

Political Systems and the Inequality of Provision of Local Public Goods

Laurent Bouton ^{*},
Micael Castanheira [§], Garance Genicot^{*} and Dario Sansone ^{*}.

January 2018

ABSTRACT

This paper studies the provision of *local* public goods and the effect of the political system on its allocation. Recognizing that most public goods are de facto local, we propose a model of allocation of local public good under political competition. We derive predictions regarding the relationship between public good provision and population in localities that differ depending on the regime: majoritarian and proportional representation systems. Using the satellite nightlight data as a proxy for local public good provision, we show that the predicted patterns are observed. Our finding raises interesting questions regarding the measurement of inequality in public good allocation.

EXTREMELY PRELIMINARY. PLEASE DO NOT POST OR DISTRIBUTE.

JEL Classification Numbers:

Key Words: Public Good, Inequality, Political System, Nightlight.

* Georgetown University; § Universite Libre de Bruxelles. Correspondence: lb910@georgetown.edu, micael.casta@gmail.com, garance.genicot@georgetown.edu and ds1289@georgetown.edu.

1 Introduction

Despite increased political and academic interest in the economics of inequality, still little is known about inequality in access to public goods and services. The theoretical concept of “public goods” is naïve in *assuming* that they benefit everyone equally. The twin assumptions of non-rivalry and non-excludability evacuates the question of public good accessibility. However the question matters because most public goods are local in nature. Take electricity or sanitation provision for instance: given capacity constraints in poor countries, providing access to the largest possible population may require to first focus supply on high population density areas.

The purpose of this paper is to study the provision of *local* public goods and the impact of the political system on their allocation.

There is a large existing empirical and theoretical literature on how and which national institutions gear policy either towards general public good provision or particularistic “pork barrel” targeted. A recurrent theme in the literature (Persson and Tabellini (1999, 2000); Persson (2002); Lizzeri and Persico (2001, 2005); Milesi-Ferretti, Perotti, and Rostagno (2002), and Myerson (1993)) is that politicians have incentives to target a smaller fraction of the population under majoritarian systems than under proportional representation and that therefore there are fewer public goods and more inequality under majoritarian systems.

Most empirical analyses at the international level must make strong assumptions about which items in the government budget can reasonably be thought to represent public goods as opposed to transfers (see Section 4.2 in Bouton, Castanheira, and Genicot (2016)). Moreover, these distinctions rest on the assumption that there exists something like a “universal public good”. Instead, with some exceptions such as nuclear deterrence, one is bound to admit that “public goods” are typically geographically targetable. The key question is then to identify when governments exploit their margins of action to target them in practice or not. 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 instead 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); ?); Gagliarducci, Nannicini, and Naticchioni (2011); Min (2015); Strömberg (2008), and Golden and Min (2013) for a survey).

This paper revisits the question of the economic impact of political institutions assuming that politicians decide on the allocation of geographically targetable public goods. We introduce a model of political competition where politicians promise local public goods in order to gain votes and contract majoritarian and proportional representation. We show that proportional representation systems give strong incentive to politicians to allocate more public good to densely populated areas. This is so because PR does not impose any constraint on where these votes should be coming from. The parties are thus allowed to concentrate resources in areas with higher population density. In contrast, in majoritarian systems, the winning party needs to win districts. This provides incentives to politician to (1) target relatively populated areas within a district (but not necessarily the most populated areas in the country) ; (2) give up on localities in non swingable districts. We show that under reasonable conditions this implies a tilt in the relationship between public good levels and population, with a steeper slope for proportional representation systems.

Empirically, we use pixel-level satellite measures of luminosity at night to assess the location of public light provision in a country. Combining these data with population data from the LandScan project, we have, for all countries, information at the *local* level both about the population and the public light provision.

In essence, this produces about 28 million observations that allow us to precisely track how night lights are geographically targeted across virtually all population groups on Earth. This allows us to construct new indicators of the inequality in public light supply across the population groups at the country and at the subdis-

trictal levels.

Pursuing the analysis further, we then exploit pixel level data in developing democratic countries to analyze how night lights are distributed across each population density levels . We observe that, in comparison with majoritarian elections, proportional representation are more responsive to population: they provide more light to densely populated cities, while the opposite is true in rural areas.

In addition to this, we test empirically the predictions from the theoretical section. We provide evidence that population in the districts matters more in majoritarian elections, while national population affects the relative importance of pixel population more in countries with proportional representation. We show that these conclusions remain valid in several robustness checks such as different functional forms, different years, different controls, and different definitions of electoral systems.

Section 2 introduces the model of allocation of public provision and discusses two normative benchmarks for the case of public good provision, and the question of how to measure inequality in public good provision. Section 3 studies the allocation of public provision under political competition and compares the allocation of public goods under majoritarian and proportional representation systems and Section 4 derives theoretical predictions regarding the relationship between public good provision and population. Next, Section 5 tests the predicted patterns using the nightlights data. Finally Section 6 concludes.

2 Local Public Good Allocation

2.1 Premises of the Model

Consider one country with a continuum of voters of total mass 1. The population is partitioned into localities

$l \in \{1, 2, \dots, L\}$ of size n_l , *s.t.* $\sum_l n_l = 1$. Each locality belongs to an *electoral district* $d \in \{1, 2, \dots, D\}$. The size of the population in district d is $m_d = \sum_{l \in d} n_l$. Naturally, $\sum_d m_d = 1$.

An elected government has to allocate resources y to the provision of locality-specific public goods $\mathbf{q} = \{q_1, \dots, q_L\}$. This implies that public goods 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).

For simplicity, the cost of providing public goods to some locality l is assumed to be linear in the quantity provided: $k_l(q_l) = q_l$, such that the aggregate budget constraint of the government can be expressed as $\sum_l q_l = y$, where y is a parameter representing the total budget of the government.

Individuals of locality l have preferences $u_l(\mathbf{q})$ for the public good, 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, $\partial u_l(\mathbf{q}) / \partial q_k = 0, \forall k \neq l$, meaning that public goods are purely local and do not produce spillovers across localities.

2.2 Normative Benchmark: A Thick or Thin Veil of Ignorance

Before introducing political competition, we want to consider the politics-free benchmark. The question is how would the social planner allocate public goods to the different locations l under the budget constraint that $\sum_l q_l = y$ (where we assume that the marginal cost $k = 1$). To answer this question, we first need to determine what is the objective of the social planner.

Harsanyi (1953, 1955) and Rawls (1971), theories of social justice argue that societies make choices under what Rawls terms the *original position*, behind a “veil of ignorance” that prevents people from “knowing their own social and economic

positions, their own special interests in the society, or even their own personal talents and abilities (or their lack of them) (Harsanyi, 1975; p.594). The question is how thick is that veil of ignorance? Kurtulmus (2012) argues that Rawls’s veil of ignorance is thicker than Harsanyi’s.

For our purpose, the relevant dimension is the distribution of the population. What does an individual behind the veil of ignorance know about the distribution of the population? Behind a ‘thick’ veil of ignorance, individuals do not know the likelihood that they may end up in any location and therefore may assign the same probability of being born in any possible location.

In this case then this benchmark defines the average of the expected individual citizens’ preferences in each location as the social planner’s objective :

$$\mathcal{W}^G(\mathbf{q}) = \frac{1}{L} \sum_l u_l(\mathbf{q}). \quad (1)$$

We call this the *geographical utilitarian social welfare*. In this case, the planner’s ideal would be to divide the budget equally across the different locations.

Now equal distribution of public good is an optimum under the assumption, made in (1), that all locations are as likely as each other under the veil of ignorance. As a result highly populated and nearly uninhabited areas carry the same weight.

However, this interpretation of the veil of ignorance may be too extreme. Behind Harsanyi’s arguably thinner veil of ignorance, individuals know that they have an equal chance of being in any person’s position. This implies that they know about the distribution of positions, in our case, locations, in society. One does not know where he or she will be born but believe that the likelihood to be born in a given location is proportional to the population actually living in the given location.

Taking this interpretation, the social planner’s objective under the veil of ignorance

is given by a weighted average of individual citizens' preferences:

$$\mathcal{W}^P(\mathbf{q}) = \sum_l n_l u_l(\mathbf{q}), \quad (2)$$

We call this the *population utilitarian social welfare*. In this case, localities that are more populated will receive more local public good.

2.3 Inequality of Local Public Good Provision

What does this imply for measuring inequalities of provision of local public goods? Atkinson (1970, 1983) proposed a welfare-based measure of inequality. Assuming CRRA preferences, *Atkinson's index of inequality* compares the actual mean income in society to minimum mean income needed to achieve the same amount of welfare. We can adapt this concept to derive two measures of inequality of provision, one for each of the welfare objectives discussed in 2.2.

Following Atkinson, we work under the assumption that individuals have CRRA preferences:

$$u_l(\mathbf{q}) = u(q_l) = \begin{cases} \ln(q_l) & \text{if } \rho = 1; \\ \frac{q_l^{1-\rho}}{1-\rho} & \text{if } \rho \neq 1 \text{ and } \rho > 0. \end{cases}$$

For the geographical utilitarian social welfare (1), a planner could achieve the highest utility possible with a budget y , $\widetilde{W}^G(y)$, by giving to each location a share $s^G = 1/L$ of the budget. This allows us to define y^G as the smallest budget needed to reach the same level of social welfare as the actual allocation of public goods:

$$y^G = \widetilde{W}^{G^{-1}}(\mathcal{W}^G(\mathbf{q})) = \begin{cases} L \Pi_l (q_l)^{1/L} & \text{if } \rho = 1; \\ L^{\frac{-\rho}{1-\rho}} (\sum_l q_l^{1-\rho})^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases}$$

The equivalent to the Atkinson index of inequality is then

$$A_\rho^G \equiv 1 - \frac{y^G}{y} = \begin{cases} 1 - \frac{(\Pi_l q_l)^{1/L}}{y/L} & \text{if } \rho = 1; \\ 1 - \left(\frac{1}{L} \sum_l \frac{q_l}{y/L} \right)^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \quad (3)$$

This measure of inequality of public provision is 0 when local public goods are equally provided in all localities, and it is maximal if one location receives the entire budget. This is what we call a measure of *geographical inequality in public good provision*.

In contrast, under the population utilitarian social welfare more populated localities receive more. With CRRA utilities, maximizing (2) implies that a locality l would receive a share $s_l^P = \frac{n_l^{\frac{1}{\rho}}}{\sum_j n_j^{\frac{1}{\rho}}}$ of the budget y . The level of social welfare corresponding to the planner's ideal allocation of a budget y is therefore:

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

Following the same logic as before, we define the *equivalent* budget y^P as the budget for local public goods that would be needed to achieve the same welfare as the actual allocation of public goods \mathbf{q} , that is $y^P = \widetilde{W}^{P-1}(\mathcal{W}^P(\mathbf{q}))$.

The resulting measure of inequality *à la* Atkinson is then

$$A_\rho^P \equiv 1 - \frac{y^P}{y} = \begin{cases} 1 - \frac{1}{y} \Pi_l (q_l/n_l)^{n_l} & \text{if } \rho = 1; \\ 1 - \left[\frac{\sum_l n_l (q_l/y)^{1-\rho}}{(\sum_j n_j^{1/\rho})^\rho} \right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \quad (4)$$

This is what we call the *population based measure of inequality in public good provision*. If the differences in local public provision are only reflecting inequalities in the population distribution but are optimal according to (2) then this measure deems

that there is no inequality in the provision of public goods.

3 Political Equilibria

3.1 A Model of Competition

Two parties, A and B , compete for seats in the election. Their objective is to maximize their expected number of seats in the national assembly. Our purpose is to see how the electoral system influences the allocation of public goods: we will compare incentives under **PR**, where seats are attributed proportionately to the *fraction of national votes* garnered by each party, and under **MAJ**, where seats are proportional to the *fraction of districts* won by each party. The party winning a district is the one with the most votes in that district. Thus, electoral districts are of secondary importance in **PR**, but they are highly relevant under **MAJ**.

The assumption that parties *maximize their seat share in the national assembly* independently of the electoral system deserves discussion. Many political economy models assume that, in MAJ, parties maximize the probability of obtaining a majority of seats in the national assembly whereas they maximize the vote share under proportional representation (see, *e.g.* Lizzeri and Persico 2001, Stromberg 2008). This has the implication 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 pivotality of a seat/district in the national assembly. However, there are arguments in favor of assuming that, in MAJ too, parties maximize the number of seats (for instance it captures the benefits of having larger coalitions) and there is some empirical support for this assumption (see, *e.g.* Jacobson and Kernell 1985b and Incerti 2015 *vs.* Snyder 1989). In this paper, we follow this second approach, which has the added benefit of tractability, by considering that under both systems parties maximize the expected number of seats.

To maximize their seat share, both parties non-cooperatively make simultaneous and binding budget proposals \mathbf{q}^A and \mathbf{q}^B that detail the provision of all locality-specific public goods. These proposals must satisfy the government budget constraint. Then, given these proposals, voter i in locality l cast her ballot for party A if:

$$\Delta u_l(\mathbf{q}) - \nu_{i,l} - \delta_d \geq 0 \tag{5}$$

where $\Delta u_l(\mathbf{q}) := u_l(\mathbf{q}^A) - u_l(\mathbf{q}^B)$ is the policy component of the preferences and:

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

capture the political preferences that are *ex ante* unknown to the parties in the probabilistic voting tradition (see, e.g., Enelow and Hinich 1982, Lindbeck and Weibull 1987, Dixit and Londregan, 1995, and Persson and Tabellini 2001). From the standpoint of politicians, each individual voter in a locality l has political preferences that are the results of two random shocks. The first is an individual specific preference shock $\nu_{i,l}$. These are independent and identically distributed draws from a locality specific distribution. We assume that the uncertainty surrounding each voter’s political preferences is locality specific: ϕ_l may differ across localities. The second shock is common to all voters in a given district, δ_d . Note that each district is characterized by a deterministic bias β_d in favor of B (when positive; against B when negative) and a district-specific random component, that has density γ_d . We call γ_d the *swingness* of district d .¹

For a given district shock δ_d , (5) identifies the “swing voter” in locality l as:

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

Voters who experience a shock $\nu_{i,l} < (>) \nu_l(\mathbf{q}, \delta_d)$ strictly prefer to vote A (B).

¹Note that, for our purpose, adding locality-specific biases and a national shock would only complicate the notation without adding insight.

Throughout the paper, we assume that there are voters to be swung in any locality:

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

With this Assumption, locality-level vote shares aggregate into the *vote share of A in district d* as:

$$\pi_d^A(\mathbf{q}; \delta_d) = \frac{1}{2} + \sum_{l \in d} \frac{\phi_l n_l}{m_d} [\Delta u_l(\mathbf{q}) - \delta_d]. \quad (6)$$

3.2 Proportional Representation System

Under PR, maximizing the expected number of seats in the national assembly is equivalent to maximizing the country-wide expected vote share:

$$\pi_{PR}^A(\mathbf{q}) \equiv \mathbb{E} \left[\sum_d m_d \pi_d^A(\mathbf{q}; \delta_d) \right],$$

subject to the aggregate budget constraint $\sum_l k_l(q_l) = y$. As shown in Appendix, this leads to the following objective:

$$\max_{\mathbf{q}^A | \sum_l k_l(q_l) = y} \pi_{PR}^A(\mathbf{q}) = \frac{1}{2} + \sum_l s_l (\Delta u_l(\mathbf{q}) - \beta_{d(l)}),$$

where $d(l)$ is the district to which locality l belongs and s_l captures the *electoral sensitivity* of locality l :

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

The first order conditions are thus:

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

where λ^{PR} is the Lagrange multiplier of the budget constraint under PR. Following the same steps for party B shows that $\mathbf{q}^A = \mathbf{q}^B$ in equilibrium.

It follows that localities that are more electorally *sensitive* – with a higher s_l either because they are more populated or more homogeneous – receive *ceteris paribus* a larger provision of public goods.

3.3 Majoritarian System

Under MAJ, seats are attributed in proportion to the number of districts won by each party. The probability that A wins district d , $p_d^A(\mathbf{q})$, is given by $Pr(\pi_d^A(\mathbf{q}; \delta_d) \geq 1/2)$, or

$$p_d^A(\mathbf{q}) = Pr\left(\delta_d \leq \tilde{\delta}_d(\mathbf{q}) \equiv \sum_{l \in d} \frac{s_l}{\sum_{k \in d} s_k} \Delta u_l(\mathbf{q})\right). \quad (9)$$

Following Persson and Tabellini (1999) and Galasso and Nannicini (2011), we consider only two types of districts: *contestable* or *non-contestable* districts. Contestable districts are such that, for any allocation, each party has a positive probability of winning: $C \equiv \{d | p_d^A(\mathbf{q}) \in]0, 1[\forall \mathbf{q}\}$. In contrast, non-contestable districts are such that, for any allocation, one of the party has a zero probability of winning: that is, $N \equiv \{d | p_d^A(\mathbf{q}) = 0 \text{ or } p_d^A(\mathbf{q}) = 1 \forall \mathbf{q}\}$.

Assumption 2 (Contestability) $\forall d, d \in C \cup N$.

The precise conditions on the parameters required for this assumption to hold can be found in Appendix.

Notice that in contestable district, the probability that $\delta_d = \tilde{\delta}_d$ and therefore that a small increase in incentive is pivotal is given by γ_d while in non-contestable districts the probability of being pivotal is 0:

$$\psi_d = \begin{cases} \gamma_d & \text{if } d \in C \\ 0 & \text{otherwise} \end{cases}$$

A maximizing his expected seat share yields the following objective:

$$\max_{\mathbf{q}^A} \pi_{MAJ}^A(\mathbf{q}) = \frac{1}{2} + \frac{1}{D} \sum_d \psi_d \left[\sum_{l \in d} \frac{s_l}{\sum_{k \in d} s_k} \Delta u_l(\mathbf{q}) - \beta_d \right]$$

subject to the budget constraint: $\sum_l k_l(q_l) = y$.

The first order conditions for any locality l in a contestable district are:

$$\frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\lambda^{MAJ}}{\gamma_{d_l}} \frac{\sum_{k \in d_l} s_k}{s_l}, \forall l. \quad (10)$$

where λ^{MAJ} is the Lagrange multiplier associated with the budget constraint under MAJ. Again, we have $\mathbf{q}^A = \mathbf{q}^B$ in equilibrium.

In any non-contestable district, local public good provision will be null.

3.4 Comparing the Systems

The difference between PR and MAJ are summarized in Proposition 1.

Proposition 1 *In PR, $q_l \geq q_{l'}$ iff $s_l \geq s_{l'}$. Under MAJ, for $\psi_{d_l} = \psi_{d_{l'}}$ then $q_l \geq q_{l'}$ iff $\frac{s_l}{\sum_{k \in d(l)} s_k} \geq \frac{s_{l'}}{\sum_{k' \in d(l')} s_{k'}}$; and for $\frac{s_l}{\sum_{k \in d(l)} s_k} = \frac{s_{l'}}{\sum_{k' \in d(l')} s_{k'}}$ then $q_l \geq q_{l'}$ iff $\psi_{d_l} \geq \psi_{d_{l'}}$.*

The most straightforward implication of this proposition concerns the horizontal equality of treatment of two localities. Under PR, any two localities with the same electoral sensitivity will be treated equally. In contrast, under MAJ, similar localities can receive widely different amount of local public goods for two reasons.

First, take two localities l and l' with the same electoral sensitivity in districts with the same probability of being pivotal. Assume that l is surrounded in its district by localities with low electoral sensitivity (due to low population, turnout or swingness)

while l' is surrounded in its district by localities with high electoral sensitivity. Under MAJ, it is the *relative* electoral sensitivity of locality l in the district, as opposed to the *absolute* sensitivity that matters. Hence, following the adage that “in the land of the blind the one-eyed is king”, the public good provision will be higher in l than l' . In addition, the probability of pivotability of the district under MAJ affect how much it receives. Two localities with the same relative sensitivity will be treated unequally, with the most pivotal receiving more. The fact that voters who are located in non-pivotal districts are more likely to be abandoned under MAJ has been stressed in the literature (CITE XXXX) and is a source of horizontal inequality here too.

Above, we have discussed the implications of Proposition 1 in terms of horizontal inequalities in public good provision: how public good provision differs across localities with the same electoral sensitivity under MAJ and PR. The same Proposition also has implications in terms of vertical inequalities: how public good provision differs across localities with different electoral sensitivities.

To understand how MAJ and PR affect vertical inequalities, it is useful to consider an extreme case in which all localities have the same relative electoral sensitivity but not necessarily the same electoral sensitivity. In particular, let us consider a country in which each district is composed of only one locality. This guarantees that all localities have a relative electoral sensitivity equal to 1. From (10) we then have that, under MAJ, any two localities l and l' with the same swingness ($\gamma_{d(l)} = \gamma_{d(l')}$), but potentially very different electoral sensitivities (say, $s_l > s_{l'}$), receive the same level of public good ($q_l = q_{l'}$). There is thus no vertical inequalities in the provision of public goods. This case illustrates how, by tallying the votes at the district level, MAJ may induce parties to ”sprinkle” public goods all over the country.

The situation is different under PR. From (8), we have that parties react to the electoral sensitivity of the localities, even if they all have the same relative electoral sensitivity. As a result, the locality with the higher electoral sensitivity receives

more public good ($q_l > q_{l'}$). We may thus end up with large vertical inequalities in the provision of public goods. This case illustrates how, by tallying the votes at the country level, PR may induce parties to concentrate most of the public good provision in a small number of localities.

The contrast between the allocation of resources under MAJ and PR comes in stark contrast with the traditional view that MAJ induces parties to target resources while PR induces them to provide public goods broadly.

4 Application: Population Patterns

In this Section, we focus on the implication of electoral competition on the relationship between public good provision and population. For this we abstract from other sources of heterogeneity (typically harder to measure than population) and assume going forward that $\phi_l = \phi, \forall l$ and $\gamma_d = \gamma, \forall d$.

4.1 Population and Public Good Provision

With this assumption, the first order conditions (8) and (10) become:

$$PR : \frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\lambda^{PR}}{n_l}, \forall l, \quad (11)$$

$$MAJ : \frac{\partial u_l(\mathbf{q}^A)}{\partial q_l^A} = \frac{\lambda^{MAJ}}{n_l/m_{d(l)}}, \forall l. \quad (12)$$

From these FOCs, it is immediate to see how Proposition 1 applies to the version of the model of this section: comparing public good provision in two localities, l and k , it appears that it only depends on local population sizes n_l/n_k in PR and on *relative* population sizes $\frac{n_l}{m_{d(l)}} / \frac{n_k}{m_{d(k)}}$ in MAJ.² This is the basis for our first empirical

²Interestingly, this implies that the two systems produce the same allocation of public goods if there is no heterogeneity in population across districts.

prediction:

Empirical Prediction 1 *Within countries, public good provision in a locality is a strictly increasing function of local population size in both **MAJ** and **PR** but decreasing in total district population only under **MAJ**.*

What does this result implies for the relationship between the provision of local public goods in a locality and its size? Since **PR** “track” population size (or density, for that matter) more closely than **MAJ**, we expect a higher correlation between public good provision and local population.

When thinking about empirical implications, it is also important to realize that district lines were often drawn much before the intense urbanization that has happened since the 1960s in advanced economies and is still much in progress in developing countries. Imagine that, at the time district lines were drawn, all localities were identical, and $L_D \equiv L/D$ such localities were grouped into also identical districts: $n_l = n, \forall l$, and $m_d = m, \forall d$.

At the time an observer looks at public good allocation, instead, some localities and districts have grown faster than others. We consider two very stylized cases. First, population increases are district specific:

Case 1: Pure district heterogeneity. Population in all localities of district d has been multiplied by χ_d , such that $\chi_1 < \chi_2 < \dots < \chi_D$:

$$n_l = n \cdot \chi_{d(l)}, \forall l \in d$$

In that case, in **PR**, we would have that localities in “higher” districts receive more public goods than localities in “lower” districts, whereas all localities would receive

the same public good provision in **MAJ**. Indeed:

$$PR : \frac{u'_l(q_l^{PR})}{u'_k(q_k^{PR})} = \frac{n_k}{n_l} = \frac{\chi_{d(l)}}{\chi_{d(k)}} \quad (13)$$

$$MAJ : \frac{u'_l(q_l^{MAJ})}{u'_k(q_k^{MAJ})} = \frac{n_k/m_{d(k)}}{n_l/m_{d(l)}} = \frac{n\chi_{d(k)}/(nL_D\chi_{d(k)})}{n\chi_{d(l)}/(nL_D\chi_{d(l)})} = 1 \quad (14)$$

More generally, if different localities in a same district had different population sizes, the ratio of public good supplies would no longer be equal to 1 in **MAJ**, but it would still be lower than in **PR**.

Case 2: Pure locality heterogeneity.

What happens if, starting from the purely equal localities and districts, some localities grow much faster than others? Consider some locality l in district $d(l)$ and compare it to another locality k in another district $d(k)$. For simplicity, assume that, initially, l and k are identical, and that $m_d = m$ in all districts. This implies that the supply of public goods is initially identical in **MAJ** and **PR**.

We consider that starting from districts that are otherwise identical is a good proxy to say that, on average, two randomly sampled districts should have grown at the same pace.

Now, how would both electoral systems react to an increase in the size of locality l ? Remember that $m_d = \sum_{l \in d} n_l$. It is therefore obvious that $\partial m_d / \partial n_l = 1$.

Taking the log ratio of (13) and (14) and differentiating with respect to n_l , we have:

$$\frac{u''_l}{u'_l} (d \log q_l^{PR} - d \log q_l^{MAJ}) = -\frac{dn_l}{n_l} + \left(\frac{dn_l}{n_l} - \frac{dm_{d(l)}}{m_{d(l)}} \right) \quad (15)$$

$$d \log q_l^{PR} - d \log q_l^{MAJ} = \left| \frac{u'_l}{u''_l} \right| \cdot \frac{dn_l}{m_{d(l)}} \quad (16)$$

In other words, a locality that sees its population size increase will typically receive more in **PR** than in **MAJ**.

This leads us to our second empirical prediction:

Empirical Prediction 2 *Within countries, the slope of the relationship between public good provision and population size should be steeper under **PR** than under **MAJ**.*

4.2 Political System and Inequality

Proposition 2 *Pop Ineq lower under PR. Can always get more geo ineq under PR*

5 Empirical Analysis

5.1 Data

In this section, we have focused on democratic countries³. Moreover, since the variation in night lights in developed countries is limited and may be due to other factors unrelated to economic development (e.g. environmental and light pollution concerns), we have restricted the analysis to developing countries. Finally, we have not considered countries that are very small in terms of size and/or population. The resulting number of countries is 71.

The local public good considered is night light. Indeed, as discussed in Appendix A.1.2, several researchers have convincingly argued that satellite data on luminosity at night are a good proxy for the geographic distribution of public spending. These data have been obtained from satellite images and have been combined with altitude

³Appendix A.1.3 discusses how we have classified democratic and non-democratic countries.

and population data at the pixel level. Each pixel is a square with a width of 30 arc seconds. The length of an arc second depends on the latitude and longitude of the pixel. At the equator, an arc second is around 30 meters, thus each pixel is around 1 km^2 .

DESCRIBE SIZE WHEN AWAY FROM EQUATOR

(I GUESS BECOME RECTANGLE LARGER THAN 1 SQ. KM)

Detailed information on these pixel-level data sources is available in Appendix A.1.2.

Our main results concern the comparison of PR versus MAJ countries, as defined by the Database of Political Institutions (DPI2012). Our sample includes 41 PR countries and 30 MAJ countries, for a total of around 7.4 million pixels in MAJ countries and 5.6 million in PR countries. We have also checked the consistency of our results using district magnitude. Appendix A.1.3 lists all the country-level data sources, while Appendix A.1.4 lists all countries considered for each continent.

Although we have also checked whether our results hold at more aggregate levels, our main level of observation is the pixel level. Cross-country regressions cannot exploit to the full the wealth of data we have. Each country, independently of its size and of its urbanization rate, count as one observation. Pixel-level regressions operate differently. Beyond providing us with a much larger number of observations, one of the advantages of changing the level of the analysis to the pixel level is to take account of these differences. For instance, India is a large country with comparatively more sparsely spread population and lower levels of light than, say, Brazil, which has higher levels of lighting on average. Pixel-level regressions can thus extract more information from the Indian pixels for lowly lit areas, and from the Brazilian pixels for the highly lit areas.

5.2 Descriptive Statistics

The box-and-whisker plots in Figure 1 provide the first descriptive statistics for the main relation of interest in this section, *i.e.* the link between light and population in PR and MAJ countries⁴. We have divided the population above the median (computed considering all PR and MAJ countries) in 10 different groups. Then, within each group and for each electoral system, we have plotted the median light (the white bar), the interquartile range (the gray rectangle), and the upper and lower adjacent lines. The advantage of this graph compared to a linear fit is that it allows us to verify whether a positive trend is due to a few outliers in highly populated pixels, or it rather reflects a general positive relationship between light and population. Moreover, it highlights in which groups the differences between PR and MAJ are more pronounced. The median is zero for the first four groups in MAJ countries, while the same is true for the first five groups in PR countries. On the other hand, highly populated pixels receive more light under PR than under MAJ.

The positive trend and the difference between PR and MAJ are amplified when we plot the same graphs for the pixels in the top quartile of the population distribution (Figures A5). However, if we exclude India, MAJ countries have lower median light in all groups (Figure A4-A6). The conclusions are also less clear-cut when looking at district magnitude instead of PR: densely populated pixels do receive more light in countries with higher district magnitude, but the relation is less well-defined when focusing on the first groups, *i.e.* on the pixels between the median and the 75th percentile of the population distribution (Figures A7-A8).

A more general overview of the correlation between light and population can be obtained by regressing light on a quadratic polynomial of population. Such quadratic fit is shown in Figure A9, while Figure A10 focuses only on low-medium densely populated pixels. Based on these raw correlations, we can see a positive relation

⁴Appendix A.2.1 provides additional descriptive statistics on the distribution of population in PR and MAJ countries.

between light and population: the curve is almost linear for MAJ countries, and **exponential** for PR countries. Moreover, the PR curve is always above the MAJ one, implying that pixels receive more light in PR than in MAJ at all population levels, even if the gap is larger among densely populated pixels.

However, these differences may be spurious and due to other factors such as GDP or geographical features. We address these concerns in the next section, but a first graphical step is to regress light on a set of control variables, compute the residuals, and plot the relation between such light residuals and pixel population. The control variables are altitude (squared), latitude (in absolute value), lagged GDP per capita (squared), whether the country is an oil producer or in war, country size, macro-districts, as well as ethnic, religious and linguistic fractionalization. Therefore, Figure 3 shows how population is linked with the portion of light which is not explained by these controls. While densely populated pixels still have higher (residual) light under PR than under MAJ, the opposite is true for low-populated pixels.

5.3 Multivariate analysis

One clear conclusion from the descriptive statistics in the previous section is that the slope of the light-population curve is steeper under PR than under MAJ. The aim of this section is to analyze this relation in more depth.

5.3.1 Econometric Framework and OLS results

Using subscript p to indicate the pixel, d the administrative district (district), and c the country, a good starting point is to regress light ($light_{pd}$) on pixel population ($lpop_{pd}$), the electoral system (PR_c), and the interaction between these two

regressors. The estimated equation is the following:

$$light_{pdc} = \beta_0 + \beta_1 lpop_{pdc} + \beta_2 PR_c + \beta_3 PR_c * lpop_{pdc} + \alpha_1 lpop_{dc} + \delta_1 lpop_c + \gamma_1 x_{pdc} + \gamma_2 x_c + \varepsilon_{pdc}, \quad (17)$$

where the main coefficient of interest is β_3 . As for Figure 3, we have controlled for pixel-level variables (x_{pdc}): altitude (squared), latitude (in absolute value). We have also included several country-level variables (x_c): lagged GDP per capita (squared), whether the country is an oil producer or in war, country size, macro-districts, as well as ethnic, religious and linguistic fractionalization. In addition to these, we have also included districtal (pop_{dc}) and national population (pop_c)⁵.

The OLS estimates are shown in the first two columns of [Table 1](#). Since data on fractionalization are not available for all countries, we have reported the estimates with (Column 2) and without (Column 1) such controls to test whether our results are robust to the change in the sample size. Moreover, in Column 2 we have also added the average light, population and altitude around the pixel to the other controls in x_{pdc} . These variables have been obtained by computing the average light, population and altitude in the 11x11 pixel square around each observation. As in all the empirical section, only pixels with population of at least 10 have been considered. To avoid simultaneity issues, the pixel itself has been excluded from the average. We have included this variable in order to control for the spatial correlation in light that can arise due to at least three factors:

1. Blurring: if pixels nearby are brightly lit, blurring in the satellite camera may increase the measured luminosity in the pixel.
2. Fixed cost and economies of scale: the cost of providing light in a pixel decreases with the light provision in neighboring cells. Moreover, the cost may be affected by nearby geographical obstacles, such as mountains.
3. Budget allocation: increasing light in the pixels nearby leaves fewer resources

⁵The Appendix A.1.5 includes a detailed description of these variables.

to increase light in the pixel itself.

Notice that this last effect goes in the opposite direction of the first two.

Before presenting out results, it is important to briefly discuss how we have constructed the standard errors. Using heteroscedasticity-robust standard errors is not enough in this case since observations are geographically linked, so the error terms are not independently distributed. We have taken this correlation into account by adding the average light, population and altitude around the pixel, but this may not be enough. For instance, all the pixels in an administrative district may be correlated because of geographical or political reason. Therefore, following Angrist and Pischke (2009), we have clustered the standard errors at the country level. The number of countries/clusters is in our case sufficiently high, greater than the number of clusters typically used in US studies (50).

Nevertheless, in our case we do not have a sample of observation: we do have data on the whole population of interest. As a consequence, as also stressed in (Abadie et al., 2017), if we were to compare the mean light between PR and MAJ countries, such difference is known with certainty, so the standard error should be zero. The common procedure in this case is to assume that there exists a data generating process - a superpopulation - from which the actual population has been drawn. As discussed in (Manski and Pepper, 2017), the issue here is that it is difficult to imagine what sampling process may be reasonable in this case. In other words, it is difficult to assume that there exists a random process which has generated the current division of the world into countries with its distribution of light and population from a set of possible (IID) alternative universes. Therefore, whether and how to actually compute standard errors in these cases is still an open question in the econometric literature. In conclusion, while we have followed the convention and we have decided to report clustered standard errors, we also offer this note of caution in interpreting them.

Both OLS specifications confirm the result from the descriptive statistics: the coefficient of the interaction between PR and pixel population is positive and significant. Therefore, the link between light and population is stronger in PR than MAJ countries. The picture that emerges from this exercise is one where proportional systems are much more responsive to the population in terms of provision of light.

5.3.2 Instrumental Variable and Fixed Effects

One concern is that the average light around the pixel is endogenous. Hence, we have instrumented the average light around the pixel with the total light in the 21x21 outer square surrounding the 11x11 square. The rationale is that those pixels are too far to directly affect the light level in pixel p , but they can affect it indirectly through the 11x11 square. As reported in Table 1 Column 3, the coefficient of the interaction between PR and pixel population remains positive and with similar magnitude.

Our results are also consistent to the inclusion of country fixed effects, thus controlling for all time-invariant country characteristics. In other words, Column 4 Table 1 reports the estimates for the following regression:

$$light_{pdc} = \beta_0 + \beta_1 lpop_{pdc} + \beta_3 PR_c * lpop_{pdc} + \alpha_1 lpop_{dc} + \gamma_1 x_{pdc} + \mu_c + \varepsilon_{pdc}, \quad (18)$$

where μ_c are the country fixed effects and we have continued to instrument average light around the pixel. The results are qualitatively and quantitatively similar to the OLS and IV results. The interaction between PR and pixel population remains positive even if the country fixed effects take into account factors common to all the pixel in a country, such as national culture, religion, language, colonial history, and social capital.

Finally, Column 5 in Table 1 replicates Column 4 using as dependent variable an

indicator equal to one if the pixel is lit, zero otherwise. The results are in line with those obtained using the continuous outcome variable: more populated pixels are more likely to be lit, and such probability increases with population more under PR than under MAJ.

5.3.3 Additional interactions

Table 2 extends Table 1 by adding the interaction between PR and district population, as well as between PR and country population when country fixed effects are not included (Columns 1-3). In addition to check whether the inclusion of such variables affects our previous conclusions, the aim of this extension is to test Proposition 1.

Indeed, the implication from that model is that in PR the key variable is the ratio between pixel population and national population, while districtal population should not matter since, unlike MAJ, the competition is at the national level, not at the district level. Therefore, since we are taking the logarithm of population, we would expect the coefficient of districtal population (\log_{dc}) to be similar in magnitude and with opposite sign to the coefficient of the interaction between PR and districtal population ($PR * \log_{dc}$). On the other hand, national population does matter in PR more than in MAJ, and a larger national population decreases the relative power of the population within a pixel. As a result, we would expect a negative sign when estimating the coefficient of the interaction between PR and country population ($PR * \log_c$).

As shown in Table 2, the interaction between PR and pixel population is positive in all specifications. This is true when looking at the OLS estimates (Columns 1-2), the IV ones (Column 3), when adding country fixed effects (Column 4), as well as when looking at the binary variable *lit* instead of the continuous *light* measure (Column 5). These results confirm our conclusions from the descriptive statistics and Table

1: PR countries are more responsive to population than MAJ countries.

As predicted, the coefficients of districtal population and its interaction with PR mirror each other in the IV and fixed effect specifications. The symmetry is not perfect, but it supports the prediction from Proposition 1. Moreover, it is worth stressing that, due to data limitations, we are using administrative districts, not electoral ones, so ours is just a proxy for the true pivotal geographical level. Finally, as expected, the interaction between PR and country population is always negative.

6 Conclusions

TBW

References

- Atkinson, A.B. 1970. “On the Measurement of Inequality.” *Journal of Economic Theory* .
- . 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.
- 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.
- Bouton, Laurent, Micael Castanheira, and Garance Genicot. 2016. “Inequalities in Government Interventions: Literature Review.” Wdr background paper, World Bank.
- 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 Tommaso Nannicini. 2011. “Competing on Good Politicians.” *American Political Science Review* 105:79–99.
- Golden, Miriam A. and Brian K. Min. 2013. “Distributive Politics Around the World.” *Annual Review of Political Science* 16:73–99.
- Harsanyi, John. 1953. “Cardinal Utility in Welfare Economics and in the Theory of Risk-Bearing.” *Journal of Political Economy* :434–35.

- . 1955. “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility.” *Journal of Political Economy* :309–21.
- 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).
- Kurtulmus, A. Faik. 2012. “Uncertainty Behind the Veil of Ignorance.” *Utilitas* 24 (01):41–62.
- 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.
- 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.
- 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.

Rawls, J. 1971. *A Theory of Justice*. Harvard paperback. Harvard University Press.

Strömberg, David. 2008. “How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida.” *American Economic Review* 98 (3):769–807.

7 Appendix:

Objective in PR:

As shown in (6), the vote share of A in district d is:

$$\pi_d^A(\mathbf{q}; \delta_d) = \sum_{l \in d} \frac{\phi_l n_l}{m_d} \cdot \left(\nu_l(\mathbf{q}, \delta_d) + \frac{1}{2\phi_l} \right) = \frac{1}{2} + \sum_{l \in d} \frac{\phi_l n_l}{m_d} (\Delta u_l(\mathbf{q}) - \delta_d).$$

Aggregating these vote shares at the national level yields:

$$\sum_d m_d \pi_d^A(\mathbf{q}; \delta_d) = \frac{1}{2} + \sum_d \sum_{l \in d} \phi_l n_l \Delta u_l(\mathbf{q}) - \sum_d \delta_d \sum_{l \in d} \phi_l n_l$$

Taking expectations over δ_d give us the expected vote share for A :

$$\begin{aligned} \pi_{PR}^A(\mathbf{q}) &= \frac{1}{2} + \sum_d \sum_{l \in d} \phi_l n_l (\Delta u_l(\mathbf{q}) - \mathbb{E}[\delta_d]) \\ &= \frac{1}{2} + \sum_l s_l (\Delta u_l(\mathbf{q}) - \beta_{d(l)}). \end{aligned}$$

where $d(l)$ is the district to which locality l belongs and $s_l = \phi_l n_l$ is the electoral sensitivity of locality l .⁶

Assumptions 1 and 2 :

Following Persson and Tabellini (1999) and Galasso and Nannicini (2011), we consider two types of districts: *contestable* and *non-contestable* districts. *Contestable* districts are such that both candidates always have a strictly positive chance of winning a majority of the votes in that district. In a *non-contestable* district, irrespective of the allocation of resource and the particular realization of the shock,

⁶For this last equality, note that $\sum_d \beta_d \sum_{l \in d} \phi_l n_l$ can be rewritten as $\sum_l \phi_l n_l \beta_{d(l)}$.

a given party is guaranteed to win a majority of the votes in the district. As we will clarify below, these districts are non-contestable because their distribution of ideological voters is sufficiently biased towards one of the two parties (*i.e.* $|b_d|$ large enough).

We denote by C the set of contestable districts, and by N the set of non-contestable districts. Naturally, districts could be neither contestable nor non-contestable. In which case, the marginal impact of allocating more local public good to a location can be increasing and the problem is non convex. We rule out these cases that would require the introduction of lotteries by making Assumption 2.

Contestable districts.

Let's define the set of contestable districts as $C \equiv \{d | p_d^A(\mathbf{q}) \in]0, 1[\forall \mathbf{q}\}$ that is:

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

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

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

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

$$\overline{\Delta U}_d + |\beta_d| \leq \frac{1}{2\gamma_d}.$$

To be contestable the variance in the district shock must be large enough compared to the bias.

Non Contestable districts.

Let's define the set of non-contestable districts as $N \equiv \{d | p_d^A(\mathbf{q}) = 0 \text{ or } p_d^A(\mathbf{q}) = 1 \forall \mathbf{q}\}$. That is, either

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

Non-contestable districts are therefore d such that

$$|\beta_d| \geq \overline{\Delta U}_d + \frac{1}{2\gamma_d}.$$

Assumption [interior].

Importantly, whether a district is contestable or not does not prevent some voters in that district to be swung. Assumption 1 actually posits that there are swingable voters in any localities, that is:

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

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

$$-\bar{\Delta} - \beta_d - \frac{1}{2\gamma_d} \geq -\frac{1}{2\phi_l} \text{ \& } \bar{\Delta} - \beta_d + \frac{1}{2\gamma_d} \leq \frac{1}{2\phi_l}.$$

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

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

Hence, Assumption 1 requires the variance in the individual preference to be large enough compare to the bias.

Objective in MAJ:

Under MAJ, seats are attributed in proportion to the number of districts won by each party. Equation (9) gives us the probability that A wins district d :

$$pr_d^A(\mathbf{q}) = Pr \left(\delta_d \leq \sum_{l \in d} \frac{s_l}{\sum_{k \in d} s_k} \Delta u_l(\mathbf{q}) \right).$$

We need to calculate the probability of that event. In contestable district there are no corner solution so that the probability of winning district d is:

$$F_{\delta_d} \left[\sum_{l \in d} \omega_l \Delta u_l(\mathbf{q}) \right] = \frac{1}{2} - \gamma_d \beta_d + \gamma_d \sum_{l \in d} \frac{s_l}{\sum_{k \in d} s_k} \Delta u_l(\mathbf{q}).$$

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

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

Proofs:

Proof of the tilt in Section ??: From the first-order conditions, we have that

$$\begin{aligned} \frac{n_l}{k'_l(q_l)} \frac{\partial u_l(\mathbf{q}^P)}{\partial q_l^P} &= \frac{n_{l'}}{k'_{l'}(q_{l'})} \frac{\partial u_{l'}(\mathbf{q}^P)}{\partial q_{l'}^P}, \forall l \neq l', \text{ or} \\ \frac{\partial u_l(\mathbf{q}^P)}{\partial q_l^P} / \frac{\partial u_{l'}(\mathbf{q}^P)}{\partial q_{l'}^P} &= \frac{n_{l'}}{n_l}. \end{aligned}$$

It implies that

$$\frac{dq_l}{dq_{l'}} = \frac{-\frac{\partial^2 u_{l'}(\mathbf{q}^P)}{\partial^2 q_{l'}^P} / \frac{\partial u_{l'}(\mathbf{q}^P)}{\partial q_{l'}^P}}{-\frac{\partial^2 u_l(\mathbf{q}^P)}{\partial^2 q_l^P} / \frac{\partial u_l(\mathbf{q}^P)}{\partial q_l^P}}.$$

Suppose that $n_{l'} > n_l$. This implies that $q_{l'} > q_l$. An increase in budget flows disproportionately towards small localities if and only if $\frac{dq_l}{dq_{l'}} > 1$. This requires that the

absolute risk aversion, *i.e.* $-\frac{\partial^2 u_l(\mathbf{q}^P)}{\partial^2 q_l^P} / \frac{\partial u_l(\mathbf{q}^P)}{\partial q_l^P}$, not to decrease in q . ■