\right], \end{aligned} \tag{21}$$

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:

$$F_{\delta_d}\left[\frac{\sum_{l\in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j\in d} s_j}\right] = \gamma_d \times \left[\frac{\sum_{l\in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j\in d} s_j} + \frac{1}{2\gamma_d} - \beta_d\right]$$
$$= \tfrac{1}{2} + \gamma_d \times \left[\frac{\sum_{l\in d} s_l \Delta u_l(\mathbf{q})}{\sum_{j\in d} s_j} - \beta_d\right] \qquad (22)$$

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

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

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:

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

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:

$$q_l^* = y\left(\frac{n_l^{\frac{1}{\rho}}}{\sum_k n_k^{\frac{1}{\rho}}}\right). \qquad (23)$$

Under PR , the FOCs in (7) and some straightforward manipulations produces the following allocation of public goods:

$$q_l^{PR} = y\left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^{L}(s_k)^{1/\rho}}\right), \qquad (24)$$

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

Finally, the FOCs for MAJ (11) yield:

$$q_l^{MAJ} = y\left(\frac{\left(\gamma_{d(l)} s_l / \boldsymbol{s}_{d(l)}\right)^{1/\rho}}{\sum_{k=1}^{L}\left(\gamma_{d(k)} s_k / \boldsymbol{s}_{d(k)}\right)^{1/\rho}}\right), \qquad (25)$$

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

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 $A\left(\mathbf{q}^{PR}\right) < A\left(\mathbf{q}^{MAJ}\right)$ iff $y^{E}\left(\mathbf{q}^{PR}\right) > y^{E}_{\rho}\left(\mathbf{q}^{MAJ}\right)$ where

$$y^{E}(\mathbf{q}) = \begin{cases} \Pi_l \left(q_l/n_l\right)^{n_l} & \text{if } \rho = 1; \\ \left[\frac{\sum_l n_l (q_l)^{1-\rho}}{\left(\sum_j n_j^{\frac{1}{\rho}}\right)^{\rho}}\right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \tag{26}$$

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

$$\begin{aligned} \Pi_l \left(\frac{t_l \phi_l}{\sum_{k=1}^{L} s_k}\right)^{n_l} &> \Pi_l \left(\frac{t_l \phi_l \gamma_{d(l)} / \boldsymbol{s}_{d(l)}}{\sum_{k=1}^{L} s_k \gamma_{d(k)} / \boldsymbol{s}_{d(k)}}\right)^{n_l} \\ \Pi_l \left(\frac{1}{\sum_{k=1}^{L} s_k}\right)^{n_l} &> \Pi_l \left(\frac{\gamma_{d(l)} / \boldsymbol{s}_{d(l)}}{\sum_{k=1}^{L} s_k \gamma_{d(k)} / \boldsymbol{s}_{d(k)}}\right)^{n_l}. \end{aligned} \tag{27}$$

Note that the denominator on the RHS of (27) is equivalent to $\sum_d \frac{\gamma_d}{\boldsymbol{s}_d} \sum_{k \in d} s_k = \sum_d \gamma_d$. Similarly, we can re-write the denominator on the LHS of (27) as $\sum_d \boldsymbol{s}_d$.

Substituting for these into (27), we get

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

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

$$\begin{aligned} -\log\left[\sum_{d'} s_{d'}\right] &> \sum_l n_l \log\left[\frac{\gamma_{d(l)}}{s_{d(l)}}\right] - \log\left[\sum_{d'} \gamma_{d'}\right] \quad \textit{or} \\ -\log\left[\sum_{d'} s_{d'}\right] &> \sum_d \boldsymbol{n}_d \log \gamma_d - \sum_d \boldsymbol{n}_d \log \boldsymbol{s}_d - \log\left[\sum_{d'} \gamma_{d'}\right] \quad \textit{or} \\ \sum_d \boldsymbol{n}_d \log\left[\frac{s_d}{\sum_{d'} s_{d'}}\right] &> \sum_d \boldsymbol{n}_d \log\left[\frac{\gamma_d}{\sum_{d'} \gamma_{d'}}\right], \end{aligned} \tag{28}$$

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

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 $\left(\sum_j n_j^{\frac{1}{\rho}}\right)^{\rho/(1-\rho)}$. This tells us that $A\left(\mathbf{q}^{PR}\right) < A\left(\mathbf{q}^{MAJ}\right)$ iff:

$$\left[\sum_l n_l \left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^{L}(s_k)^{1/\rho}}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} > \left[\sum_l n_l \left(\frac{\left(s_l \gamma_{d(l)}/\boldsymbol{s}_{d(l)}\right)^{1/\rho}}{\sum_{k=1}^{L}\left(s_k \gamma_{d(k)}/\boldsymbol{s}_{d(k)}\right)^{1/\rho}}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} \tag{29}$$

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 ($\boldsymbol{n}_d = 1/D\ \forall d$), inequality (28) 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 $\frac{\gamma_d}{\sum_{d'} s_{d'}}$ is a mean preserving spread of $\frac{s_d}{\sum_{d'} s_{d'}}$ (and vice versa).

Next, consider the case $\rho \neq 1$ and $L = D$. With one locality per district, all relative sensitivities are 1 $(s_l/s_{d(l)} s_l = 1)$. Simplifying for $n_l = 1/L$, inequality (29) becomes:

$$\left[\sum_l \left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^{L}(s_k)^{1/\rho}}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} > \left[\sum_l \left(\frac{\left(\gamma_{d(l)}\right)^{1/\rho}}{\sum_d (\gamma_d)^{1/\rho}}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}.$$

or

$$\frac{1}{1-\rho}\sum_l \left(\frac{(s_l)^{1/\rho}}{\sum_{k=1}^{L}(s_k)^{1/\rho}}\right)^{1-\rho} > \frac{1}{1-\rho}\sum_l \left(\frac{\left(\gamma_{d(l)}\right)^{1/\rho}}{\sum_d (\gamma_d)^{1/\rho}}\right)^{1-\rho}. \tag{30}$$

Again, the strict concavity of the CRRA function implies (Atkinson (1983)) that this inequality holds if $\frac{\gamma_d}{\sum_{d'} s_{d'}}$ is a mean preserving spread of $\frac{s_d}{\sum_{d'} s_{d'}}$ (and vice versa).

Targeted versus Universal Spending

The objectives in MAJ and PR are similar to (6) and (10) adjusting for the utility $\Delta w_l(\mathbf{q}, G)$ and the budget constraint. In PR, the first order conditions are thus:

$$\sum_l s_l u'(G) = T\lambda^{PR} \tag{31}$$

and

$$\frac{s_l}{n_l} = T\lambda^{PR} \ \forall l \text{ with } q_l > 0 \tag{32}$$

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. If some transfers are given, then

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

A necessary and sufficient condition for some transfer to arise in equilibrium is:

$$u'(y) < \max_l \frac{s_l}{n_l} \frac{1}{\sum_{k=1}^{L} s_k}. \tag{34}$$

In MAJ, the first order conditions become:

$$\sum_{d \in D} \gamma_d u'(G) = D\lambda^{MAJ}, \tag{35}$$

and

$$\frac{\gamma_{d(l)}}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} = D\lambda^{MAJ} \ \forall l \text{ with } q_l > 0 \tag{36}$$

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

In a symmetric equilibrium, then only individuals in the locality with the highest left hand side in (36) could receive a transfer. If some transfers arise in equilibrium:

$$\max_l \frac{\gamma_{d(l)}}{\sum_{d \in D} \gamma_d} \frac{1}{n_l} \frac{s_l}{\sum_{k \in d(l)} s_k} = u'(G). \tag{37}$$

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:

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

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:

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

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

Note that:

$$\sigma_E^2(\mathbf{q}) = \frac{D}{4} - \sum_{d \in C} \gamma_d^2 \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d\right]^2 \tag{38}$$

which implies:

$$S(\mathbf{q}) = \frac{-\sum_d \gamma_d \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d\right]}{\left(\frac{D}{4} - \sum_d \gamma_d^2 \left[\sum_{l \in d} \frac{s_l \Delta u_l(\mathbf{q})}{\sum_{j \in d} s_j} - \beta_d\right]^2\right)^{1/2}} \tag{39}$$

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

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

$$\text{s.t.} \; \sum_l n_l^\alpha q_l = y,$$

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 (18).

9 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 dataset[15] 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.

[15] Coppedge, 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.

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

$$E_{it} = \beta_0 + \beta_1 Pr_{it} + \beta_2 X_{it} + \delta_t + \beta_4 R_i + \epsilon_{it} \quad (40)$$

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 presidential[16], 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.[17]

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.

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

[17]The 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.

Table 1: Proportional Representation and Encompassing 1992-2013

Equation No.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
PR	-0.122	-0.001	0.015	-0.362**				
	(-0.78)	(-0.01)	(0.11)	(-2.13)				
Presidential	-0.025	0.075	0.044	0.127*	-0.057	0.073	0.057	0.127*
	(-0.13)	(0.46)	(0.33)	(1.87)	(-0.29)	(0.46)	(0.45)	(1.86)
District Magn					-0.396*	-0.198	-0.095	0.013
					(-1.92)	(-0.73)	(-0.48)	(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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- *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.[18]

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

- 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;

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

- 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: [19]

- POLITY2, a measure of democracy,
- NYYRSDEM: 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.

[19]Blume 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.

Appendix 3: Online Appendix

Should we expect government spending to be higher under MAJ or under PR? To address this question, the model in this appendix endogenizes the size of the government budget. For the sake of simplicity and comparability with previous results in the literature (see, *e.g.*, Persson (2002)), we assume that the government finances its interventions through a non-distortionary and linear tax τ on income.

To isolate the effect coming from the targeting of resources by the government, we need to remove all possibility of local targeting using taxation. To do so, we assume that all individuals have the same income y. The government budget constraint is thus: $\sum_l n_l^\alpha q_l \leq y\tau$.

Individuals of locality l have the following preferences for taxation τ and the government interventions $\mathbf{q}$:

$$w_l\left(\mathbf{q}, \tau\right) = v(y(1-\tau)) + u_l(\mathbf{q}), \tag{41}$$

with u, v strictly increasing and strictly concave.

We focus on the case of local public goods ($\alpha = 0$) and assume $\gamma_d = \gamma > 0$, $\forall d$. This last assumption guarantees that the effects we identify are not driven by differences in contestability across districts, a mechanism already identified in the existing literature (see, *e.g.*, Persson and Tabellini (1999), and Milesi-Ferretti, Perotti, and Rostagno (2002)).

The optimization problems of the parties in PR and MAJ are similar to (6) and (10) respectively but for three differences: (i) Δu_l replaced by Δw_l, as defined in (41), (ii) the total budget is $y\tau$, and (iii) parties choose both $\mathbf{q}$ and τ.

In PR, the first order conditions become:

$$\sum_l s_l \cdot v'(y(1-\tau)) = T\lambda^{PR}, \tag{42}$$

$$s_l u_l'(\mathbf{q}) = T\lambda^{PR} \; \forall l, \tag{43}$$

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

While in MAJ, we have :

$$\sum_d \gamma \; v'(y(1-\tau)) = D\lambda^{MAJ}, \tag{44}$$

$$\gamma \frac{s_l}{\sum_{k\in d(l)} s_k} u'(q_l) = D\lambda^{MAJ} \; \forall l, \tag{45}$$

where λ^{MAJ} is the multiplier associated with the budget constraint in MAJ.

In each of these sets of first order conditions, we can average the second equality over all localities and equate the left-hand-sides to see that in a symmetric equilibrium the following must be satisfied:

$$v'(y(1-\tau^{PR})) = \frac{1}{L}\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u_l'(\mathbf{q}^{PR}) \text{ in PR;} \tag{46}$$

$$v'(y(1-\tau^{MAJ})) = \frac{1}{LD}\sum_{l=1}^{L}\frac{s_l}{\sum_{k\in d(l)}s_k}u_l'(\mathbf{q}^{MAJ}) \text{ in MAJ.} \tag{47}$$

From these two conditions, we see that the total budget depends positively on a weighted average of the electoral sensitivities in PR and on a weighted average of the *relative* electoral sensitivities in MAJ.

To explore the effect of the electoral system on the size of the government, we focus on two extreme cases that give full power to the forces that distinguish MAJ and PR in our model. These two cases exacerbate the difference between absolute and relative electoral sensitivity. In case 1, electoral sensitivities are the same, $s_l = s_{l'}\ \forall l, l'$, but relative electoral sensitivities can differ. In case 2, there is one locality per district so that relative electoral sensitivities are the same but electoral sensitivities can differ.

The overall message that stems out of the analysis of those two cases is that when the electoral system incentivizes parties to distribute resources more equality across localities, parties also have incentives to increase the size of the government. This is encapsulated in the following proposition:

Proposition 2 *If $u''' \leq 0\ \forall l$, then the government is (i) larger in* PR *than in MAJ, $\tau^{PR} > \tau^{MAJ}$, in case 1; (ii) smaller in* PR *than in MAJ, $\tau^{PR} < \tau^{MAJ}$, in case 2.*

The intuition is that when politicians want to spread the resources unequally, the diminishing marginal utility of q is decreasing the impact of the marginal unit of public good in the localities that they want to target. The restriction on $u''' \leq 0$ ensures that changes in the expected marginal utility when resources are spread do not counteract this effect. While we cannot prove a general result when $u''' > 0$, it is easy to show that the results in Proposition 2 are not necessarily reversed. For instance, they hold for the case of logarithmic utility.

Proof of Proposition 2

Case 1: In this case, all s_l are the same though relative sensitivities differ.

From the first order conditions (43) and $s_l = s_{l'}$ $\forall l, l'$, we have that $q_l^{PR} = q_{l'}^{PR}$ $\forall l, l'$ and therefore that

$$u'(q_l^{PR}) = \frac{1}{L}\sum_{l=1}^{L} u'(q_l^{PR}) \tag{48}$$

Conditions (46) and (47) boil down to:

$$v'(y(1-\tau^{PR})) = \frac{1}{L^2}\sum_{l=1}^{L} u_l'(\mathbf{q}^{PR}); \tag{49}$$

$$v'(y(1-\tau^{MAJ})) = \frac{1}{L}\sum_{l=1}^{L} \frac{1}{DL_{d(l)}} u_l'(\mathbf{q}^{MAJ}), \tag{50}$$

where $L_{d(l)}$ is the number of localities in district $d(l)$. These two conditions together with the concavity of v give us that $\tau^{PR} > \tau^{MAJ}$ if and only if

$$\frac{1}{L}\sum_{l=1}^{L} u_l'(\mathbf{q}^{PR}) > \sum_{l=1}^{L} \frac{1}{DL_{d(l)}} u_l'(\mathbf{q}^{MAJ}). \tag{51}$$

The first order condition (45) implies that q_l^{MAJ} is positively correlated with $\frac{1}{L_{d(l)}}$ and therefore that

$$\sum_{l=1}^{L} \frac{1}{DL_{d(l)}} q_l^{MAJ} > \sum_{l=1}^{L} \frac{1}{L} q_l^{MAJ}. \tag{52}$$

Now, it follows from Jensen's inequality and $u''' \leq 0$ $\forall l$ that

$$u'(\sum_{l=1}^{L} \frac{1}{DL_{d(l)}} q_l^{MAJ}) > \sum_{l=1}^{L} \frac{1}{DL_{d(l)}} u_l'(\mathbf{q}^{MAJ})). \tag{53}$$

We are now in position to prove the result stated in the proposition. We proceed by contradiction. Let us suppose that $\tau^{PR} < \tau^{MAJ}$. This directly implies that $\sum_{l=1}^{L} \frac{1}{L} q_l^{MAJ} > q^{PR}$. From (52), this implies that $\sum_{l=1}^{L} \frac{1}{DL_{d(l)}} q_l^{MAJ} > q^{PR}$ and thus that $u'(\sum_{l=1}^{L} \frac{1}{DL_{d(l)}} q_l^{MAJ}) < u'(q^{PR})$. Combining this with (48) and (53), we obtain:

$$\frac{1}{L}\sum_{l=1}^{L} u_l'(\mathbf{q}^{PR}) > \sum_{l=1}^{L} \frac{1}{DL_{d(l)}} u_l'(\mathbf{q}^{MAJ}). \tag{54}$$

From (51), this gives $\tau^{PR} > \tau^{MAJ}$, a contradiction. We must thus have $\tau^{PR} > \tau^{MAJ}$.

Case 2. For this case relative electoral sensitivities are all 1 while sensitivities differ.

Conditions (46) and (47) then boil down to:

$$v'(y(1-\tau^{PR})) = \frac{1}{L}\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u_l'(\mathbf{q}^{PR}); \tag{55}$$

$$v'(y(1-\tau^{MAJ})) = \frac{1}{L^2}\sum_{l=1}^{L}u_l'(\mathbf{q}^{MAJ}). \tag{56}$$

These two conditions together with the concavity of v give us that $\tau^{PR} > \tau^{MAJ}$ if and only if

$$\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u_l'(\mathbf{q}^{PR}) > \frac{1}{L}\sum_{l=1}^{L}u_l'(\mathbf{q}^{MAJ}). \tag{57}$$

From the first order conditions (45), we have that $q_l^{MAJ} = q_{l'}^{MAJ}\ \forall l, l'$ and therefore that

$$\sum_{l=1}^{L}u'(q_l^{MAJ}) = \sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u'(q_l^{MAJ}). \tag{58}$$

The first order condition (43) tells us that q_l^{PR} is positively correlated with s_l, so that

$$\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}q_l^{PR} > \frac{1}{L}\sum_{l=1}^{L}q_l^{PR}. \tag{59}$$

This and the concavity of u imply

$$u'(\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}q_l^{PR}) < u'(\frac{1}{L}\sum_{l=1}^{L}q_l^{PR}). \tag{60}$$

Now, $u''' \leq 0\ \forall l$ and Jensen's inequality mean that

$$\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u'(q_l^{PR}) < u'(\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}}q_l^{PR}). \tag{61}$$

We are now in position to prove the result stated in the proposition. We proceed by contradiction. Let us suppose that $\tau^{PR} > \tau^{MAJ}$. This directly implies that $\sum_{l=1}^{L}\frac{1}{L}q_l^{PR} > q^{MAJ}$ and thus that $u'(\sum_{l=1}^{L}\frac{1}{L}q_l^{PR}) < u'(q^{MAJ})$. Combining this with conditions (60) and (61), we obtain:

$$\sum_{l=1}^{L}\frac{s_l}{\sum_{k=1}^{L}s_k}u_l'(\mathbf{q}^{PR}) < u'(q_l^{MAJ}) \tag{62}$$

From (57), given that $q_l^{MAJ} = q_{l'}^{MAJ}\ \forall l, l'$, this implies $\tau^{PR} < \tau^{MAJ}$, a contradiction. We must thus have $\tau^{PR} < \tau^{MAJ}$.