Population Patterns and the Economic Effects of Constitutions

Laurent Bouton *,
Micael Castanheira §, Garance Genicot* and Dario Sansone *.

September 2017

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 nighlight 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 its 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 **??** 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 region (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 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 necerssarily the most populated areas in the country) ; 2. give up on localities in non swingable districts. We show that under reasonable implications 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 subregional 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 presents a model of allocation of public provision under political competitions and compare majoritarian and proportional representation systems and 3 derives theoretical predictions regarding the relationship between public good provision and population. Next, 4 tests the predicted patterns using the nightlights data. Section 5 discusses the welfare consequences of our finding and the question of how to measure inequality in public good provision. Finally Section **??** concludes.

2 A model of political competition

2.1 Setup

Consider one country with a total population of size 1. The population is partitioned into localities $l \in \{1, 2....L\}$ of size n_l. Localities belong to districts $d \in \{1, 2....D\}$. The size of the population in district d is $m_d = \sum_{l \in d} n_l$. Naturally $\sum_l n_l = 1$ and $\sum_d m_d = 1$.

There is a continuum of voters in each locality. Two parties A and B compete by promising the provision of locality-specific public goods: q_l. The cost is assumed to be weakly increasing in the population size in the group and strictly increasing in the quantity provided:

$$k_l(q_l) = k(n_l)q_l.$$

with $1 \geq k'(n_l) \geq 0$.

Two cases of particular interest:

- $k'(n_l) = 0$: the cost is then $k_l(q_l) = kq_l$ which is independent of the population size. It is an interesting case because it captures the case of a targeted public good.
- $k'(n_l) = 1$: the cost is then $k_l(q_l) = q_l n_l$, which captures the case of pure transfers (or: *publicly provided private goods* in the words of P&T).

We assume that parties maximize the expected number of seats in the parliament.

Under proportional representation (PR), districts are irrelevant. Parties maximize the expected vote share (each vote corresponds to a pre-specified fraction of a seat and we assume that seats are continuously divisible).

Under a majoritarian system (MAJ), seats are allocated on a first-past-the-post basis at the district level. Each district has one seat.

The two parties simultaneously choose their policy platforms: $\mathbf{q}^A$ and $\mathbf{q}^B$ while respecting the budget constraint:

$$\sum_l k_l(q_l) = \tau.$$

Individuals of locality l have preferences $W_l(\mathbf{q})$ for the public good, with $W_l(\mathbf{q})$

strictly increasing and strictly concave in q_l and independent of $q_{l'}$ when $l' \neq l$. $W_l(\mathbf{q}) = 0$ if $q_l = 0$.

In addition individuals have non-policy preferences. $\nu_{i,l}$ is an individual-specific parameter which measures the individual ideological bias toward party B of voter i in locality l. We assume that $\nu_{i,l} \sim U[\beta_l - \frac{1}{2\phi_l}, \beta_l + \frac{1}{2\phi_l}]$ where β_l is the non-stochastic part of locality l's bias toward B.

We denote by δ the average (relative) popularity of party B in the population as a whole. We assume that parties do not know the realization of δ but they know that $\delta \sim U[-\frac{1}{2\gamma}, \frac{1}{2\gamma}]$.

Therefore, voter i in locality l prefers A to B iff

$$\Delta W_l(\mathbf{q}) - \nu_{i,l} - \delta \geq 0,$$

where $\mathbf{q} = (\mathbf{q}^A, \mathbf{q}^B)$ and $\Delta W_l(\mathbf{q}) = W_l\left(\mathbf{q}^A\right) - W_l\left(\mathbf{q}^B\right)$.

2.2 Preliminaries

Let us identify the "swing voter" in locality l as:

$$\widetilde{\nu}_l(\mathbf{q}, \delta) = \Delta W_l(\mathbf{q}) - \delta.$$

All voters i in locality l with $\nu_{i,l} < \tilde{v}_l(\mathbf{q}, \delta)$ vote A. If interior (i.e. $\tilde{v}_l(\mathbf{q}, \delta) \in [\beta_l - \frac{1}{2\phi_l}, \beta_l + \frac{1}{2\phi_l}]$), the fraction of individuals i in locality l with $\nu_{i,l} < \tilde{v}_l(\mathbf{q}, \delta)$ is

$$\pi_l^A(\mathbf{q}; \delta) = \frac{\nu_l(\mathbf{q}, \delta) - (-\frac{1}{2\phi_l} - \psi_l)}{\frac{1}{2\phi_l} - \psi_l - (-\frac{1}{2\phi_l} - \psi_l)} = \phi_l(\tilde{v}_j(\mathbf{q}, \delta) + \frac{1}{2\phi_l} - \beta_l).$$

or

$$\pi_l^A(\mathbf{q}; \delta) = \tfrac{1}{2} + \phi_l\left[\Delta W_l(\mathbf{q}) - \delta - \beta_l\right]$$

In a proportional representation system, politicians maximize the expected vote share. Taking the expectation over δ and aggregating over the localities, the expected fraction of voters in the country voting for A is given by:

$$\pi^A(\mathbf{q}) = E_\delta \sum_l n_l \pi_l^A(\mathbf{q};\delta)$$

or

$$\pi^A(\mathbf{q}) = \sum_l n_l \left(\tfrac{1}{2} + \phi_l\left(\Delta W_l(\mathbf{q}) - \beta_l\right)\right) \tag{1}$$

In a majoritarian system, politicians maximize expected seats (or districts). Given δ, A wins district d if $\sum_{l\in d} n_l \pi_l^A(\mathbf{q};\delta) \geq \frac{m_d}{2}$. If interior,

$$\begin{gathered}\sum_{l\in d} \frac{n_l}{m_d}\left(\tfrac{1}{2} + \phi_l\left(\Delta W_l(\mathbf{q}) - \delta - \beta_l\right)\right) \geq \frac{1}{2} \\ \delta \sum_{l\in d} \frac{n_l}{m_d}\phi_l \leq \sum_{l\in d} \frac{n_l}{m_d}\phi_l\left(\Delta W_l(\mathbf{q}) - \beta_l\right).\end{gathered}$$

Hence the probability that district d votes for A (more than 1/2 of the votes in that region go to A):

$$\Pr\left(\delta \leq \bar{\delta}_d(\mathbf{q}) \equiv \sum_{l\in d} \omega_l\left(\Delta W_l(\mathbf{q}) - \beta_l\right)\right)$$

where $\omega_l = \frac{\phi_l n_l}{\sum_{k\in d \text{ s.t.} d\ni l} \phi_k n_k}$. Note that $\sum_{l\in d}\omega_l = 1$ and that if ϕ_l is constant across all l, $\omega_l = \frac{n_l}{m_d}$ when $d \ni l$.

Hence, A wins district/region r with probability $F_\delta\left[\bar{\delta}_r(\mathbf{q})\right]$ where F_δ is the cumulative distribution of the common preference shock δ. When interior, this is:

$$p_d^A(\mathbf{q}) = \frac{1}{2} + \gamma \sum_l \omega_l\left(\Delta W_l(\mathbf{q}) - \beta_l\right) \tag{2}$$

2.3 PR System

Under PR, there is one national district composed of all regions and party A maximize his expected vote share:

$$\max_{\mathbf{q}^A} \pi^A(\mathbf{q}) \ \text{ s.t. } \sum_l k_l(q_l) = \tau.$$

Using (1), the maximization becomes:

$$\max_{\mathbf{q}^A} \sum_l n_l \left(\tfrac{1}{2} + \phi_l \left(\Delta W_l(\mathbf{q}) - \beta_l\right)\right) \ \text{ s.t. } \sum_l k_l(q_l) = \tau.$$

The first order conditions are (λ^{pr} is the Lagrange multiplier associated with the budget constraint):

$$n_l \phi_l \frac{\partial W_l\left(\mathbf{q}^A\right)}{\partial q_l^A} = \lambda^{pr} k_l'(q_l), \ \forall l \tag{3}$$

If $k(n_l) = kn_l$ then parties do not take the size of the localities into account, but if $k(n_l)$ is less than proportional to n_l then parties give more to large localities (i.e. n_l large).

Note that $\mathbf{q}^A = \mathbf{q}^B$ in equilibrium (see LW and PT for formal proofs).

2.4 Majoritarian

As shown in appendix, the problem of party A is:

$$\max_{\mathbf{q}^A} \sum_d p_d^A(\mathbf{q}) \quad s.t. \quad \sum_l k_l(q_l) = \tau.$$

Plugging in (2) , the problem is as follows:

$$\max_{\mathbf{q}^A} \sum_d \left(\frac{1}{2} + \gamma \sum_{l \in d} \omega_l \left(\Delta W_l \left(\mathbf{q} \right) - \beta_l \right) \right) \quad \text{s.t. } \sum_l k_l(q_l) = \tau.$$

First order conditions are (λ^m is the Lagrange multiplier associated with the budget constraint):

$$\omega_l \gamma \frac{\partial W_l \left(\mathbf{q}^A \right)}{\partial q_l^A} = \lambda^m k_l'(q_l), \ \forall l, \tag{4}$$

3 Localities and Government Intervention

3.1 Importance of the district characteristics

Let's consider the provision of public good so that $\frac{k(n_l)}{n_l}$ is decreasing in n_l, that is the cost in the locality increases less than proportionally with the population (some element of public good).

First, we show that the relative population in the district is important under a majoritarian system.

Assume that $\phi_l = \phi \ \forall l$ (to eliminate any influence of the swingness of the different localities). Under PR, the first order conditions (3) become

$$n_l \frac{\partial W_l \left(\mathbf{q}^A \right)}{\partial q_l^A} = \frac{\lambda^{pr}}{\phi} k_l'(q_l), \ \forall l. \tag{5}$$

while under MAJ, the first order conditions (4) become

$$\frac{n_l}{m_{d(l)}} \frac{\partial W_l \left(\mathbf{q}^A \right)}{\partial q_l^A} = \frac{\lambda^m}{\gamma} k_l'(q_l), \ \forall l \tag{6}$$

where $d(l)$ is the district in which locality l is situated.

It is easy to see that PR and MAJ *do not* produce the same outcome when there is malaportionment. Since $\frac{k(n_l)}{n_l}$ decreases in n_l, politicians under PR privilege populated localities: the larger n_l the higher q_l. Under a majoritarian regime we see that the average relative size within districts, $\frac{n_l}{m_{d(l)}}$, matters. If $k_l'(q_l)$ increases in n_l then parties give more to localities with a large average relative size within districts ($\frac{n_l}{m_d}$ large) but will privilege smaller localities for the same relative size. In the case of a pure (local) public good the absolute size of the district does not matter, only the relative size.

This is summarized in the following proposition.

Proposition 1 *Under PR, the provision of public good in municipalities monotonically increases in their population. In contrast in MAJ, the provision of public good in municipalities monotonically increases in their relative population within the district.*

Example: Assume a pure local public good with a marginal cost k and $W_j(q) = V\ln(1+q)$ In this case, we get

$$\begin{aligned} \text{PR:} \qquad & q^A_{j,r} = n_j[\tau/k + L] - 1 \\ \text{MAJ:} \qquad & q^A_l = \frac{n_l}{D\ m_{d(l)}}[\tau/k + L] - 1 \end{aligned}$$

Note that the average size of a district is $1/D$.

Notice that this effect disappears if there is no malapportionment. Indeed, if all districts have the same population, $m_d = m$ for all d, then the ranking of localities within the country in terms of relative and absolute population coincide.

However, remark that the point that the characteristics of other localities in the district will matter in MAJ but not in PR is more general. For instance, assume that all localities have the same population but differ in terms of swingness ϕ_l. Then

(3) tells us that the provision of local public good is increasing in the swingness of the location. In contrast, (4) become

$$\frac{\partial W_l\left(\mathbf{q}^A\right)}{\partial q_l^A}\frac{\phi_l}{\sum_k \phi_k} = \frac{\lambda^m}{\gamma}k_l'(q_l),\ \forall l. \tag{7}$$

The same point could be done with the access to information of localities (a la Strömberg (2004)) or taste for local public goods.

What does Proposition 1 implies for the relationship between the provision of the local public good in a locality and its size?

If districts are clones of each other with a population that is either scaled up or down then we can show that the relationship between q_l and $n_l$under MAJ is flatter than under PR.

Indeed, let $n_1, ...n_{L_1}$ be the populations of the L_1 locations in district 1. Assume that for any other district d, the distribution of population into localities is given by $\chi_d n_1, \chi_d n_2 ... \chi_d n_{L_1}$ for some χ_d. Without loss of generality assume that $\chi_D \geq \chi_{D-1} \geq \chi_1 = 1$.

In the first order conditions for a majoritarian system (6), we see directly that any scaling factor χ_d cancels so that the distribution of public good provision $q_1, ...q_{L_1}$ will be identical in each district irrespective of the χs. As we increase $\frac{\chi_d}{\chi_{d-1}}$ for all d and make districts more heterogenous the relationship between q_l and n_l over the entire set of localities flattens. In contrast, in the conditions (5) for proportional representation, we see that increasing $\frac{\chi_d}{\chi_{d-1}}$ for all d and make districts more heterogenous increases the provision of public goods in the most populated districts and therefore decrease it in the less populated states. Once $\frac{\chi_d}{\chi_{d-1}} > \frac{n_{L_1}}{n_1}$ then any locality l in district $d > d'$ gets more public good than l' in d'.

Now consider any initial distribution of the population. Start from the first order conditions (5) and (6) in MAJ and PR, for $k_l(q) \equiv q$, and $W_l(\mathbf{q}^A) = W(q_l)$ and take the logarithms. Differentiating these FOCs around the equilibrium tells us how a change in population size in one locality should affect public good allocation across all districts:

$$\rho\left(q_l\right)dq_l^{pr}=\frac{dn_l}{n_l}-\frac{d\lambda^{pr}}{\lambda^{pr}},\ \forall l,\ \text{where } \rho\left(q_l\right)\equiv-\frac{W''(q_l)}{W'(q_l)}. \tag{8}$$

$$\rho\left(q_l\right)dq_l^{m}=\frac{dn_l}{n_l}-\frac{dm_d}{m_d}-\frac{d\lambda^{m}}{\lambda^{m}},\ \forall l. \tag{9}$$

Thus, in a country with a large number of localities and district, such that $\frac{d\lambda^s}{\lambda^s}\simeq 0$, we expect to observe a positive but smaller slope in majoritarian systems if and only if m_d tends to increase in n_l – which is what we would predict if we were to assign localities to districts randomly.

This intuition can be formalized exactly in the case of the CRRA utility function $q^{1-\rho}/\left(1-\rho\right)$: then, in MAJ,

$$\log q_l=\rho\left(\log n_l-\log m_l\left(n_l\right)-\log\lambda\right),\ \text{with } m_l\left(n_l\right)=\sum_{i\in d_l}n_i.$$

Thus, for a given population and district distribution, the expected observed elasticity of q_l with respect to n_l is:

$$\rho\left(1-\mathtt{E}\left(\frac{dm_l}{dn_l}\right)\right).$$

Given that the partial derivative $\partial m_l\left(n_l\right)/\partial n_l=1$, we expect this elasticity to be smaller than ρ in MAJ, instead of exactly ρ in PR.

3.2 Competitive and non competitive districts

As shown by Persson and Tabellini (1999) and Galasso and Nannicini (2011), another difference between PR and MAJ arises when some districts (the so-called *non-competitive* districts) are always supporting the same party, but the extent of the support for that party depends on the parties platforms. In such situations, under MAJ, parties abandon non-swing districts and devote all resources to the *competitive* districts (that may end up supporting one or the other party). The same is not true under PR: given that there are still voters to persuade in non-swing districts, parties allocate resources to all districts: both swing and non-swing. In this section, we formally explore how this difference affects the relation between population density and the provision of local public goods under PR and MAJ.

There are now two types of districts: the competitive and the non-competitive ones. We denote by C the set of competitive districts, and by NC the set of non-competitive districts. Importantly, it is not because a district is non-competitive that no voter in that region can be swung. Actually, the interesting case requires that there are swing voters in non-competitive districts. As we clarify in the appendix, the assumption that there are non-competitive districts with swing voters can be microfounded. It requires that (i) voters in the district are sufficiently biased towards one of the two parties (i.e. $|\beta_l|$ large enough $\forall l \in d$), and that locality-level popularity shock is more important than the national-level popularity shock (i.e. $\forall l \in d$, ϕ_l is sufficiently smaller than γ).

Given that $p_d^A(\mathbf{q}) \in \{0,1\}$ $\forall d \in NC$, we have that the problem of party P under MAJ is:

$$\max_{\mathbf{q}^P} \sum_{d \in C} p_d^P(\mathbf{q}) \quad s.t. \quad \sum_l k_l(q_l) = \tau.$$

Plugging in (2), the problem is as follows:

$$\max_{\mathbf{q}^A} \sum_{d \in C} \left(\frac{1}{2} + \gamma \sum_{l \in d} \omega_l \left(\Delta W_l(\mathbf{q}) - \beta_l \right) \right) \quad \text{s.t.} \quad \sum_l k_l(q_l) = \tau.$$

Lemma 1 follows directly:

Lemma 1 . *Under MAJ, all resources go to localities in competitive districts:* $q_l^P = 0$ $\forall l \in d \in NC$. *Resources are allocated to localities in competitive districts according* *to the following conditions:*

$$\omega_l \gamma \frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P} = \lambda^m k_l'(q_l), \forall l \in d \in C.$$

Given that there are swing voters in all districts, the problem of party P under PR remains the same as in the previous section. Resources are allocated to localities according to the following conditions:

$$n_l \phi_l \frac{\partial W_l\left(\mathbf{q}^A\right)}{\partial q_l^A} = \lambda^{pr} k_l'(q_l),\ \forall l.$$

This means that parties privilege populated localities: the larger n_l the higher q_l.

In Section 3.1, we proved that, when all districts are competitive, MAJ and PR produce the exact same outcome if (i) all district have the same population, $m_d = m\ \forall d$. and (ii) all localities have the same swingness, $\phi_l = \phi\ \forall l$. This is not true anymore when some districts are non-competitive:

Proposition 2 *If $m_d = m\ \forall d$ and $\phi_l = \phi\ \forall l$:*
(i) $\forall l \in d \in NC$, $q_l^{P,MAJ} = 0 < q_l^{P,PR}$;
(ii) $\forall l \in d \in C$, $q_l^{P,MAJ} > q_l^{P,PR} > 0$.

Proof: From Lemma 1, we have that $q_l^{P,MAJ} = 0\ \forall l \in d \in NC$.

Under PR, the first order conditions determining the allocation of local public goods among all districts are:

$$\frac{n_l}{k_l'(q_l)} \frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P} = \frac{\lambda^{PR}}{\phi\gamma}, \forall l. \tag{10}$$

Given the concavity of W_l, we have $q_l^{P,PR} > 0\ \forall l$.

Under MAJ, the first order conditions determining the allocation of local public goods among competitive districts are:

$$\frac{n_l}{k_l'(q_l)}\frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P} = m\frac{\lambda^{MAJ}}{\gamma}, \forall l \in d \in C. \tag{11}$$

We can now prove, by contradiction, that $q_l^{P,MAJ} > q_l^{P,PR}$ $\forall l \in d \in C$. Suppose that $\exists l \in d \in C$ such that $q_l^{P,MAJ} < q_l^{P,PR}$. By (10) and (11), we have that this requires $\frac{\lambda^{PR}}{\phi\gamma} < m\frac{\lambda^{MAJ}}{\gamma}$. But, then, it must be that $q_l^{P,MAJ} < q_l^{P,PR}$ $\forall l \in d \in C$. Given that $q_l^{P,MAJ} = 0$ $\forall l \in d \in NC$, we have that the budget constraint cannot be binding under both PR and MAJ, a contradiction. Therefore, it must be that $q_l^{P,MAJ} > q_l^{P,PR}$ $\forall l \in d \in C$.

Proposition 3 *Under MAJ, population size has a positive and monotonic effect on the provision of public goods in competitive-district localities, but no effect in non-competitive district localities. Under PR, population size has a positive and monotonic effect on the provision of public goods in all districts.*

The proof follows directly from Lemma 1 and the first order conditions (10) and (11).

If there is no correlation between the population size of localities in a district and its competitiveness, we can show that the relationship between q_l and n_l is flatter under MAJ than under PR. The flattening comes through two channels. First, $q_l^{P,MAJ} = 0$ $\forall l \in d \in NC$ and $q_l^{P,PR} > 0$ $\forall l$. Second, given the concavity of $W_l\left(\mathbf{q}\right)$ $\forall l$, the extra budget allocated to competitive districts under MAJ will flow more than proportionally to localities with smaller population.

What does Proposition 3 implies for the relationship between the provision of the local public good in a locality and its size?

First, note that given that $q_l^{P,MAJ} = 0$ $\forall l \in d \in NC$ and $q_l^{P,PR} > 0$ $\forall l$, the budget allocated to localities in competitive districts is higher under MAJ than under PR.

If this extra budget is divided equally among all localities in competitive districts, i.e. $q_l^{P,MAJ} = q_l^{P,PR} + \frac{\sum_{l \in d \in NC} q_l^{P,PR}}{\# l \in d \in C}$ $\forall l \in d \in C$, the relationship between q_l and n_l is necessarily flatter under MAJ than under PR. Indeed, in that case, the difference between MAJ and PR stems entirely from the localities in non-competitive districts. For those localities, the relationship between q_l and n_l is completely flat under MAJ, i.e. they receive no public good, but not under PR (monotonic and positive relationship). However, the extra budget will generally not be divided equally among all localities in competitive districts. Whether it flows towards small or large localities depends on the form of the utility function $W_l(\mathbf{q})$. In particular, it flows disproportionally towards small localities in competitive districts if and only if the absolute risk aversion does not decrease in q. (See Appendix for the proof).

If the extra budget flows disproportionally towards small localities in competitive districts, then the flattening effect of MAJ is reinforced. Indeed, on top of a completely flat relationship between q_l and n_l for localities non-competitive districts, the same relationship is flatter under MAJ than PR for localities in competitive districts: $\forall l \in d \in C$, $q_l^{P,MAJ} - q_l^{P,PR}$ (> 0) is increasing in n_l. If the extra budget flows disproportionally towards large localities in competitive districts, then the two effects go in opposite directions: the localities in non-competitive districts make the relationship between q_l and n_l flatter, but the way the localities in competitive districts make the relationship steeper. In that case, the overall effect is indeterminate.

4 Empirical

4.1 Data

In this section, we have focused on democratic countries[1]. Moreover, since the variation in night lights in developed countries is limited and may be due to other

[1] Appendix A.1.3 discusses how we have classified democratic and non-democratic countries.

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 very small in term of size and/or population. The resulting number of countries is 71.

The local public good considered is nigh 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 and population data at the pixel level. Each pixel is a square with width equals to 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 1km^2. Detail 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 list 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.

4.2 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 countries[2]. 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 compare to a linear fit is that it allows us to verify whether a positive trend is due to 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: high 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 75^{th} percentile of the population distribution (Figures A7-A8).

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

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 this raw correlations, we can see a positive relation 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 have addressed 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-regions, 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.

4.3 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.

4.3.1 Econometric Framework and OLS results

A good starting point is to regress light ($light_{pdc}$) on pixel population ($lpop_{pdc}$), electoral system (PR_c), and the interaction between these two regressors. The estimated equation is the following:

$$light_{pdc} = \beta_0+\beta_1 lpop_{pdc}+\beta_2 PR_c+\beta_3 PR_c * lpop_{pdc}+\alpha_1 lpop_{dc}+\delta_1 lpop_c+\gamma_1 x_{pdc}+\gamma_2 x_c+\varepsilon_{pdc} \tag{12}$$

Where the subscript p indicates the pixel, d the administrative district (region), and c the country. 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-regions, as well as ethnic, religious and linguistic fractionalization. In addition to these, we have also included regional (pop_{dc}) and national population (pop_c)[3].

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 as controls in $x_p dc$ the average light, population and altitude around the pixel. 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:

[3]The Appendix A.1.5 includes a detailed description of these variables.

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 affect by nearby geographical obstacles, such as mountains.

3. Budget allocation: increasing light in the pixels nearby leaves fewer resources 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.

4.3.2 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 affecting 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:

$$light_{pdc} = \beta_0 + \beta_1 lpop_{pdc} + \beta_3 PR_c * lpop_{pdc} + \alpha_1 lpop_{dc} + \gamma_1 x_{pdc} + \mu_c + \varepsilon_{pdc} \quad (13)$$

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 commons 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.

4.3.3 Additional interactions

Table 2 extends Table 1 by adding the interaction between PR and regional 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 regional 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 regional population (log_{dc}) to be similar in magnitude and with opposite sign to the coefficient of the interaction between PR and regional 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 regional 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.

5 Welfare and Inequality in Public Good Provision

The previous Sections have been devoted to the discussion of how the political system affects the relationship between provision of local public goods and population. The question is what does this mean in terms of inequality of public good provision and even how should we measure that inequality?

Key to understanding why inequality matters is to understand its impact on *welfare*. Hence, this section develops a politics-free benchmark to gain insights into the welfare question.

Following Atkinson (1970, 1983), we assume that individuals have CRRA preferences

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

The question is how to allocate public goods to the different locations l under the budget constraint that $\sum_l q_l = \tau$.

Harsyani (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).

For our purpose, the relevant dimension is location. If under the veil of ignorance any particular individual had equal probability of being born in any possible location, then this benchmark defines the average of the expected individual citizens' preferences in each location as the social social planner's objective :

$$\mathcal{W}^L = \frac{1}{L} \sum_l W_l(\mathbf{q}). \tag{14}$$

In this case, the planner's ideal would be to divide the budget equally across the different locations. Let τ^L be the smallest budget needed to reach the same level of welfare as the actual allocation of public goods:

$$\tau^L = \begin{cases} L\Pi_l(q_l)^{1/L} & \text{if } \rho = 1; \\ L^{\frac{-\rho}{1-\rho}} \sum_l (q_l{}^{1-\rho})^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases}$$

The equivalent to the Atkinson index of inequality would then be

$$A^T \equiv 1 - \frac{\tau^L}{\tau} = \begin{cases} 1 - \frac{\Pi_l (q_l)^{1/L}}{\tau/L} & \text{if } \rho = 1; \\ 1 - \frac{1}{L}\sum_l (\frac{q_l}{\tau/L}^{1-\rho})^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \tag{15}$$

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

Now equal distribution of public good is an optimum under the assumption, made in (14), that all locations are as likely as each other under the veil of ignorance. As a result highly populated and nearly inhabited areas carry the same weight. However, this interpretation of the veil of ignorance may be too extreme. A more natural assumption might be that individuals actually know the distribution of the population in the country. One may 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:

$$\mathcal{W}^P = \sum_l n_l W_l(\mathbf{q}), \tag{16}$$

and the highest level of welfare possible is given by

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

We can define the *equivalent* budget as the budget for local public goods that would be needed to give the same welfare as the current allocation of public goods as

$$\tau^P = \mathcal{W}^{*-1}(\mathcal{W}^P)$$

The resulting measure of inequality *a la* Atkison is then

$$A^P \equiv 1 - \frac{\tau^P}{\tau} = \begin{cases} 1 - \frac{1}{\tau}\Pi_l(\frac{q_l}{n_l})^{n_l} & \text{if } \rho = 1; \\ 1 - \left[\frac{\sum_l n_l \frac{q_l}{\tau}^{1-\rho}}{\left(\sum_k n_k^{\frac{1}{\rho}}\right)^{\rho}}\right]^{\frac{1}{1-\rho}} & \text{if } \rho \neq 1. \end{cases} \tag{17}$$

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 the inequalities in the population distribution but are optimal according to 16 then we say that there is no inequality in the provision of pubic goods.

Now recall our findings in Sections 3 and 4 that, compared with majoritarian systems, proportional representation regimes allocated more local public goods to highly populated areas and less to more scarcely populated areas. Clearly, measuring inequality in public good provision according to (15) would then tell us that there is more inequality in the allocation of public good provision in proportional representation systems. This is sharp contrast with the typical finding in the literature Persson and Tabellini (1999, 2000); Persson (2002); Lizzeri and Persico (2001, 2005); Milesi-Ferretti, Perotti, and Rostagno (2002), and Myerson (1993) that majoritarian systems lead to more inequality than under proportional representation. No if we use 17 to measure inequality, then the answer of which systems is more unequal is less clear.

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.

Harsyani, 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).

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. 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.

6 Appendix:

Competitive and non-competitive districts:

Following Persson and Tabellini (1999) and Galasso and Nannicini (2011), we consider two types of regions: *swing* and *non-swing* regions. Swing regions are such that if A (B) spends as much as he can on region r while B (A) spends 0, then A (B) has a strictly positive chance of winning a majority of the votes in that region. The other regions are non-swing: one of the two parties has no chance of winning. As we will clarify below, these regions are non-competitive because their distribution of ideological voters is sufficiently biased towards one of the two parties (i.e. $|b_j|$ large enough). We assume that half of the non-swing regions favor party A, while the other half favor party B. We denote by S the set of swing regions, and by C the set of non-swing regions. Importantly, it is not because a region is non-swingable that no voter in that region can be swung. Actually, the interesting case requires that there are swing voters in non-swing regions.

Let us identify the " swing voter" and the "ideology neutral" voter in locality l respectively as:

$$\widetilde{\nu}_l(\mathbf{q},\delta) = \Delta W_l\left(\mathbf{q}\right) - \delta \text{ and } \nu_l^0(\delta) = -\delta.$$

A locality l has always swing voters if $\widetilde{\nu}_l(\mathbf{q},\delta) \in [\beta_l - \frac{1}{2\phi_l}, \beta_l + \frac{1}{2\phi_l}]$ for all $\mathbf{q}$ and δ, and it always has "ideology neutral" voters if $\nu_l^0(\delta) \in [\beta_l - \frac{1}{2\phi_l}, \beta_l + \frac{1}{2\phi_l}]$ for all δ.

There are always "ideology neutral" voters if

$$-\frac{1}{2\gamma} \geq \beta_l - \frac{1}{2\phi_l} \;\;\&\;\; \frac{1}{2\gamma} \leq \beta_l + \frac{1}{2\phi_l}$$

or

$$\ni \frac{1}{2\phi_l} - \frac{1}{2\gamma} \geq \beta_l \geq -\frac{1}{2\phi_l} + \frac{1}{2\gamma}, \tag{18}$$
$$\text{or} \quad : \quad \frac{1}{2\phi_l} \geq \beta_l + \frac{1}{2\gamma}$$

which requires that ϕ_l be sufficiently smaller than γ.

Let $w_l(\tau)$ is the utility when all the resources go to locality l: $q_l = \tau$. There are always some swing voters in l if

$$-w_l(\tau) - \frac{1}{2\gamma} \geq \beta_l - \frac{1}{2\phi_l} \;\; \& \;\; w_l(\tau) + \frac{1}{2\gamma} \leq \beta_l + \frac{1}{2\phi_l}$$

or

$$\frac{1}{2\phi_l} - w_l(\tau) - \frac{1}{2\gamma} \geq \beta_l \geq -\frac{1}{2\phi_l} + w_l(\tau) + \frac{1}{2\gamma}$$
$$\text{or} \quad : \quad \frac{1}{2\phi_j} \geq \beta_l + \frac{1}{2\gamma} + w_l(\tau)$$

We are now in position to clarify the meaning of *non-swingable regions*. Let's define the set of non-swingable regions as $C \equiv \{r | p_r^A(\mathbf{q}) = 0 \text{ or } p_r^A(\mathbf{q}) = 1 \; \forall \mathbf{q}\}$.

Let $\bar{W}_{j,r}$ be the level of $W_j(\mathbf{q})$ reached when the entire budget τ is distributed to maximize $\sum_j \omega_{j,r} W_j(\mathbf{q})$. Clearly, $\sum_j \omega_{j,r} (\Delta W_j(\mathbf{q}) - \beta_j)$ is maximized (minimized) at $\sum_j \omega_{j,r} \left(\bar{W}_{j,r} - \beta_j\right)$ $\left(\sum_j \omega_{j,r} \left(-\bar{W}_{j,r} - \beta_j\right)\right)$.

Hence, $p_r^A(\mathbf{q}) = 0$ for all $\mathbf{q}$ iff

$$\sum_j \omega_{j,r} \beta_j \geq \sum_j \omega_{j,r} \bar{W}_{j,r} + \frac{1}{2\gamma}.$$

Similarly, $p_r^A(\mathbf{q}) = 1$ for all $\mathbf{q}$ iff

$$\sum\nolimits_j \omega_{j,r}\beta_j \leq -\left(\sum_j \omega_{j,r}\bar{W}_{j,r} + \frac{1}{2\gamma}\right).$$

Non-swingable regions are therefore r such that

$$\left|\sum\nolimits_j \omega_{j,r}\beta_j\right| \geq \frac{1}{2\gamma} + \sum_j \omega_{j,r}\bar{W}_{j,r}. \tag{19}$$

Conversely, a region r does not belong in C if:

$$|\sum\nolimits_j \omega_{j,r}\beta_j| < \frac{1}{2\gamma} + \sum_j \omega_{j,r}\bar{W}_{j,r}.$$

If $\beta_j = \beta\ \forall j$ then non-swing states are such that $\sum_j \omega_{j,r}\bar{W}_{j,r} \leq |\beta| - \frac{1}{2\gamma}$.

Objective in MAJ:

In a majoritarian system, there are D single-member districts. As discussed, A wins district d with probability $F_\delta\left[\bar{\delta}_d(\mathbf{q})\right]$ where F_δ is the cumulative distribution of the common preference shock δ.

Let $i(n,\mathbf{q})$ be a function that gives the index of the district with the n^{th} lowest threshold $\bar{\delta}_d(\mathbf{q})$ for $n \in \{1, 2...D\}$.[4] Party A's objective is to maximize the expected number of seats:

$$\begin{aligned}&(D)\, F_\delta(\bar{\delta}_{i(1,\mathbf{q})}(\mathbf{q})) + (D-1)\left[F_\delta(\bar{\delta}_{i(2,\mathbf{q})}(\mathbf{q})) - F_\delta(\bar{\delta}_{i(1,\mathbf{q})}(\mathbf{q}))\right] + \\ &\quad (D-2)\left[F_\delta(\bar{\delta}_{i(3,\mathbf{q})}(\mathbf{q})) - F_\delta(\bar{\delta}_{i(2,\mathbf{q})}(\mathbf{q}))\right] + ... + \left[F_\delta(\bar{\delta}_{i(D,\mathbf{q})}(\mathbf{q})) - F_\delta(\bar{\delta}_{i(D-1,\mathbf{q})}(\mathbf{q}))\right].\end{aligned}$$

[4] Assume that it picks the lowest index if the thresholds are the same for two districts.

Simplifying this expression gives us the following objective

$$\sum_{n=1}^{D} F_\delta \left[\bar{\delta}_{i(1,\mathbf{q})}(\mathbf{q})\right].$$

Since this objective in symmetric in the index, maximizing it is equivalent to maximizing

$$\sum_{r=1}^{D} F_\delta(\bar{\delta}_d(\mathbf{q})). \tag{20}$$

Hence, the choice of $\mathbf{q}_A$ by party A maximizes (20) subject to the budget constraint: $\sum_l k_l(q_l) = \tau$.

Proofs:

Proof of the tilt in Section 3.2: From the first-order conditions, we have that

$$\begin{aligned} \frac{n_l}{k_l'(q_l)} \frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P} &= \frac{n_{l'}}{k_{l'}'(q_{l'})} \frac{\partial W_{l'}\left(\mathbf{q}^P\right)}{\partial q_{l'}^P}, \forall l \neq l', \text{ or} \\ \frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P} \Big/ \frac{\partial W_{l'}\left(\mathbf{q}^P\right)}{\partial q_{l'}^P} &= \frac{n_{l'}}{n_l}. \end{aligned}$$

It implies that

$$\frac{dq_l}{dq_{l'}} = \frac{-\frac{\partial^2 W_{l'}\left(\mathbf{q}^P\right)}{\partial^2 q_{l'}^P} \Big/ \frac{\partial W_{l'}\left(\mathbf{q}^P\right)}{\partial q_{l'}^P}}{-\frac{\partial^2 W_l(\mathbf{q}^P)}{\partial^2 q_l^P} \Big/ \frac{\partial W_l(\mathbf{q}^P)}{\partial q_l^P}}.$$

Suppose that $n_{l'} > n_l$.This implies that $q_{l'} > q_l$. An increase in budget flows disproportionally towards small localities if and only if $\frac{dq_l}{dq_{l'}} > 1$. This requires that the absolute risk aversion, i.e.$-\frac{\partial^2 W_l\left(\mathbf{q}^P\right)}{\partial^2 q_l^P} \Big/ \frac{\partial W_l\left(\mathbf{q}^P\right)}{\partial q_l^P}$, not to decrease in q. ■

Main Figures and Tables

Figure 1

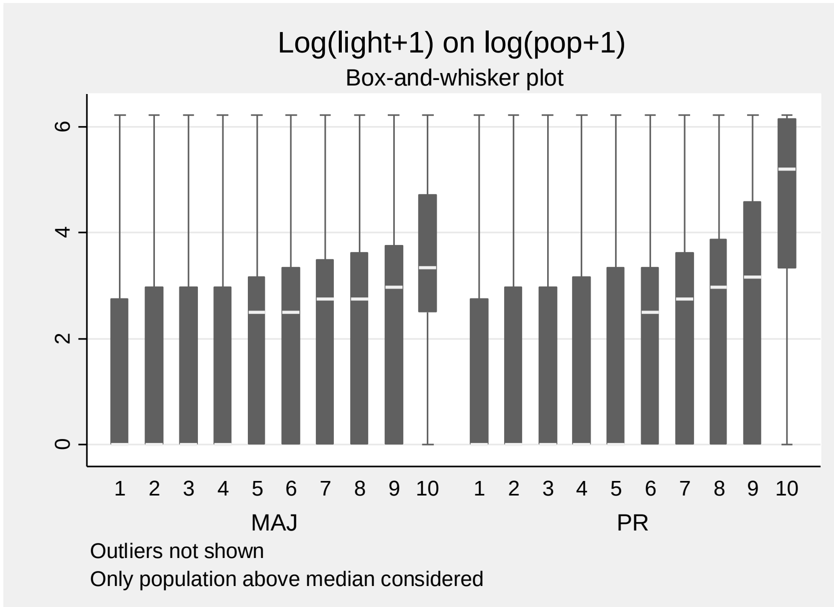

Figure 2

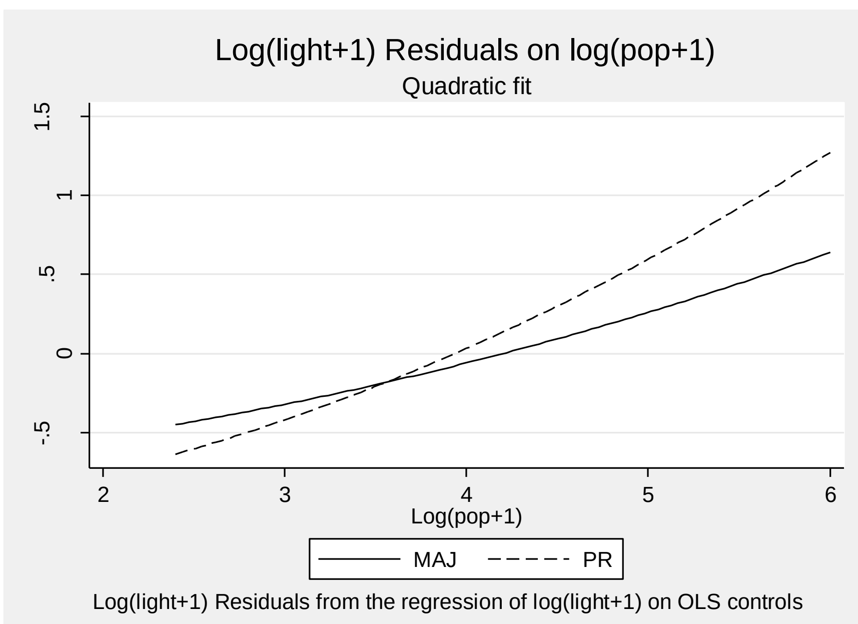

Table 1: Effect of PR on Light and Lit - OLS and FE

	(1) Light OLS	(2) Light OLS	(3) Light IV	(4) Light IV-FE	(5) Lit IV-FE
PR	-59.350**	-21.400*	-27.430*		
	(24.483)	(12.136)	(14.247)		
Pixel Pop	18.153***	4.251***	4.972***	5.049***	0.055***
	(2.734)	(0.827)	(0.991)	(1.009)	(0.007)
Regional Pop	-0.503	-0.519***	-0.589**	-0.656*	-0.002
	(1.324)	(0.151)	(0.283)	(0.382)	(0.009)
Country Pop	-8.205***	-4.738***	-5.636***		
	(2.899)	(1.346)	(1.410)		
PR*PixelPop	14.955**	5.873*	7.353*	7.896**	0.008
	(6.756)	(3.238)	(3.843)	(3.992)	(0.017)
Latitude	Yes	Yes	Yes	Yes	Yes
Altitude	Yes	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	No	No
Fractionalization	No	Yes	Yes	No	No
11x11 Pop	No	Yes	Yes	Yes	Yes
11x11 Altitude	No	Yes	Yes	Yes	Yes
11x11 Light	No	Yes	Yes	Yes	Yes
Country FE	No	No	No	Yes	Yes
N pixels	13,029,666	12,950,167	12,950,167	13,029,666	13,029,666
N countries	71	66	66	71	71
R^2	0.258	0.822	0.811	0.812	0.393
Adjusted R^2	0.258	0.822	0.811	0.812	0.393

Standard errors in parentheses clustered at country level. Only pixels with population greater than 10 have been considered. Population (Pop) is the Ln(population in the pixel+1). Only democratic countries considered. Original question for PR (DPI2012): "Which electoral rule (proportional representation or plurality) governs the election of the majority of House seats?". Country Controls: GDP per capita (lagged and squared), oil producer, war, country size, macroregions. Fractionalization: linguistic, ethnic, religious division. We have used as IV for average light around the pixel (11x11 square) the total light in the 21x21 square (excluding the total light in the 11x11 square).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effect of PR on Light and Lit - OLS and FE - Additional interactions

	(1) Light OLS	(2) Light OLS	(3) Light IV	(4) Light IV-FE	(5) Lit IV-FE
PR	-35.280	3.777	6.240		
	(49.635)	(12.967)	(16.818)		
Pixel Pop	17.980***	4.209***	4.894***	5.052***	0.055***
	(2.700)	(0.814)	(0.973)	(1.006)	(0.007)
Regional Pop	-1.433	-0.325**	-0.545	-0.740*	-0.008
	(1.582)	(0.156)	(0.336)	(0.411)	(0.009)
Country Pop	-6.476**	-4.409***	-5.019***		
	(3.006)	(1.198)	(1.239)		
PR*PixelPop	15.495**	5.987*	7.580*	7.891**	0.008
	(6.773)	(3.306)	(3.912)	(3.989)	(0.017)
PR*RegPop	3.569	-0.324	0.226	0.207	0.013
	(2.548)	(0.406)	(0.532)	(0.783)	(0.018)
PR*CountryPop	-5.118**	-1.087	-2.158**		
	(1.955)	(0.929)	(1.073)		
Latitude	Yes	Yes	Yes	Yes	Yes
Altitude	Yes	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	No	No
Fractionalization	No	Yes	Yes	No	No
11x11 Pop	No	Yes	Yes	Yes	Yes
11x11 Altitude	No	Yes	Yes	Yes	Yes
11x11 Light	No	Yes	Yes	Yes	Yes
Country FE	No	No	No	Yes	Yes
N_pixels	13,029,666	12,950,167	12,950,167	13,029,666	13,029,666
N_countries	71	66	66	71	71
R^2	0.260	0.822	0.811	0.812	0.393
AdjR^2	0.260	0.822	0.811	0.812	0.393

Standard errors in parentheses clustered at country level. Only pixels with population greater than 10 have been considered. Population (Pop) is the Ln(population in the pixel+1). Only democratic countries considered. Original question for PR (DPI2012): "Which electoral rule (proportional representation or plurality) governs the election of the majority of House seats?". Country Controls: GDP per capita (lagged and squared), oil producer, war, country size, macroregions. Fractionalization: linguistic, ethnic, religious division. We have used as IV for average light around the pixel (11x11 square) the total light in the 21x21 square (excluding the total light in the 11x11 square).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$