Group Size and Political Representation Under Alternate Electoral Systems*

Sugat Chaturvedi
Indian Statistical Institute

Sabyasachi Das
Ashoka University

March, 2019

Abstract

We examine the effect of group size of minorities on their representation in national government under majoritarian (MR) and proportional (PR) electoral systems. We provide a theoretical framework that models spatial distribution of *multiple* minority groups. It predicts that a minority's population share has *no effect* on its representation and per capita resource allocation under PR, but has an *inverted-U* shaped relation under MR. We compile an ethnicity level panel dataset comprising 87 democracies that remarkably exhibits the same causal relationship for political representation and resource allocation. The results show how electoral systems can starkly affect inter-group inequality.

JEL Classification: D72, D78, H11, J15

Keywords: Electoral systems, minorities, political representation, settlement patterns

*Chaturvedi: Economics and Planning Unit, Indian Statistical Institute, Delhi, India (e-mail: sugatc15r@isid.ac.in); Das: Economics Department, Ashoka University, Delhi, India (e-mail: sabyasachi.das@ashoka.edu.in). We thank Helios Herrera, Ebonya Washington, Nikhar Gaikwad, Sourav Bhattacharya, Rohini Somanathan, Izaskun Zuazu, Diane Coffey, Dean Spears, Abhiroop Mukhopadhyay, Bharat Ramaswami, Tridip Ray along with seminar and conference participants at the European Winter Meeting of the Econometric Society (Naples, 2018), Royal Economic Society Annual Conference (Sussex, 2018), NEUDC (Cornell, 2018), the 3^{rd} Conference on Political Economy of Dictatorship and Democracy (Münster, 2019), the 13th Annual Conference on Growth and Development (ISI Delhi, 2017), IIM-Calcutta, DSE, South Asian University for comments and suggestions.

1 Introduction

Representation of ethnic groups in democratic governments is an important determinant of their welfare. This is especially true for smaller groups or "minorities" as they are potentially more vulnerable to exclusion.[1] Political representation of these groups ensures that they can voice their interests and desires to the government, which makes policymaking more inclusive and improves allocation of public resources towards them.[2]

However, our data show that on average, only a third of these groups get any representation in the national executive of democracies during the post World War II period. In contrast, the "majority" group is almost always represented.[3] Moreover, representation of minorities is highly unequal both within as well as across democracies. In this context it may be useful to ask whether different electoral institutions provide different incentives for political parties to represent some minorities but not others, and to what extent the size of a group matters for this differential incentive.

In this paper, we examine this issue first theoretically and then empirically by looking at how population share of minorities affects their representation in the national government, and how this relationship depends on the electoral system. We focus on two broad categories of electoral systems—majoritarian (MR), where elections are typically contested over single member districts, and proportional representation (PR), where seats are allocated to parties in proportion to their vote share in multimember districts.

To contrast PR and MR elections, we propose a model with three groups (i.e., one majority and two minorities) and two parties in a probabilistic voting setup. In our model, political parties compete for votes from all groups

[1]For ease of exposition, we refer to all the groups that are not the largest group of a country to be "minorities" and define the largest group to be the "majority."

[2]Previous works, for example, show that representation fosters trust and approval in government decision-making (Banducci, Donovan and Karp, 2004), engenders greater political participation among minority group's members (Bobo and Gilliam, 1990), and consequently, improves allocation of public resources towards them (see Cascio and Washington, 2013 for the case of African Americans in the US and Besley, Pande and Rao, 2004, 2007 etc., for the case of minority caste and tribe groups in Indian village governments).

[3]This is based on the dataset that we compile for this paper, which covers more than 400 ethnicities across 87 countries for the period 1946-2013. We discuss this later in greater detail.

and promise representation in the government to each group as platforms. Representation determines the per capita private transfer of government resources targeted towards group members. We look at equilibrium representation (and consequently, per capita resource allocation) for various population share compositions of the two minorities (keeping the majority group's size fixed). We show that under PR, group size of minorities has *no effect* on their representation in the national government, whereas it has an *inverted-U shaped* effect under MR. Therefore, under MR there is an "optimal" minority size above which its representation and per capita resource allocation begin to fall.[4]

The result is in contrast to the theoretical predictions of Trebbi, Aghion and Alesina (2008) who study a similar problem in the context of US municipalities following the Voting Rights Act, 1965. They model the representation of two groups—the white majority and the black minority in US cities and compare the welfare levels across the two electoral systems for minority of varying size.[5] In their model, access to power for the minority never falls with its population share within any electoral system. In PR, it first remains unresponsive to population share and then increases eventually, and in MR, it first increases and then becomes unresponsive as minority's share becomes large enough. We show that this gets modified when we allow political parties to compete for all groups in a multiple minority context.[6] Since we are concerned with representation in national governments, the assumption of multiple minorities seems reasonable.[7] Further, in most countries the major national parties do attempt to court mul-

[4]The majority group under both systems gets higher representation and larger per capita resource allocation compared to both the minorities.

[5]The paper explains the *choice* of electoral system by the incumbent whites after the effective enfranchisement of black population in the southern US municipalities. We, on the other hand, examine how minorities of differing sizes fare under a given electoral system.

[6]Trebbi, Aghion and Alesina (2008) abstract away from parties and assume that group members always vote for the candidate belonging to their own group. One way to conceptualize their assumption in terms of our model would be to say that, in their model voters are partisan and voters' ethnic identity and party support are aligned—all whites vote for one party and all blacks for the other. In our model we relax this assumption.

[7]In the context involving three or more groups, the majority group, the way we have defined it, may not have absolute majority in the population. This however is not an important issue for us. More than 80 percent of the majority groups in our dataset indeed have absolute majority in their respective countries. Our results, both theoretical as well as empirical, do not change if we restrict attention to countries where the largest group has absolute majority.

tiple groups, even though certain groups may have inclinations towards specific parties. Therefore, our assumption of parties attempting to get votes from all groups, we believe, is not an unreasonable way to model national electoral politics. As we discuss later, this plays a crucial role in shaping the nature of the equilibrium representations in our model.

The result we get for the PR system is a straight forward implication of the standard probabilistic voting model with multiple groups. In a PR system parties essentially maximize votes. Now, there are two opposing forces in action that result in a group's representation being unresponsive to its population share. Consider two minority groups with one being larger than the other. Though offering higher representation (and hence, per capita transfers) to the larger group gets a party more total votes, it is cheaper for a party to attract a higher *share* of voters from the smaller group. When representations are equal, these two forces balance each other out across groups.

In MR, on the other hand, parties want to win constituencies and hence, they have to consider settlement patterns of groups across constituencies, i.e., over space. Our equilibrium characterization result shows that representations of groups depends on how exactly they are distributed across constituencies. However, for any given population shares of groups at the national level, there are a lot of possible ways they can be distributed over space. Therefore, deriving general results regarding comparative statics can be hard. We propose a parsimonious framework that models the spatial distribution of groups across an arbitrary number of constituencies to address this issue.[8] We use insights from the *settlement scaling theory* (Bettencourt, 2013) in the literature on urban geography to map population share of a group at the national level to its settlement area over space. We postulate that the area occupied by a minority group has a *concave* relationship with its population share. Intuitively, if the benefit of living in an area is increasing in the density of own group members living in the area (due to positive network effects), then we should observe that larger groups

[8]Milesi-Ferretti, Perotti and Rostagno (2002) also model MR and PR systems with 3 or more groups with general number of constituencies. However, they assume that the population composition of groups are identical across all constituencies. Our model relaxes this assumption to allow the group shares to be different across constituencies.

live more densely, giving rise to the concave relationship.[9] We take the concave relation as exogenously given. Majority groups are assumed to be present in all parts of the country. This allows us to characterize the groups' distribution across constituencies by imposing minimal structure on the problem. This implies that for the majority group with a given population share, if the minority groups are unequal in their population shares, they in aggregate would occupy less area than if they were all equally sized. Therefore, if minority groups are too unequal in size (i.e., say, one "too small" and one "too large"), they *both* suffer a geographical disadvantage against the majority group in MR. Parties respond to this by diminishing the promised representations to both minorities.[10] This is at the core of the inverted-U shaped result that emerges as the equilibrium in our model.

In the empirical section, we show evidence in favor of our concavity assumption and then test the comparative static results that the model delivers using measures of political representation as well as per capita resource allocation. We compile an ethnicity level panel dataset comprising 421 minorities across 87 countries for 1946–2013 by triangulating various sources that we describe in section 4.1. Our main measure of political representation is an indicator that takes value one if a group has *any* representation in the national government and zero if it is either powerless or discriminated by the state. The indicator, therefore, captures the extensive margin of political representation. We use nightlight luminosity per unit area for each group calculated using GIS maps of settlement areas of ethnic groups as a proxy for allocation of public resources by the government. As we argue in section 7.2, existing evidence shows that nightlight luminosity is highly correlated with provision of electricity—a publicly provided good which is subject to political influence—as well as the provision of other public goods.

[9] The theoretical model in Bettencourt (2013) generates a further prediction that the elasticity of the relationship between area of settlement and population should be 0.67. We get an estimate of 0.63 for elasticity in our data and which is statistically indistinguishable from 0.67. We elaborate on this and discuss empirical evidence in favor of this estimate from other contexts in sections 2.3 and 7.1, respectively.

[10] In Trebbi, Aghion and Alesina (2008) this equilibrium response by parties does not occur as the voters are partisan by design, and therefore, can not be induced by changing the platforms. This is why in their model when a minority size becomes large, the representation becomes constant under the MR system.

Further, we use cross-sectional data on road length per unit area for ethnic groups as an alternative measure of public resource allocation to show robustness of our results.

In our empirical analysis, we compare groups *within* a country-year by using fixed effects for country-year pairs.[11] This is a strong specification which controls for a host of time invariant as well as time varying country level observable and unobservable factors that may affect the relation. Consistent with the predictions of our theoretical model, the result shows a statistically significant inverted-U shaped relationship between population share and political inclusion under MR and no relationship under PR. The predicted "optimum" population share for minorities in MR countries is estimated to be 0.26. Importantly, the result is replicated with logarithm of nightlight emissions per unit area in the settlement area of a group as the dependent variable.

However, the electoral system of a country is not exogenous. Political actors in positions of power may strategically choose electoral systems that maximize their chances of winning, as Boix (1999) and Trebbi, Aghion and Alesina (2008) show. This means that the electoral system at the time of democratization of a country, and even changes in it later may depend on existing power distribution across groups (Colomer, 2004; Persson and Tabellini, 2003). We address this endogeneity issue by looking at a subsample of erstwhile colonies. Consistent with Reynolds, Reilly and Ellis (2008), we show that electoral system of the former colonial rulers systematically predicts electoral system of the colonies post-independence. We, therefore, use this as an instrument for the electoral system of a colony.[12] The two-stage-least-squares estimates replicate our results for both political representation *and* nightlight luminosity.

Remarkably, the result also holds up when we compare *same* group present in more than one country within a continent and exploit the plausibly exogenous variation in its population share across countries. In this strategy, the varia-

[11]The analysis, therefore, only considers countries with multiple minority groups and is consistent with our modeling assumption.

[12]We restrict our sample to colonies which democratized not too long after independence. We use a maximum lag of 30 and 50 years between independence and democratization for our analysis. We do this to improve the predictive power of the first stage. See Sections 6 and 7.3 for a detailed discussion about this.

tion comes primarily from a group falling unequally on two sides of the national boundary.[13] This strategy heavily restricts our sample and consequently, our sample size falls by more than 80 percent. Even in the reduced sample, we find a statistically significant inverted-U shaped relation in MR and no relation in PR for political representation. Importantly, the nightlight regression mirrors the pattern observed for political representation. The coefficients for the nightlight regression are, however, imprecise presumably due to small sample size. As a robustness check, we show that the patterns remain similar when we use road length per km^2 in the settlement area of an ethnic group as our alternate measure of public resource allocation. Since the road data is cross-sectional, we have fewer observations and therefore, the results are not as precise. However, the pattern holds up both under the baseline specification and the IV strategy. Additionally, in appendix section E we show that, consistent with the model's predictions, the inverted-U shape relation in MR countries is driven by groups which are geographically concentrated.[14]

Our work is related to the large literature examining the effect of electoral systems on public policy and other political outcomes. Myerson (1999) and Persson and Tabellini (2002) discuss and extensively review the literature on theoretical aspects of electoral systems. Some of the outcome variables that have been studied with regard to effects of electoral systems are public goods provision and redistribution (Lizzeri and Persico, 2001; Milesi-Ferretti, Perotti and Rostagno, 2002; Persson and Tabellini, 2004), trade policy (Rickard, 2012b), corruption (Kunicova and Ackerman, 2005), public attitude towards democracy (Banducci, Donovan and Karp, 1999), voter turnout (Herrera, Morelli and Palfrey, 2014; Kartal, 2014), and incentive to engage in conflict (Fjelde and Hoglund, 2014).

Some papers such as Moser (2008) and Wagner (2014) have compared differences in the level of minority representation across the two systems by exploiting the variation in electoral systems over space and time in specific countries

[13] Dimico (2016) uses a similar identification strategy to identify the effect of group size on its level of economic performance in the African continent.

[14] This result is also consistent with the papers that argue that geographic concentration of groups matters for policies under MR system (See Rickard (2012a), Moser (2008), Wagner (2014)).

(Russia and Macedonia, respectively). In both cases the authors argue that settlement pattern of minorities is an important factor to consider when analyzing change in minority representation across electoral systems. Our analysis also highlights this concern and points out the exact nature of this influence, both theoretically and empirically. Moreover, while these papers look at the level of power enjoyed by minorities, our paper focuses on difference in the *slope* of the relationship between group size and political power across electoral systems. This allows us to look at differential access to power received by minorities of differing sizes *within* a system.[15] Our result has important implications for power inequality between minorities. It suggests that PR distributes power more equally across minority groups, and hence, their (per capita) resource inequality is also minimal. The implication for inequality in the MR system is more nuanced. Our result suggests that small and large minorities might enjoy similar level of power and material well-being in MR countries while the mid-sized groups enjoy a greater access to and benefit from the government.

The rest of the paper is organized as follows: section 2 develops the model and generates testable predictions, section 3 discusses the two electoral institutions that we consider and their relevance for government formation, section 4 elaborates on the various datasets used and summarizes the main variables, section 5 explains the empirical methodology, section 6 describes the identification strategy, and section 7 discusses the results. We show robustness of our empirical results in section 8 and finally, section 9 concludes.

2 Model

2.1 Set Up

We develop a probabilistic voting model with two parties *à la* Persson and Tabellini (2002). There are three groups of voters. Each group has a continuum of voters of mass n_j with $\sum_{j=1}^{3} n_j = 1$. We will treat group 3 as the majority

[15] In order to compare a group of a given size across two systems, we need to know about the *intercept* along with the slope. The focus of our model and empirics is, however, mostly on the slope. We briefly discuss in the empirical section about the intercept estimate and its implication.

group and groups 1 and 2 as the minorities. Therefore, $n_3 \in (0.33, 1)$. Voters have preferences over private transfers made by the government. These transfers can be targeted at the level of groups but not at the individual level. We represent individual preference of any voter in group j by the utility function:

$$U(f_j) \;=\; \ln f_j$$

where f_j denotes per capita private transfers to group j. Our assumption of the specific utility function ensures that the utility function is strictly increasing, strictly concave, and satisfies the Inada conditions—requirements which are necessary for our results to hold. f_j is completely determined by the political processes of a country. Before election takes place, the two political parties A and B simultaneously announce the group composition of the government that they will form in the event of an election win. Therefore, we can define group j's representation in the government promised by party h, G_j^h, as simply the total number of government positions announced by party h in favor of group j. G_j^h, determines how much per capita transfer voters of group j will get if party h comes to power. We denote this as follows:

$$f_j^h = f(G_j^h) \quad \text{or} \quad G_j^h = f^{-1}(f_j^h).$$

More representation in government is always beneficial for group members, i.e., $f'(G_j^h) > 0$. Since representation in government determines the individual level payoff of the voters, the political parties commit to allocation of government positions as their platforms during the election. In the following analysis, we use f_j^h directly as a choice variable of the parties, since representation in government (G_j^h) and per capita transfer (f_j^h) are synonymous in our model. Any voter i belonging to group j votes for party A if:

$$U(f_j^A) > U(f_j^B) + \delta + \sigma_{i,j}$$

where $\delta \sim U[\frac{-1}{2\psi}, \frac{1}{2\psi}]$ and $\sigma_{i,j} \sim U[\frac{-1}{2\phi_j}, \frac{1}{2\phi_j}]$ are preference shocks to the voter.

This is a standard probabilistic voting set up where δ can be interpreted as population wide wave of support in favor of party B (relative to A). $\sigma_{i,j}$ represents (relative) ideological bias of a member i of group j towards party B. ϕ_j is the height of the p.d.f. of the $\sigma_{i,j}$ distribution. It measures the responsiveness of group j voters to private transfers by parties. A larger value of ϕ_j would imply that for the same increase in promised per capita transfer by any party, a greater proportion of group j voters would sway in favor of that party. For simplicity, we assume that minority groups 1 and 2 are identical in their political responsiveness to transfers, i.e., $\phi_1 = \phi_2 = \phi$. Group 3 (the majority group) is more responsive to transfers compared to the minorities, i.e., $\phi_3 > \phi$. This assumption is motivated by the observation that the minorities often have stronger attachments to specific parties owing to historical factors. Consequently, this makes them less pliable compared to the majority group from the parties' point of view.[16] Values of ψ and ϕ_j are known to both the parties. The government has a total budget which is exogenously fixed at S. Each party h maximizes the probability of forming government p_h by choosing f_j^h subject to the budget constraint:

$$\sum_{j=1}^{3} n_j f_j^h \leq S$$

In proportional system p_h is the probability that vote share is larger than 0.5, while in the majoritarian system it is the probability of winning more than half of the electoral districts. We assume that in majoritarian system there are K equally sized electoral districts of population $\frac{1}{K}$ each. We denote by n_j^k the population share of group j relative to population in district k. Therefore,

$$\sum_{j=1}^{3} n_j^k \; = \; 1 \quad \text{for all } k = 1, 2, \ldots, K$$

$$\text{and} \quad \frac{1}{K} \sum_{k=1}^{K} n_j^k \; = \; n_j \quad \text{for } j = 1, 2, 3.$$

[16] Trebbi, Aghion and Alesina (2008), who also develop a model of group based representations under the two electoral systems, assume that candidates can not sway voters in their favor by announcing higher transfers, i.e., voters are not pliable at all. We relax this assumption by allowing the parties to compete for all voting groups, and introduce heterogeneity in their degree of pliability depending on the nature of the groups.

We compare equilibrium political representation in single district PR system with that in K district MR voting system.

2.2 Equilibrium Characterization

Since the parties are symmetric, we have policy convergence in equilibrium, i.e., both parties choose the same equilibrium policy in any system. The following two propositions characterize the equilibrium allocation of resources (and hence, equilibrium representation) under the two systems.

Proposition 1 *Under a single district proportional representation voting system, group size n_j of a minority has no effect on equilibrium representation G_j^* and equilibrium transfer f_j^*. In equilibrium:*

$$\phi_j U'(f_j^*) = \phi_l U'(f_l^*) \quad \forall\, j \neq l. \tag{1}$$

We relegate all proofs to appendix section H. Proposition 1 implies that under PR, minority groups 1 and 2 would receive identical per capita transfers irrespective of their population shares, i.e., $f_1^* = f_2^*$ for all n_1 and n_2. To understand the result intuitively, let's consider the case where group 1 is the larger minority, i.e., $n_1 > n_2$. Suppose that f_1 and f_2 are the initial transfers promised by any party. Further, consider the party taking away $\epsilon > 0$ per capita transfer from group 1 and reallocating it to group 2. The per capita transfer of group 2, therefore, would increase by $\frac{n_1\epsilon}{n_2} > \epsilon$. This highlights the fact that it is always cheaper to increase per capita transfer of the smaller group. This reallocation, for a small ϵ, would cost the party $n_1\phi U^{'}(f_1)$ votes from group 1 and would increase votes from group 2 by $n_2\phi U^{'}(f_2)\frac{n_1}{n_2}$. Since in PR the political parties maximize votes, the party would prefer to reallocate as long as the gain and the loss from reallocation are different. It is obvious that when $f_1 = f_2$, they equalize. Therefore, even though vote *shares* of the smaller group are cheaper to buy, the return to a party for doing this (in terms of *total* votes) is lower, precisely because the group is small. These two opposing forces balance each other out in equilibrium, giving us the result.

Moreover, we get that the majority group gets higher per capita transfer compared to minorities, i.e., $f_3^* > f_1^* = f_2^*$. This is a direct result of our assumption that majority group voters are easier to sway through electoral commitments and hence, parties compete more fiercely for their votes.

The following result characterizes the equilibrium transfers in MR:

Proposition 2 *Under the majoritarian voting system with K districts, the following set of equations characterizes the equilibrium transfers (f_1^*, f_2^*, f_3^*) announced by both parties:*

$$\phi_j U'(f_j^*) \sum_{k=1}^{K} \frac{n_j^k / n_j}{\sum\limits_{j'=1}^{3} \phi_{j'} n_{j'}^k} = \phi_l U'(f_l^*) \sum_{k=1}^{K} \frac{n_l^k / n_l}{\sum\limits_{j'=1}^{3} \phi_{j'} n_{j'}^k} \quad \forall\, j \neq l \tag{2}$$

This is a general characterization result for any arbitrary distribution of groups across constituencies. We emphasize two aspects of the result above. Firstly, the characterization implies that equilibrium representation and transfer to groups under MR depends on the population shares. Importantly, the transfer also depends on distribution of groups across electoral districts, suggesting that *settlement patterns* of groups across districts or over space are important in determining the exact nature of the relation between group size and transfers. Moreover, if all groups have the same responsiveness to transfers, i.e., if $\phi_1 = \phi_2 = \phi_3$, then equation (2) collapses to equation (1). Therefore, heterogeneity in responsiveness across groups, especially across majority and minority groups is critical for group size to matter in MR system.

To explore this issue a little further we rewrite equation (2) as the following:

$$\phi_j U'(f_j^*) \frac{\sum_{k=1}^{K} \omega^k n_j^k}{n_j} = \phi_l U'(f_l^*) \frac{\sum_{k=1}^{K} \omega^k n_l^k}{n_l} \quad \text{where } \omega^k = [\sum_{j'=1}^{3} \phi_{j'} n_{j'}^k]^{-1}.$$

ω^k is therefore the inverse of the average responsiveness of district k, and $\sum_{k=1}^{K} \omega^k n_j^k$ is the weighted average of the group j's shares across districts with ω^k as the weights. Thus, the proposition above states that in MR, a group will get higher political representation and private transfers relative to another group if it is concentrated more in districts having a less responsive mass of voters, i.e., if the group has a higher correlation between n_j^k and ω^k. Since the majority group is more responsive, it therefore follows that a minority group would gain if it is concentrated more in districts with low majority group population. This happens because parties in MR wish to win electoral districts (as opposed to votes). Therefore, if a minority is settled in districts where the majority group is relatively scarce, the group becomes attractive to the political parties for the purposes of winning those districts. This logic plays an important role in determining the nature of the comparative static exercise we perform in the following section.

2.3 Spatial Distribution of Groups and Comparative Statics

In this section, we study equilibrium representation and transfers in MR for minorities of differing group sizes. Specifically, we see how equilibrium outcomes change when we change n_1 and n_2, keeping the majority population share n_3 fixed. Our comparative static exercise, therefore, looks at the effect of changing n_1, holding n_3 constant. Now, any change in the composition of population shares of minorities at the national level changes their distribution across districts, i.e., the values of n_1^k and n_2^k for all k. Therefore, even though proposition 2 characterizes the equilibrium for any given profile of population shares of groups, it is hard to comment on the nature of comparative static result without specifying how changes in the population shares of groups lead to consequent change in their spatial distribution across electoral districts. Below we provide a framework to incorporate this concern in our model.

We first normalize the total area of the country to 1. We denote by A_j the measure of the area where group j has presence and postulate that $A_j = n_j^{\alpha_j}$ for some $\alpha_j \geq 0$.[17] We assume that for group 3 (i.e., majority group) $\alpha_3 = 0$, or $A_3 = 1$, i.e., the majority group is dispersed over all the space in the country. For the groups 1 and 2, we consider two possibilities. In one case, we assume

[17] Note that the same space can have presence of multiple groups, and therefore, $\sum_{j=1}^{3} A_j$ need not be one. If groups overlap over space, $\sum_{j=1}^{3} A_j$ would in fact be larger than one.

$\alpha_1 = \alpha_2 = \alpha > 0$, i.e., both minorities are geographically concentrated in some region of the country. In the alternative scenario we allow group 2 to be dispersed and group 1 to be concentrated, i.e., $\alpha_1 = \alpha$ and $\alpha_2 = 0$.[18]

Importantly, we take $\alpha < 1$ for groups that are geographically concentrated, i.e., the area of settlement of a group has a *concave* relationship with its population share. This assumption is motivated by the insight from the literature on urban geography. Specifically, Bettencourt (2013) provides a parsimonious theoretical framework to predict the relationship between population and area of settlement (and other characteristics of the population, such as network length, interactions per capita etc) in the context of cities. He argues that the benefit of living in a city is increasing in the population density of the area. This would be true because for the same distance travelled, an individual will have larger number of productive interactions with people. On the other hand, the cost of living is increasing in the diameter of the city, i.e., it is proportional to the square root of the area. The city size is in equilibrium when the benefit and cost are equalized. The equilibrium relationship is therefore given by $A = c_0 n^{\frac{2}{3}}$, for some constant c_0. Bettencourt (2013), therefore, provides a theoretical prediction of the elasticity of the relationship. He further shows that for a sample of cities in the USA, the prediction is indeed valid. We assume that the concave relationship holds in the context of settlement of ethnic groups as well, since the basic forces highlighted by Bettencourt (2013) should be at play in our context as well.[19] This assumption will turn out to be important for the result we derive below.

Now we consider dividing the country in K equally sized electoral districts. Note that in the case where both minorities are geographically concentrated, we have three types of districts: (i) group 3 is present with only one minority group in the district, (ii) all the three groups are present, and (iii) only group 3 is present. The last type of district will not be there if group 2 is also dispersed. For us the most important type of district is the one where all groups are present.

[18] If all groups are dispersed then the population distribution of groups in the country is replicated in each of the districts individually and consequently, the result for MR collapses again to the PR case.

[19] Subsequent to the findings of Bettencourt (2013), several papers show that the relationship holds true in other contexts as well. We also estimate the value of α in our data and, surprisingly, find the same result. We discuss this in section 7.1.

Since the majority group is present everywhere, the proportion of this type of district is determined by the overlap region of the settlement areas of the two minorities. We denote by $A_{1\cap 2}$ the measure of the area where groups 1 and 2 overlap and correspondingly we define the overlap coefficient (also known as the Szymkiewicz-Simpson coefficient) as:

$$O = \frac{A_{1\cap 2}}{\min\{n_1^\alpha, n_2^\alpha\}}$$

We, therefore, have $O \in [0, 1]$. With these objects defined, we state the main result that establishes the relationship between group size and political representation for minorities in MR systems.

Proposition 3 *We state the results separately for the two cases that we consider:*

1. *If group 2 is also concentrated, then G_1^* follows an inverted-U shaped relation with n_1 with the peak of political representation at $n_1^* = \frac{(1-n_3)}{2}$ if and only if $O > O^*$ for some $O^* \in (0, 1)$.*

2. *If group 2 is geographically dispersed, equilibrium political representation of group 1, G_1^*, follows an inverted-U shaped relation with n_1 with the peak of political representation at $n_1^* = (1-\alpha)^{\frac{1}{\alpha}}$.*

The result implies that when both groups are concentrated, the equilibrium representation of (and consequently, transfers to) both groups have an inverted-U shaped relationship with group size. The intuition behind this result follows from the discussion of proposition 2. Our assumption about concave relationship between group population share and area occupied implies that the total area occupied by the two minorities together would be largest if they are equal sized (i.e., $n_1 = n_2 = \frac{(1-n_3)}{2}$). As their population shares diverge from each other, i.e., as one becomes larger and the other smaller, their total settlement area would fall. Now consider the type of electoral districts where all groups are present (the type (ii) district, as mentioned above). Divergence in the population shares of minorities away from the "mid-size" would imply that in those districts the *relative* share of the majority group would go up, since this is the only type of district where all groups are present. This, according to the discussion above, harms both minorities, as they become concentrated in the districts with larger (relative) majority share. The minority group which is getting smaller, therefore, loses out in both types (i) and (ii) of districts. The group which is getting larger faces opposing forces on its representation. It becomes more important in type (i) districts, but less important in type (ii) districts. Therefore, overall increase in population share would harm the group if most of its population is settled in the type (ii) districts, i.e., if the overlap coefficient is high enough.[20]

An alternative way to think about it is to notice the fact that the concave relationship between population share and area occupied implies that larger minorities, on average, have higher population density than smaller ones. For minorities which are not dispersed through out the country, there is an "optimal" density that maximizes their presence across districts. If a minority is too dispersed, they become less important everywhere. If they become too concentrated, their importance remains clustered around few districts only. Our model shows that the large minorities suffer from the latter problem by becoming "too large" in type (i) districts and "too small" in type (ii) districts.

3 Electoral Systems and Government Formation

The decline of colonialism and autocratic rule, and a transition towards democracy has characterized the world in the post World War II period. An interesting aspect of this wave of democratization is the choice of electoral system made by the newly emerging democracies. On one hand, we have MR in which elections are typically contested over single member districts. The candidate or party with a plurality or an absolute majority in a district wins the district and the parties generally attempt to win as many districts as possible. Among MR systems, single member district plurality (SMDP)—where individuals cast vote for one candidate in single member district and the candidate with the most votes is elected—is the most common. SMDP system is currently followed for legislative

[20]This force is absent in Trebbi, Aghion and Alesina (2008) since candidates in their model can not change platforms in response to changing concentration of groups in certain districts. This partly explains the result of their model that when the minority becomes large, representation under MR becomes unresponsive to group size.

elections in countries such as India, Nigeria and United Kingdom among others. Around 63 percent of country-year observations that follow MR have this system in our dataset.[21]

In contrast, in PR system, parties typically present list of candidates and seats are allocated to parties in proportion to their vote share in multimember districts. This reduces disparity in vote share at the national level and seat share of a party in the parliament. Examples of countries that currently have PR system are Argentina, Belgium, South Africa and Turkey among others.[22] Appendix figure F1 depicts the countries with MR and PR systems in 2013.

We discuss the trends in choice of MR and PR by countries over the decades in appendix section A. However, one aspect of the choice is worth highlighting here—namely the role played by colonial history in shaping the electoral systems of the colonies. Most of the countries that were once British and French colonies adopted MR while those that had been colonized by Belgium, Netherlands, Portugal and Spain adopted PR. We discuss this aspect of the choice of electoral systems in the empirical analysis to address causality.

It is important to note here that the electoral system pertains to the legislature while we look at representation of minorities in the national government (or the executive). Our analysis includes countries with both parliamentary and presidential systems. The fact that in parliamentary systems representation in the legislature has a bearing on the executive is understandable, since the executive is selected from the legislature itself. The case for non-parliamentary systems, however, is less obvious and needs an explanation. The first thing to note is that a significant proportion of such countries have a semi-presidential system where the cabinet is either formed by the legislature, or faces the threat of no confidence vote from the legislature, or both. France, Poland, Sri Lanka, Peru, and Senegal are examples of such countries. The difference in the strategic incentives of parties across MR and PR, therefore, would be relevant in such countries. Among

[21]Another variant of MR systems is a two-round system (TRS). In TRS candidates or parties are elected in the first round if their proportion of votes exceeds a specified threshold. Otherwise, a second round of elections takes place—typically one or two weeks later—among the top candidates. France and Mali currently employ TRS for parliamentary elections.

[22]Some countries also use mixed systems which are a combination of both MR and PR. However, we do not include them in our empirical analysis.

the countries with a presidential system, some still need formal approval of the legislature to form the cabinet. In fact, even in countries where the president can appoint and dismiss the cabinet freely without any legislative approval, there is a high correlation between seat share of parties in the legislature and seat share in the cabinet.[23] Therefore, the electoral strategies of the parties to form the government seem to be similar to the strategies for legislative elections even in presidential systems. This is understandable given that legislative and executive elections are often held simultaneously and consequently, political parties have consistent platforms (in terms of group representation) for both elections.[24]

4 Data Description

4.1 Data Sources

In this section, we briefly describe the various data sources that we have put together for this project. To conserve space, the full description of each of the datasets is provided in appendix section B.

EPR: The information on political representation and demographic details at ethnic group level comes from Ethnic Power Relations (EPR) core dataset 2014 (Vogt et al., 2015). The EPR dataset provides a measure of political representation in the national government for every ethnic group in a country for all years from 1946–2013. The measure, called the "power rank," can belong to one of six categories signifying the degree of representation. These are, in descending order of power, *monopoly*, *dominant*, *senior partner*, *junior partner*, *powerless*, and *discriminated by the state*. The first two categories refer to cases where a group has substantial representation in the government, for the next two, some representation, and the final two categories refer to cases where the group has no representation. The categorization is created by the scholars in the field after taking inputs from over one hundred country experts. It is nonetheless a subjective

[23] Silva (2016), for example, shows that in Brazil even though the party of the president gets an advantage in the cabinet, the cabinet portfolio share increases by 0.9 percent for every percentage point increase in legislative seat share even for non-presidential parties.

[24] All our empirical results remain the same if we do not consider countries with the presidential system where the president doesn't require any approval from the legislature for cabinet formation.

measure of representation and therefore, could potentially be biased. We discuss in section 4.2 how we deal with this issue of subjectivity in our measurement. Apart from the power rank measure, EPR also provides annual group-country level data on population shares, settlement patterns and other characteristics of groups.

Nightlight Luminosity: The EPR dataset is complemented with GeoEPR dataset (Wucherpfennig, 2011) which consists of GIS maps of the settlement areas of a subsample of ethnic groups in the EPR dataset which are geographically concentrated in a region. These maps are overlaid with DMSP-OLS Nighttime Lights Time Series to measure average nightlight luminosity in an ethnic group's settlement area.

Electoral Systems Data: The data for electoral rules come from two sources —the Democratic Electoral Systems (DES) and the IDEA Electoral System Design Database. For any given year, the electoral system in a country is the electoral system used in the most recent election. We restrict our analysis to Majoritarian and Proportional systems.

Polity IV: Polity IV Project allows us to identify periods of autocratic and democratic rule for a country. We define democracy as country-year pairs where the position of the chief executive is chosen through competitive elections and include only those observations in the sample.[25]

Colonial History: The ICOW Colonial History Dataset 1.0 (Hensel, 2014) is used to identify the primary colonial ruler and the year of independence for each country that was colonized. The primary colonial ruler is typically the state that ruled the largest area of the colony or ruled it for the longest time. We use this dataset to find the electoral rule followed by the primary colonial ruler in the colony's year of independence for our identification strategy.

[25] Our results are robust to using the more conventional definition of democracy based on the polity score.

Road Network: The Global Roads Inventory Project (GRIP) (see Meijer et al. (2018)) provides raster data at a resolution of 5 arcminutes (approximately 8 x 8 km^2 at the equator) on road density (road length per unit area) for various kinds of roads across several countries as they exist currently. We overlay the road network map on the maps of ethnic groups and national boundaries to calculate a cross-sectional measure of road length (in kilometers) per square kilometer of the settlement area of an ethnicity. We use this measure for our robustness exercise (Section 8.2).

4.2 Subjectivity in Political Representation Measurement

One concern with the power rank variable is that it is a subjective measure and therefore, could potentially be biased. We address the concern in three different ways. First, we use the dataset created by Francois, Rainer and Trebbi (2015) for 15 countries in Africa. It contains share of cabinet positions held by ethnic groups within each country for every year during 1960–2004. This could be considered to be a more objective measure of representation. We, therefore, match the ethnic groups from that dataset to the EPR data. We are able to match 90% of groups. We then correlate the power ranking (from the EPR data) with cabinet shares. Figure F2 graphically shows this correlation using a binned scatter plot. We observe that the two variables are highly positively correlated (r = 0.56) and also, the nature of the relationship is linear. This suggests that the power rank variable is indeed informative about the real power held by groups within governments. Further, we notice that out of the six categories of power rank, the last two categories (*powerless* and *discriminated by the state*) refer to cases where the group is either has no representation in the government or is actively discriminated by the state. We consider these to be the stark cases where the problem of subjectivity is presumably minimal. We therefore coarsen the power ranking to create an indicator of *political inclusion* which takes value one if a group is neither powerless nor discriminated by the state, and zero otherwise. The political inclusion indicator therefore measures the extensive margin of representation of a group, i.e., whether the group has *any* representation in the government or not. We take the indicator of political inclusion as our main political variable.[26] Moreover, we argue in section 7 that given our empirical specification, our results are unlikely to be driven by biased measurement. Finally, we provide evidence that, consistent with model's prediction, the pattern is replicated with measures of material well-being of groups.

4.3 Summary Statistics

Appendix table F1 reports summary statistics for both the ethnicity level (Panel A) and the country level (Panel B) variables. In our final data, 43.87 percent of country-year observations have MR, whereas 56.13 percent have PR. The countries with MR are more fractionalized, have greater number of relevant groups, but allow lesser political competition and place fewer constraints on decision making powers of the chief executive compared to PR. However, these differences are not statistically significant at 10 percent level. On an average, the largest group comprises 73.5 percent of the politically relevant population and in 84.9 percent of country-year observations the largest group has an absolute majority in the country (i.e., population share over 50 percent). Overall 36.6 percent of minorities are politically included and 78.4 percent are geographically concentrated. The ethnicity level characteristics are also not significantly different between countries with MR and PR systems.

5 Empirical Methodology

We use the linear probability model to estimate the effect of group size on political inclusion under MR and PR. In the baseline specification we first check if population share of a group has any relationship with its probability of inclusion in the national executive and whether the relationship is different across the two electoral systems. The following is our preferred specification:

$$\mathbb{P}[I_{ict} = 1] = \delta_{ct} + \beta_1 n_{ict} + \beta_2 n_{ict}^2 + \beta_3 P_{ct} * n_{ict} + \beta_4 P_{ct} * n_{ict}^2 + \gamma X_{ict} + \epsilon_c \quad (3)$$

[26] We of course show robustness of our results to using the main power rank variable as the outcome variable.

where I_{ict} is a dummy indicating whether group i is politically included in country c in year t, δ_{ct} denotes fixed effects at the level of country-year pairs, n_{ict} is the population share of the group, P_{ct} is a dummy indicating whether PR system has been used in the latest national elections in country c in year t; X_{ict} is a vector of ethnicity level controls (which include years of peace, settlement patterns, trans-ethnic kin inclusion/exclusion and fraction of the group associated with the largest language and religion in the group). The error term ϵ_c is clustered at the country level. We include a square term for the population share of the group to check for non-linearity in the relation.

Given this specification, we compare groups *within* a country-year. Therefore, we only consider countries with 2 or more minorities. The specification controls for a variety of observables and unobservables that vary at the country-year level and may affect the relation we estimate. We argue that two groups of same size across two different countries or in same country but in two different years may wield different political power. This is because a group's access to state power may depend on the number and size composition of all the groups, including the majority, their explicit or implicit political alliances, electoral strategies of political parties, voters' attitudes towards the groups and any political, economic or social contingency that may affect all these factors in complex and unpredictable ways. It may depend on other historical and cultural factors as well, which may depend on time varying characteristics of the country which are often hard to observe. By comparing groups within a country-year observation, we are able to cut through all these issues which may affect a group's political representation and focus sharply on group specific features only. Our analysis, therefore, avoids any "cross-country" analysis in the sense that the coefficients are not estimated by comparing groups across countries (or by comparing the same group over time).

An alternative, though imperfect, way of estimating the relationship would be to use the panel nature of our data and compare the same minority over time, by exploiting its temporal change in population share and political inclusion status. However, the estimation strategy suffers from a major drawback. There are unobservable political factors in a country, some of which we have listed above, that can change over the years which may affect the relationship we wish to estimate. A panel regression would not be able to absorb such changes. For this reason this is not our preferred empirical specification. We therefore relegate the discussion on the panel analysis to appendix section C.

6 Identification

The baseline specification treats the electoral system of a country as exogenous. However, scholars have argued that the choice of electoral system is endogenous to the existing power structure of the country (Boix, 1999; Lijphart, 1992; Trebbi, Aghion and Alesina, 2008). In the presence of such concerns our interaction terms in specification (3) are likely to be misidentified. One potential solution to the issue could have been to focus on the few countries that switch from one electoral system to the other during the sample period. However, such switches themselves could be endogenous as they could be precipitated by the discontent of some of the groups with the current distribution of power.

Reynolds, Reilly and Ellis (2008) argue that many colonies adopted the electoral system of their colonial ruler. We, therefore, look at a subset of countries which had once been colonies. We use the primary colonial ruler's electoral system in the year of independence of a colony as an instrument for electoral system of the colony. The exclusion restriction for this specification requires that the colonialists' electoral system did not have a direct *differential* effect on the political power of minorities of different sizes. This would hold even if the electoral system of the colonial ruler is correlated with the power of minority groups on average as long as it is uncorrelated with the power inequality among minorities. For example, one concern with the IV strategy could be that the British might have had more egalitarian and permissive legal codes in their colonies as compared to the Spanish or Portuguese. This might result in differences in political representation of minorities across MR and PR countries today. However, as long as the liberal legal codes of British colonies gave similar kind of advantages to minorities of all sizes—which is likely to be the case—our identification strategy would remain valid.[27]

[27]There could be a further threat to the IV strategy if, for example, the colonial rulers with different electoral systems happened to colonize countries having different group size compositions. Appendix table F5 reports the results of regressing the indicator that colonialist's electoral system is PR on various population composition measures (fractionalization of minorities, number of minorities etc). We use the population figures of the groups and number of

For the IV strategy, we keep only those colonies in the sample which democratized not too long after gaining independence from their colonial ruler. Some countries, such as Indonesia and Brazil, became dictatorships after independence and remained so for many decades before democratizing. In such cases the colonial ruler's electoral system matters much less for a country. For example, there are 7 countries which democratized at least 50 years after becoming independent.[28] Only one of them have MR even though all except one were colonized by countries with the MR system. We use two thresholds for our selection of sample: countries which democratized within 30 and 50 years of independence.[29] We first run the following first stage regressions:

$$P_{ct} * n_{ict} = d_{ct} + a_1 n_{ict} + a_2 n_{ict}^2 + a_3 H_c * n_{ict} + a_4 H_c n_{ict}^2 + \pi X_{ict} + u_c$$

$$P_{ct} * n_{ict}^2 = e_{ct} + b_1 n_{ict} + b_2 n_{ict}^2 + b_3 H_c * n_{ict} + b_4 H_c n_{ict}^2 + \omega X_{ict} + v_c$$

where $H_c = 1$, if colonialist of country c had proportional system in the colony's year of independence. We then get the estimates of β_1–β_4 from specification (3) in the second stage regression.

7 Results

7.1 Verifying the Main Assumption of the Model

Before we discuss the empirical results, we verify one key parameter restriction of the model that we need for our main theoretical result to hold. Proposition 3 requires the minority groups' settlement areas to be inelastically related to their population shares. Moreover, Bettencourt (2013) argues that the value of α should be 0.67. We run the following specification to test this:

$$\ln S_{ict} = \alpha \ln n_{ict} + \gamma X_{ict} + \delta_{ct} + \epsilon_c \tag{4}$$

groups for the earliest period in the sample when the country was independent. The coefficients show that colonialist's electoral system is not correlated with the population composition of groups at or near the time of the colony's independence.

[28] These are Bhutan, Brazil, El Salvador, Honduras, Indonesia, Nicaragua and Panama.

[29] There are 18 countries which democratized over 30 years after independence. Of them 10 have PR, though only 2 were colonized by countries with a PR system.

Table 1—Settlement Area Expands Inelastically: $\alpha = 0.67$

	ln(Settlement area)		
	(1)	(2)	(3)
α: ln(Population share)	0.625	0.661	0.668
	(0.122)	(0.134)	(0.124)
$H_0 : \alpha \geq 1$ (one tailed p-value)	0.001	0.007	0.005
$H_0 : \alpha = 0.67$ (p-value)	0.736	0.968	0.992
Mean dependent	10.140	10.006	9.783
Observations	6,665	5,946	4,357
R-squared	0.792	0.779	0.742
Ethnicity-year controls	YES	YES	YES
Country-year FE	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. All concentrated minorities are in column (1). Minority population share in column (2) $\leq$ 0.25 and that in column(3) $\leq$ 0.10. Standard errors clustered at the country level are reported in parentheses.

where S_{ict} is the settlement area of a group i in country c in year t and n_{ict} is the population share that group. α measures the elasticity of settlement area with respect to population share of a group, and therefore, is a direct estimate of the parameter α in the model. The EPR dataset provides information about the settlement area of groups which are geographically concentrated. This allows us to estimate equation (4). We report the results in table 1. Column (1) reports the main estimate of α to be 0.625. It is statistically significant at 1 percent level and significantly lower than one, also at 1 percent level. Further, the coefficient is statistically indistinguishable from 0.67, confirming the prediction of Bettencourt (2013). Moreover, we estimate this parameter in two sub-samples—where the minority groups' population shares are smaller than 0.25 (column (2)) and smaller than 0.1 (column (3)). Both estimates are close to each other and are similar to the main estimate. This shows that the elasticity of settlement area with respect to population share of a group is indeed stable, further confirming our model's assumption. It is important to mention here that this result is in line with papers that also verify the theoretical claim of Bettencourt (2013) in various contexts (Ortman et al., 2014, 2015; Ortman et al., 2016; Cesaretti et al., 2016).

7.2 Baseline Results

Table 2 column (2) shows the results from our baseline specification. The coefficient of population share is positive and significant at 1 percent level and that of population share-squared is negative and significant at 5 percent level. Their magnitudes imply that for MR countries there is an inverted-U shaped relation between population share of a group and its probability of political inclusion. Probability of political inclusion attains its peak at population share of 0.26. The interactions of population share and its square with the PR dummy are statistically significant (at 5 percent level) and have opposite signs. F-tests for the hypotheses $\beta_1 + \beta_3 = 0$ and $\beta_2 + \beta_4 = 0$ give p-values of 0.33 and 0.96 respectively. This indicates that there is no relation between population share and political inclusion under PR. Column (1) reports the results with a weaker specification—having country and year fixed effects separately. We observe that the coefficients remain similar in magnitude. Also, the PR dummy has a positive and marginally significant coefficient. This suggests that very small minority groups presumably enjoy higher political representation under PR compared to MR.

The aforementioned result is unlikely to be driven by a systematic bias in coding of the power rank variable. Since we compare the groups within a country-year pair, we effectively control for the researcher(s) who were responsible for the power ranking of these groups. For the result to be driven by biased coding, it must be the case that the sets of researchers coding the MR and PR countries are systematically biased against subsets of minority groups with different population shares. Further, the coefficients of ethnicity level controls as reported in appendix table F2 are of the expected sign. The coefficient of peace years is positive and statistically significant at 1 percent level. An additional decade without any conflict incidence experienced by an ethnicity is associated with 4.15 percent more likelihood of its political inclusion. The coefficient of transethnic-kin exclusion dummy is positive and significant. This might be due to the fact that politically excluded ethnic groups sometimes migrate to countries where they might get political representation. An indicator of an ethnic group's cohesiveness is the fraction of its members associated with the largest language spoken by the group.

Table 2—Inverted-U shaped relation under MR and no relation under PR

	Political inclusion		Nightlight	
	(1)	(2)	(3)	(4)
β_1: Population share	4.405	4.825	10.77	11.02
	(1.239)	(1.227)	(3.741)	(3.840)
β_2: Population share - squared	-7.884	-9.276	-24.48	-24.49
	(3.883)	(3.955)	(10.67)	(11.30)
β_3: Proportional x Population share	-3.011	-3.661	-9.729	-10.11
	(1.687)	(1.721)	(5.840)	(6.103)
β_4: Proportional x Population share - squared	6.903	9.106	24.03	24.29
	(5.159)	(5.313)	(15.27)	(16.17)
Proportional	0.247		0.328	
	(0.144)		(0.382)	
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.23	0.33	0.84	0.86
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.77	0.96	0.97	0.99
Predicted optimal size	0.279	0.260	0.22	0.21
Mean dep. var.	0.366	0.366	-0.23	-0.26
Observations	9,294	8,706	3,756	3,469
R-squared	0.652	0.687	0.821	0.816
Ethnicity-year controls	YES	YES	YES	YES
Country-year controls	YES	NO	YES	NO
Country FE	YES	NO	YES	NO
Year FE	YES	NO	YES	NO
Country-year FE	NO	YES	NO	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. The dependent variable for columns (1) and (2)—political inclusion—is a dummy variable that takes value one if the group in a country in a given year is neither powerless nor discriminated by the state. Column (1) includes 438 ethno-country groups in 102 countries, and column (2) includes 421 ethno-country groups in 87 countries over the period 1946–2013. The dependent variable for columns (3) and (4) is logarithm of nightlight luminosity per unit area of groups which have well-demarcated settlement areas. Standard errors clustered at the country level are reported in parentheses.

Groups that are linguistically more cohesive find it easier to organize themselves and put forth their demands. Therefore, they are more likely to be politically included. This is supported by the result that a 10 percentage points increase in fraction of group members associated with the largest language for the group is related with a 2.10 percent increase in likelihood of political inclusion for the group. Additionally, table F3 reports the results of various robustness exercises we carry out, including using the power rank as an outcome variable, to ensure that the result is not driven by either sample selection or the chosen specification. We discuss these results in the appendix section D.

The model we develop predicts that per capita allocation of public resources to ethnic groups follows the same pattern as their political representation. We test this using the same specification. We use nightlight intensity as a proxy for public resource allocation towards groups which are settled in a geographically well demarcated region within a country. Electricity in most countries is publicly provided and is an essential public good for any region within a country. Therefore, nightlight luminosity acts as a direct proxy for government allocation of resources, in the form of electricity access, in an area. In fact, this has been shown to be the case in Senegal and Mali (Min et al., 2013), and Vietnam (Min and Gaba, 2014). We use (logarithm of) nightlight intensity per unit area as our dependent variable to test specification (3).[30] Michalopoulos and Papaioannou (2014) use the same measure to proxy for economic development of ethnic groups in the African continent. They further use micro-data from Afrobarometer surveys to confirm that the measure is a good proxy for various public goods such as "access to electrification, presence of a sewage system, access to piped water, and education" within settlement areas of ethnic groups.[31] Given the volume of evidence coming from a wide range of countries, we feel confident that our measure is a good proxy for allocation of government resources, and more generally, for the level of development of an ethnic group.[32]

[30]We add 0.01 as a constant to nightlight intensity per area measure before taking the logarithm.

[31]Nightlight luminosity is also a well-documented and widely used proxy for the level of economic development of a geographic region. For a discussion about using nightlight luminosity as a measure of economic activity see Doll (2008) and Henderson et al. (2012). The papers using nightlight data as a proxy for economic development in various contexts are too numerous to cite here. The papers that use nightlight data to answer political economy related questions include among others, Michalopoulos and Papaioannou (2013, 2014), Prakash, Rockmore and Uppal (2015), Baskaran et al. (2015), Alesina et al. (2016) etc.

[32]Henderson et al. (2012) have raised important issues with using nightlight luminosity as proxy for economic outcome. Many of these concerns are however addressed in our empirical analysis. Firstly, Henderson et al. (2012) point out that the nightlight data is captured using different satellite sensors and therefore, the luminosity data is not comparable across the years. This is addressed in our analysis since we use country-year fixed effects. Henderson et al. (2012) similarly use year fixed effects to address the issue. The other concern is that nightlight data is not captured in countries with high latitudes during summer time. Thus, Henderson et al. (2012) remove the regions above the Arctic Circle from their analysis. All the countries in the Arctic Circle, barring Russia, are not in our sample as well, since they have only one minority group. The third concern with nightlight data is the phenomenon of blurring, i.e., tendency of light to be captured beyond the exact source (due to coarse light sensors). However this is more of an issue in using nightlight data in smaller areas. The extent of blurring ranges from 4.5 km to 9 km depending on the radiance of the light source (Abrahams et al., 2018). Since the median area of ethnic groups in our sample is about 23,500 square km, we do not think

The use of nightlight luminosity imposes two restrictions in the data—it is available only from 1992 onwards and can be used only for groups which have a well-demarcated and contiguous settlement area as specified by the EPR dataset. Table 2 column (4) reports the results. It shows that the result for political inclusion is replicated with nightlight as outcome variable. The estimated population share with peak nightlight intensity in MR countries is 0.21 which is similar to what we estimated for political inclusion. Moreover, we see in column (3) that the result do not change with the weaker specification. The dummy for PR system again has a positive (though imprecise) coefficient, consistent with the column (1) result. This suggests that the patterns of political inclusion indeed have implications for the level of per capita resource allocation of the groups.

7.3 Identification Results

The IV results are reported in table 3. Panel B (column 1) of the table shows that the presence of proportional electoral system in a country is 47 percent more likely in countries that democratized within 30 years of independence if the electoral system of its primary colonial ruler was also proportional in the colony's year of independence. The coefficient is statistically significant at 1 percent level. Panel A reports the second stage results using political inclusion dummy and log of nightlight intensity per unit area as dependent variables. The first two columns report the results for countries which democratized within 30 years of being independent and the next two columns report the same with a 50 year threshold. In all four columns we find the same pattern. For MR countries we get a strong inverted-U shaped relationship. The peak is achieved at population shares 0.22 and 0.24 for political inclusion, and 0.22 and 0.26 for nightlight intensity, for the 30 and 50 year threshold regressions respectively. Moreover, the relationships are indeed flat for PR, as both the tests of $\beta_1 + \beta_3 = 0$ and $\beta_2 + \beta_4 = 0$ fail to reject the null hypothesis for all the four columns. The coefficients for political inclusion across columns (1) and (3) are similar in magnitudes and comparable to the coefficients estimated in the baseline specification (table 2, column (4)). Importantly, the Kleibergen-Paap rk LM statistic for the first stage regressions are

this to be a major source of measurement error.

Table 3—IV strategy replicates main results

Panel A: Second stage	Lag < 30 years		Lag < 50 years	
	Political inclusion	Nightlight	Political inclusion	Nightlight
	(1)	(2)	(3)	(4)
β_1: Population share	6.142	26.87	5.832	21.95
	(1.999)	(10.23)	(1.541)	(8.936)
β_2: Population share - squared	-13.72	-68.13	-12.17	-46.23
	(6.695)	(25.77)	(4.629)	(21.89)
β_3: Proportional x Population share	-4.332	-41.69	-4.049	-36.53
	(2.421)	(17.67)	(1.962)	(16.37)
β_4: Proportional x Population share - squared	14.77	92.90	13.35	69.35
	(8.698)	(49.71)	(7.158)	(44.86)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.102	0.174	0.102	0.174
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.859	0.488	0.844	0.504
Predicted optimal size	0.223	0.224	0.239	0.263
Observations	4,361	1,720	4,632	1,926
R-squared	0.700	0.773	0.711	0.766
Ethnicity-year controls	YES	YES	YES	YES
Country-year FE	YES	YES	YES	YES
Kleibergen-Paap rk LM stat	5.06	3.10	5.12	3.12
Cragg-Donald Wald F stat	172.18	43.01	188.47	50.96
F stat (Proportional*Population share)	119.51	40.47	260.33	125.57
F stat (Proportional*Population share - squared)	312.74	72.36	919.01	516.56

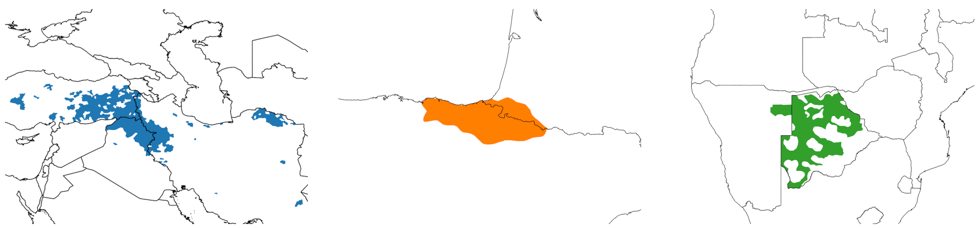
Panel B: First Stage (Country level)	Proportional		Proportional	
Colonialist proportional	0.470		0.522	
	(0.162)		(0.143)	
Observations	508		818	
R-squared	0.653		0.561	
Region-year FE	YES		YES	

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. Political inclusion (dependent variable in columns (1) and (3)) is a dummy variable that takes value one if the group in a country in a given year is neither powerless nor discriminated by the state. The dependent variable in columns (2) and (4) is logarithm of nightlight luminosity per unit area of groups which have well-demarcated settlement areas. The first two columns in Panel A and the first column in Panel B include countries which were once colonies and democratized within 30 years of gaining independence ("Lag < 30 years"). The last two columns in Panel A and the second column in Panel B has the same sample restrictions with the independence-democracy lag being changed to a maximum of 50 years ("Lag < 50 years"). Standard errors clustered at the country level are reported in parentheses.

high in all specifications, alleviating concerns related to under-identification. The F statistics for the two first stage regressions are also very large in magnitudes in each case. Finally, for the sake of transparency, we report in appendix table F4 the IV strategy results when we do not put any restrictions on the sample. Both political inclusion (column 1) and nightlight (column 2) regressions show an inverted-U shaped relationship for MR countries. We get a flat relationship for political inclusion in PR countries. For the nightlight regressions, however, the β_3 and β_4 coefficients have the wrong signs. The column (2) coefficients are also noisy. Importantly, the regressions don't pass the under-identification tests as the Kleibergen-Paap rk LM statistics are low. This suggests that our sample restrictions are indeed useful in making our specification stronger.

8 Robustness Exercises

8.1 Comparing Same Group Across Countries



(a) Kurds (b) Basques (c) San

Figure 1. Examples of groups with settlement areas across national boundaries. Panel (a) shows *Kurds* in Iran, Iraq and Turkey; panel (b) shows *Basques* in France and Spain; and panel (c) shows *San* in Botswana and Namibia.

Sometimes a group is present in more than one country and often those countries are in the same region.[33] Examples include the Kurds who are present in both Turkey and Iran (figure 1, panel A), the Basques in France and Spain (panel B) and the San in Botswana and Namibia (panel C) etc. Therefore, as a robustness exercise, we exploit the differences in the sizes of the *same* group across those countries to identify the effect of group size. When the countries have different electoral systems (as in the case of France and Spain), the differential effect of electoral systems could also be estimated by comparing the group across those countries. The idea is that the variation in population shares of the same group across countries within a region comes from the group being unequally divided into multiple national jurisdictions, and therefore, can be treated exogenously.[34] We estimate the following model:

[33] The countries belong to one of five regions: Africa, Asia, Americas, Europe and Oceania.

[34] The method is similar to Dimico (2016) who shows in the context of Africa that the partition of an ethnicity in two countries adversely affects their political representation when the resulting groups are small. We, on the other hand, show that the effect of how an ethnic group is divided in two democracies on the group's political representation and economic development depends on the electoral system.

Table 4—Comparing same group across countries replicate main results

	Political inclusion	ln(Nightlight per area)
	(1)	(2)
β_1: Population share	10.44	58.54
	(2.424)	(35.90)
β_2: Population share - squared	-26.13	-156.4
	(6.091)	(92.29)
β_3: Proportional x Population share	-8.269	-58.72
	(2.686)	(35.96)
β_4: Proportional x Population share - squared	25.79	147.7
	(10.96)	(96.88)
Proportional	0.138	0.991
	(0.0513)	(1.352)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.17	0.99
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.96	0.83
Predicted optimal size	0.200	0.187
Observations	1,370	417
R-squared	0.836	0.887
Group-year controls	YES	YES
Country-year controls	YES	YES
Group-region-year FE	YES	YES

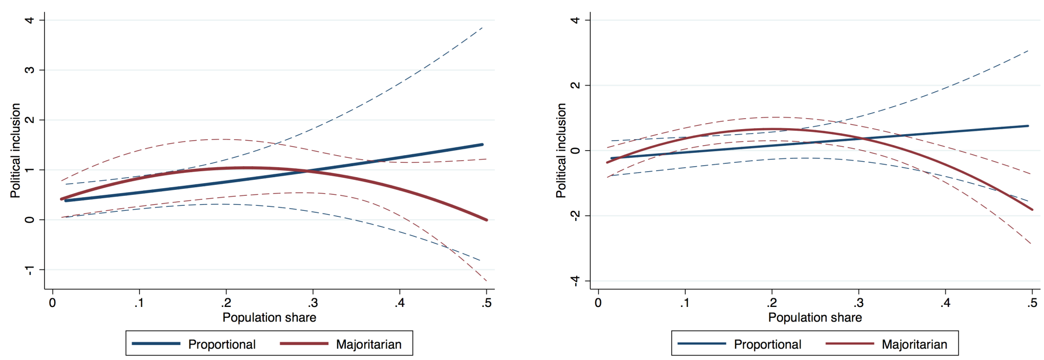
Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. Column (1) compares 21 groups in 40 countries and column (2) compares 12 groups in 30 countries. Standard errors double clustered at the group and country level are reported in parentheses.

$$\mathbb{P}[I_{ict} = 1] = \delta_{irt} + \theta P_{ct} + \beta_1 n_{ict} + \beta_2 n_{ict}^2 + \beta_3 P_{ct} * n_{ict} + \beta_4 P_{ct} * n_{ict}^2 + \gamma X_{ict} + \epsilon_{ic}$$

where δ_{irt} denotes ethnicity-region-year fixed effects, error term ϵ_{ic} is double clustered at group and country level to adjust standard errors against potential auto-correlation within group and country. The coefficient θ is the intercept of the relationship and β_1–β_4 are our other coefficients of interest, as before.

Table 4 reports the coefficients with political inclusion (column 1) and log nightlight intensity per area (column 2) as dependent variables. The within group comparison reaffirms the inverted-U shaped effect of population share on political representation under MR and no relation under PR. The coefficients reported in column (1) are a bit larger compared to those estimated in the IV regression (table 3). The peak of political representation under MR is achieved at population shares of 0.20 in this identification strategy, which is similar to what we estimated before. We also find that nightlight intensity indeed has the same pattern with the peak achieved at population share of 0.19 for MR countries.

The coefficients estimated however have large standard errors, presumably due to small sample size. Also, the coefficient of the PR dummy is positive and significant for political inclusion, suggesting that minorities of very small size get better represented in PR relative to MR. We plot the marginal effect of population share on political inclusion for the two identification methods in the figure 2. It suggests that mid-sized groups enjoy higher level of political inclusion under MR.



(a) IV Estimation (b) Same Group Across Countries

Figure 2. Marginal Effect of Group Size on Political Inclusion

8.2 Using Road Network Data as Alternate Outcome

In this sub-section, we show the robustness of our results to using a different measure of resource allocation. We do so by using a cross-sectional dataset on road construction available from Global Road Inventory Project or GRIP (Meijer et al., 2018). Road construction is widely believed to be an important activity of governments and often constitutes an important item in their annual budget. Burgess et al. (2015), for example, use road building in Kenya to show how democracy affects allocation of public resources across ethnic groups. We construct cross-sectional data on road construction across ethnicity-country pairs and use it as a proxy for allocation of public resources to test the robustness of our result. Section 4.1 describes the data and the construction of our outcome variable. The data is then matched with our main dataset for the latest year (i.e., for 2013). We then run the cross-sectional version of specification 3 for the year 2013 (with only country fixed effects). Since we have a cross section of

Table 5—Road Construction and Electoral Systems

	Road length per unit area			
	Baseline		IV	
	(1)	(2)	(3)	(4)
β_1: Population share	1.492	1.236	3.456	2.430
	(0.762)	(0.863)	(1.540)	(1.690)
β_2: Population share - squared	-4.071	-3.229	-11.35	-8.246
	(2.267)	(2.396)	(5.248)	(5.499)
β_3: Proportional x Population share	-1.500	-1.180	-3.597	-2.563
	(0.981)	(0.981)	(1.699)	(2.035)
β_4: Proportional x Population share - squared	4.480	3.571	11.22	7.424
	(2.657)	(2.797)	(5.662)	(6.932)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.99	0.89	0.79	0.85
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.77	0.76	0.94	0.76
Predicted optimal size	0.18	0.19	0.15	0.15
F stat (Proportional x Population share)	–	–	16.36	11.65
F stat (Proportional x Population share - squared)	–	–	16.23	12.37
Observations	227	227	105	105
R-squared	0.750	0.777	0.754	0.768
Group Controls	NO	YES	NO	YES
Country FE	YES	YES	YES	YES

Notes: Dependent variable is kilometers of non-local roads in the settlement area of a group per square kilometer of the area in 2013. The data is cross-sectional. Columns (1) and (3) have no group level controls while columns (2) and (4) have the same set of group level control as the previous regressions. The baseline regressions (columns (1) and (2)) have 54 countries and IV regressions (columns (3) and (4)) have 24 countries. Standard errors clustered at the country level are reported in parentheses.

groups for a subset of democracies, the number of observations in the regression is small. So we run the specification with and without group level controls. The results are reported in table 5, columns (1) and (2). The coefficients indicate that the pattern mirrors our main result—an inverted-U shaped relationship in MR, and no relationship in PR. However, when we include group level controls, the coefficients expectedly become noisier.

We then run our IV strategy specification on the sample of erstwhile colonies for the year 2013. We use the 30 year democracy lag as the sample restriction. The results without and with group level controls are reported in columns (3) and (4), respectively. The F-stats of the first stage regressions are above the commonly used threshold of 10. The second stage estimates show that the pattern is replicated even in the small cross-sectional sample. The optimal group size in MR in the baseline specification is around 19 percent which is similar to the one estimated in the second identification strategy. The optimal group size for the IV specification is smaller at 15 percent. However, given that different (small) sub-samples of countries are used in some of the regressions, getting different estimates of the optimal group size is not unlikely.

9 Concluding Remarks

This paper examines how electoral systems influence the relation between population share of a minority group and its access to power in the national government. First, we develop a model with multiple minority groups that predicts that in PR countries, population share of a minority has no effect on its political representation, while in MR countries the relation is inverted-U shaped. We then compile a large panel dataset at the ethnicity level for 87 countries for the post war period to test the predictions of our model. The empirical analysis remarkably exhibits the same pattern for both measures of political representation as well as per capita resource allocation. Our results imply that electoral systems can have stark effect on power (and welfare) inequality. We get that under PR, group size inequality does not translate into inequality in the political representation of minorities and consequently, the inequality in material well-being would also be minimal. On the other hand, power inequality among minorities in MR countries may be lower or higher than group size inequality depending on the size distribution of the groups. It is the mid-sized minority groups that enjoy maximum access to power in MR, while the small and large minorities enjoy similar levels of representation. Our work further highlights the importance of settlement patterns of groups in determining their representation in the government under MR. We, however, take settlement patterns as exogenously given. One interesting line of future enquiry can be to consider the settlement patterns of mobile minorities to be endogenous and explore if electoral system influences the settlement decisions of such minorities. We wish to take up this issue in our future work.

References

Abrahams, Alexei, Christopher Oram, and Nancy Lozano-Gracia (2018). Deblurring DMSP nighttime lights: A new method using Gaussian filters and frequencies of illumination. *Remote Sensing of Environment*. 210: 242-258.

Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou (2016). Ethnic inequality. *Journal of Political Economy*. 124(2): 428-488.

Banducci, S.A., Donovan, T., Karp, J.A. (1999). Proportional Representation and Attitudes About Politics: Evidence from New Zealand. *Electoral Studies*. 18: 533-555.

Banducci, S.A., Donovan, T., Karp, J.A. (2004). Minority representation, empowerment, and participation. *Journal of Politics*. 66(2): 534-556.

Besley, Tim, Rohini Pande, Lupin Rahman and Vijayendra Rao (2004). The Politics of Public Good Provision: Evidence from Indian Local Governments. *Journal of the European Economic Association Papers and Proceedings*. 2(2-3): 416-426.

Besley, Timothy, Rohini Pande, and Vijayendra Rao (2007). Political economy of panchayats in South India. *Economic and Political Weekly*. 661-666.

Baskaran, Thushyanthan, Brian Min, and Yogesh Uppal (2015). Election cycles and electricity provision: Evidence from a quasi-experiment with Indian special elections. *Journal of Public Economics*. 126 (June). 64-73.

Bettencourt, Luís M. A. (2013). The origins of scaling in cities. *Science*. 340(6139): 1438-1441.

Bobo, Lawrence, and Franklin D. Gilliam (1990). Race, sociopolitical participation, and black empowerment. *American Political Science Review*. 84(2): 377-393.

Boix, Carles (1999). Setting the rules of the game: the choice of electoral systems in advanced democracies. *American Political Science Review*. 93(3): 609-624.

Bormann, N.C., Golder, M. (2013). Democratic Electoral Systems Around the World, 1946-2011. *Electoral Studies*. 32: 360-369.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel (2015). The value of democracy: evidence from road building in Kenya. *American Economic Review*. 105(6): 1817-51.

Cascio, Elizabeth U., and Ebonya Washington (2015). Valuing the vote: The redistribution of voting rights and state funds following the voting rights act of 1965. *The Quarterly Journal of Economics*. 129(1): 379-433.

Cederman, L.E., Wimmer, A., Min, B. (2010). Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics*. 62 (1): 87-119.

Cesaretti, Rudolf, Josï¿œ Lobo, Luï¿œs MA Bettencourt, Scott G. Ortman, and Michael E. Smith (2016). Population-area relationship for Medieval European cities. *PloS one*. 11(10): e0162678.

Cohen, F. (1997). Proportional versus Majoritarian Ethnic Conflict Management in Democracies. *Comparative Political Studies*. 30(5): 607-630.

Colomer, Josep M. (2004). The strategy and history of electoral system choice. *The Handbook of Electoral System Choice*. Palgrave Macmillan UK. 3-78.

De Luca, G., Hodler, R., Raschky, P. A., Valsecchi, M. (2018). Ethnic Favoritism: An Axiom of Politics?. *Journal of Development Economics*. Forthcoming.

Dimico, A. (2016). Size Matters: The Effect of Size of Ethnic Groups on Development. *Oxford Bulletin of Economics and Statistics*. 1-28.

Doll, Christopher NH (2008). CIESIN thematic guide to night-time light remote sensing and its applications. *Center for International Earth Science Information Network of Columbia University*, Palisades, NY.

Farrell, D. M. (2011). *Electoral systems: A comparative introduction*. Palgrave macmillan.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of economic growth*. 8(2): 195-222.

Fjelde, H., Hoglund, K. (2014). Electoral Institutions and Electoral Violence in Sub-Saharan Africa. *British Journal of Political Science*. 46: 297-320.

Franck, R., Rainer, I. (2012). Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa. American Political Science Review. 106(2): 294-325.

Francois, P., Rainer, I., Trebbi, F. (2015). How is Power Shared in Africa. *Econometrica*. 83(2): 465-503.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil (2012). Measuring economic growth from outer space. *American economic review*. 102(2): 994-1028.

Hensel, P. R., Mitchell, S. M. (2007). Issue Correlates of War Project User Manual for ICOW Data, Version 1.0. University of North Texas and University of Iowa. URL: http://www. icow. org.

Herrera, H., Morelli, M., and Palfrey, T. (2014). Turnout and power sharing. *The Economic Journal.* 124(574): F131-F162.

Kartal, M. (2014). Laboratory Elections with Endogenous Turnout: Proportional Representation Versus Majoritarian Rule. *Experimental Economics.* 18(3): 366-384.

Kunicova, J., Ackerman, S.R. (2005). Electoral Rules and Constitutional Structures as Constraints on Corruption. *British Journal of Political Science.* 35(4): 573-606.

Lijphart, Arend (1992). Democratization and constitutional choices in Czecho-Slovakia, Hungary and Poland 1989-91. *Journal of theoretical politics.* 4(2): 207-223.

Lijphart, Arend (2004). Constitutional design for divided societies. *Journal of democracy.* 15(2): 96-109.

Lizzeri, A., Persico, N. (2001). The Provision of Public Goods Under Alternative Electoral Incentives. *The American Economic Review.* 91(1): 225-239.

Marshall, M.G. (2016). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2015. Center for Systemic Peace.

Meijer, Johan, Mark AJ Huijbregts, Kees Schotten, and Aafke Schipper (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, 13(April).

Michalopoulos, Stelios, and Elias Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica.* 81(1): 113-152.

Michalopoulos, Stelios, and Elias Papaioannou (2014). National institutions and subnational development in Africa. *The Quarterly Journal of Economics.* 129(1): 151-213.

Milesi-Ferretti, G.M., Perotti, R. and Rostagno, M. (2002). Electoral systems and public spending. *The Quarterly Journal of Economics.* 117(2): 609-657.

Min, Brian, and Kwawu Mensan Gaba (2014). Tracking electrification in Vietnam using nighttime lights. *Remote Sensing.* 6(10): 9511-9529.

Min, Brian, Kwawu Mensan Gaba, Ousmane Fall Sarr, and Alassane Agalassou (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International journal of remote sensing*. 34(22): 8118-8141.

Moser, R.G. (2008). Electoral Systems and the Representation of Ethnic Minorities: Evidence from Russia. *Comparative Politics*. 40(3): 273-292.

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

Myerson, R.B. (1999). Theoretical Comparisons of Electoral Systems. *European Economic Review*. 43: 671-697.

Norris, P. (1997). Choosing Electoral Systems: Proportional, Majoritarian and Mixed Systems. *International Political Science Review*. 18(3): 297-312.

Ortman, Scott G., Andrew HF Cabaniss, Jennie O. Sturm, and Luis MA Bettencourt (2014). The pre-history of urban scaling. *PloS one*. 9(2): e87902.

Ortman, Scott G., Andrew HF Cabaniss, Jennie O. Sturm, and Luï¿œs MA Bettencourt (2015). Settlement scaling and increasing returns in an ancient society. *Science Advances*. 1(1): e1400066.

Ortman, Scott G., Kaitlyn E. Davis, Jose Lobo, Michael E. Smith, Luis MA Bettencourt, and Aaron Trumbo (2016). Settlement scaling and economic change in the Central Andes. *Journal of Archaeological Science*. 73: 94-106.

Persson, T., and Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. Cambridge, Massachusetts: The MIT Press.

Persson, T., Tabellini, G. (2003). *The Economic Effects of Constitutions*. Cambridge, MA: MIT Press.

Persson, Torsten, and Guido Tabellini (2004). Constitutional rules and fiscal policy outcomes. *American Economic Review*. 94(1): 25-45.

Prakash, Nishith, Marc Rockmore, and Yogesh Uppal (2015). Do criminally accused politicians affect economic outcomes? Evidence from India. *Households in Conflict Network (HiCN) Working Paper*, The Institute of Development Studies, University of Sussex.

Reynolds, A. (2006). *Electoral Systems and the Protection and Participation of Minorities.* London: Minority Rights Group International.

Reynolds, A., Reilly B., Ellis, A. (2008). *Electoral System Design: The New International IDEA Handbook.* International Institute for Democracy and Electoral Assistance.

Rickard, S. J. (2012). Electoral systems, voters' interests and geographic dispersion. *British Journal of Political Science*, 42(4), 855-877.

Rickard, S. J. (2012). A non-tariff protectionist bias in majoritarian politics: Government subsidies and electoral institutions. *International Studies Quarterly*, 56(4), 777-785.

Silva, Thiago (2016). Who Gets What? *Formateur* Advantage and Salience in Portfolio Allocation in Presidential Multiparty Systems. Working Paper.

Stojanovic, N. (2006). Do Multicultural Democracies Really Require PR? Counter-evidence from Switzerland. *Swiss Political Science Review.* 12(4): 131-157.

Trebbi, F., Aghion P., Alesina, A. (2008). Electoral Rules and Minority Representation in U.S. Cities. *The Quarterly Journal of Economics.* 325-357.

Vogt, M., Bormann, N.C., Ruegger, S., Cederman, L.E., Hunziker, P., Girardin, L. (2015). Integrating Data on Ethnicity, Geography and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution.* 59(7): 1327-1342.

Wagner, W. (2014). The Overstated Merits of Proportional Representation: The Republic of Macedonia as a Natural Experiment for Assessing the Impact of Electoral Systems on Descriptive Representation. *Ethnopolitics.* 13(5): 483-500.

Wimmer, Andreas, Lars-Erik Cederman, and Brian Min (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review.* 74(2): 316-337.

Wucherpfennig, J., Weidmann, N. B., Girardin, L., Cederman, L. E., Wimmer, A. (2011). Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Conflict Management and Peace Science.* 28(5): 423-437.

Online Appendix for "Group Size and Political Representation under Alternate Electoral System"

Appendix Sections:

A Trend in Electoral System and Minority Representation

From 1950s to the 1970s, a larger fraction of countries had the MR system. However, the past few decades have seen a trend towards the adoption of PR. This can be observed in figure A1, where we plot the number of country-year observations by electoral system for each decade from 1950s through 2000s. This is mostly driven by adoption of the PR system by the new democracies in Latin America, Africa, and Mediterranean, Central and Eastern Europe in the 1970s and 1980s.[35] Several countries have also changed their existing systems to electoral formulas that are more proportional. For example, Japan and New Zealand switched from MR and held their first general elections under a mixed system in 1996. Another case in point is Russia, which changed its mixed electoral system and employed PR for the 2007 legislative election.[36]

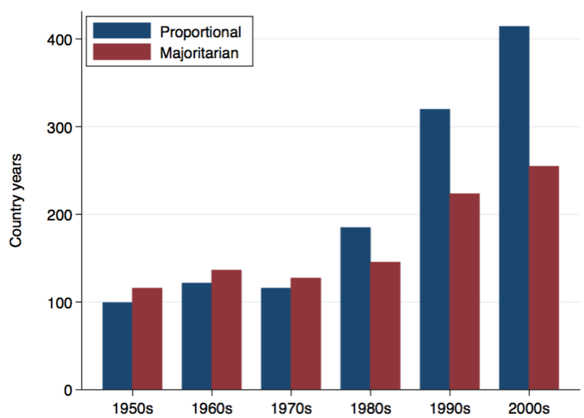


Figure A1. Electoral systems by decade

Figure A2 plots the proportion of minority groups in democracies in each power status category during 1946–2013. As it shows, there has been a gradual decline in state administered discrimination against minorities over the years. However, the share of groups in the powerless category has correspondingly increased. There is also no clear pattern in the proportion of groups in power sharing arrangements with other groups (i.e., junior and senior partner) and of

[35]The possible reasons for adoption of PR system by these countries are discussed in Farrell (2011).

[36]Other examples include Argentina, Sri Lanka and Moldova which switched directly from MR to PR for their parliamentary elections held in 1963, 1989 and 1998 respectively. There have also been a few instances of changes in the opposite direction—i.e. towards less proportionality. These include Venezuela, Madagascar and Bulgaria where PR was replaced in favor of mixed system in 1993, 1998 and 2009 legislative elections respectively.

those who rule virtually alone (dominant and monopoly groups). While this proportion was increasing during 1990s, it has remained virtually stable afterwards and was in fact declining during some of the earlier decades.

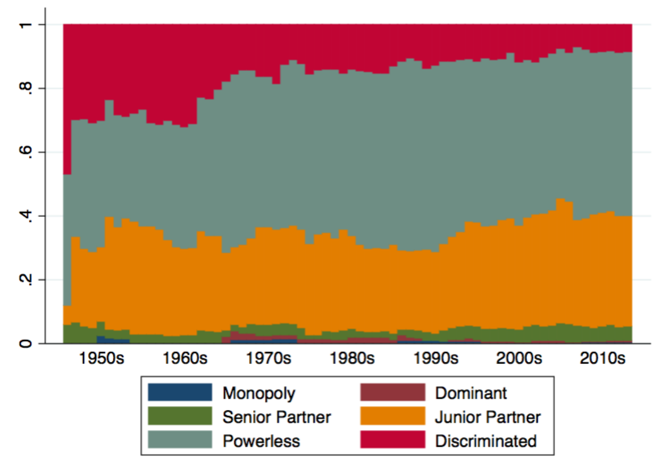


Figure A2. Minority power status over time

B Data Compilation

EPR: Our primary source of data is the Ethnic Power Relations (EPR) core dataset 2014 (Vogt et al., 2015). The dataset contains various characteristics of well-identified groups ("ethnicities") within countries for about 155 countries across the world at an annual level for the period 1946–2013. All sovereign states with a total population of at least 500,000 in 1990 are included in the dataset. The dataset defines a group "as any subjectively experienced sense of commonality based on the belief in common ancestry and shared culture."[37] (Cederman, Wimmer and Min, 2010) The dataset is concerned with groups that are politically relevant; a group is politically relevant if at least one political organization or a political party has at least once claimed to represent it at the national level or the group has been explicitly discriminated against by the state during any time in the period 1946–2013. This aligns with our interest as well. As long as there

[37] Cederman, Wimmer and Min (2010) further point out that in different countries different "markers may be used to indicate such shared ancestry and culture: common language, similar phenotypical features, adherence to the same faith, and so on." Further, in some societies there may be multiple dimensions of identity along which such "sense of commonality" may be experienced.

is some marker of identity which is salient in the society and is also politically meaningful, we should consider them in our analysis.

The demarcation of groups is intuitive and meaningful. India, a large and diverse country, for example, has 20 groups—the second highest in our sample.[38] These groups are based on religion (Kashmiri Muslims and Other Muslims), caste (SC/STs, OBCs) as well as language or ethnicity (Non SC/ST Bengalis, Non SC/ST Marathis, Mizo, Naga etc). United States, on the other hand, has 6 groups—Whites, African Americans, American Indians, Asian Americans, Arab Americans and Latinos. All the countries in our sample, barring India and Russia, have number of groups ranging from 2 to 14, with the average number of groups in the total sample being 4.6. We list in Appendix G the samples of countries used in our empirical exercises along with the respective number of minority groups and number of years in the sample, i.e., having a democratic regime.[39]

The dataset provides annual group-country level data on population shares, settlement patterns, trans-border ethnic kinship, as well as religious and linguistic affiliations for the period 1946–2013. However, most importantly for us, it also codes a group's access to national executive. A group's access to absolute power in the national government is coded based on whether the group rules alone (power status = *monopoly*, *dominant*), shares power with other groups (power status = *senior partner*, *junior partner*) or is excluded from executive power (power status = *powerless*, *discriminated by the state*). We rank these six categories in a separate variable called "power rank"; they range from 6 to 1 in decreasing order of power (i.e., from *monopoly* to *discriminated*).[40] The power ranking of the groups is evidently a subjective exercise. The researchers, however, are fairly transparent in the method that they follow in assigning power ranks. They look at the degree and nature of presence of members of a particular group in the most important political positions in the national government in determining its power rank. The details about group demarcation of the countries and the justification of the power rankings of each group is fully described on the official website of

[38]Russia with 39 groups has the highest number.

[39]It is important to note that politically relevant ethnic divisions in a country may change over time. New cleavages may emerge increasing the number of groups or some existing group may cease to be politically relevant as well. In case of South Africa, for example, racial divisions primarily between the Whites and the Blacks marked the political climate during the Apartheid era, while divisions within the black South African population along ethno-linguistic lines (such as between *Xhosa* and *Zulu*) has become more prominent in the subsequent period. The dataset recognize this fact. The number of groups in some countries, therefore, changes a little bit over the years. The number of minorities specified in Appendix G is the maximum number in the sample.

[40]There is an additional categorization in the data, known as *self-exclusion*. This applies to groups which have declared independence from the central state. They constitute only 0.7 percent of our sample and we do not consider them in our analysis.

the EPR project: https://growup.ethz.ch.

The EPR dataset also provides information about the settlement patterns of the groups. Specifically, it categorizes the groups as being *dispersed*, i.e., those who do not inhabit any particular region within a country and, *concentrated*, i.e., settled in a particular region of the country which is easily distinguishable on a map. For concentrated groups, it further gives information about the country's land area (km^2) that they occupy.[41]

The EPR dataset was created by scholars who work on group based conflict. The first version of the dataset was created as part of a research project between scholars at ETH Zurich and University of California, Los Angeles (UCLA), which was then updated and released by Vogt et al. (2015). The information about the attributes of groups, including their power status is coded by the researchers based on inputs from about one hundred country experts. This dataset has certain advantages for our paper over other existing datasets about political outcomes of groups. Some of the prominent datasets used by scholars of conflict are the Minorities At Risk (MAR) dataset, the All Minorities at Risk (A-MAR) dataset and the dataset used by Fearon (2003). Though most of these datasets give information about group sizes, none of these provide any detail about the settlement patterns of the groups. This is critical for us since we demonstrate that the pattern observed in our data is driven by groups which are geographically concentrated. Also, the EPR dataset provides information about the power status of all groups; this is in contrast to the MAR dataset which systematically excludes the groups that are in the government.

Electoral Systems Data: The data for electoral rules used for national elections come from merging two datasets. The first of these is the Democratic Electoral Systems (DES) data compiled by Bormann and Golder (2013). It contains details about electoral systems used for about 1200 national elections for the period 1946–2011. We complement this with a second source of data—the IDEA Electoral System Design Database, which gives us information about the electoral systems for some additional countries. The classification into broad electoral systems is based on the DES dataset. For any given year, the electoral system in a country is the electoral system used in the most recent election. We restrict our analysis to Majoritarian and Proportional systems.

Polity IV: Polity IV Project allows us to identify periods of autocratic and democratic rule in a country. We define democracy as country-year pairs where the position of the chief executive is chosen through competitive elections and include only those observations in the sample. We prefer this definition over

[41] The GIS shape file of their area of settlement is also provided on the EPR website.

the standard categorization based on the Polity IV score because we wish to look at all the countries that have competitive elections and have one of the two electoral systems of our interest. However, there are other aspects of a regime such as extent of checks and balances on the executive that affect the Polity IV score as well, which are of less relevance to our specific analysis. We, of course, show robustness of our result using a different definition of democracy based on the polity score.

Colonial History: The ICOW Colonial History Dataset 1.0 compiled by Hensel (2014) recognizes the primary colonial ruler and the year of independence for each country that was colonized. The dataset marks the start of colonial rule with the establishment of first permanent outpost or settlement by a state-sponsored company. *Primary colonial ruler* is defined as the one "most responsible for shaping the development" of the colony. This is typically the state that ruled the largest area of the colony or ruled it for the longest time. To obtain the electoral systems of the colonial rulers we use the data on electoral systems provided in The Handbook of Electoral System Choice (HESC) (Colomer, 2004). The HESC provides information about electoral systems of democracies since 1800. We use this to find the electoral rule followed by the primary colonial ruler in the colony's year of independence. We use this information for our identification.

Road Network: We compile a cross-sectional measure of road length at the ethnicity-country level. We use geo-spatial data from the Global Roads Inventory Project (GRIP) (see Meijer et al. (2018)) who provide GIS locations of various kinds of roads across several countries as they exist currently.[42] The dataset therefore is cross-sectional in nature. The dataset distinguishes among roads of five categories: national highways, primary roads, secondary roads, tertiary roads and local roads. We overlay the road network map on the maps of the settlement areas of ethnic groups and national boundaries to get the section of road network that lies within an ethnicity-country pair. We then aggregate the road length of the first four types of roads falling within the area of each pair. We don't consider local roads in our analysis because they are unlikely to be allocated by the national government. We use total road length per square kilometer of the settlement area of an ethnicity as our measure of public resource allocation towards that group.

[42] The dataset is freely available in the project website: http://www.globio.info/download-grip-dataset.

C Panel Analysis

The panel specification could be written as:

$$\mathbb{P}[I_{ict}=1]=\delta_{ic}+\phi_t+\beta_1 n_{ict}+\beta_2 n_{ict}^2+\beta_3 P_{ct}*n_{ict}+\beta_4 P_{ct}*n_{ict}^2+\gamma_1 X_{1ict}+\gamma_2 X_{2ct}+\epsilon_{ict} \quad (5)$$

where δ_{ic} is a group-country fixed effect, ϕ_t is a year fixed effect, X_{1ict} is a vector of ethnicity characteristics and X_{2ct} is a vector of country characteristics. However, there are two important drawbacks in this estimation strategy. Importantly, there are unobservable political factors in the country, some of which we have listed above, that can change over the years which may affect the likelihood of political inclusion of the group. The direction of this effect is uncertain as it would depend on the nature of the change in the political climate of the country. Therefore, the coefficients $\beta_1-\beta_4$ are likely to have noisier estimates. Also, the size composition of other groups, including the majority group would change over time which may affect the relationship as well. To partly account for the last factor we run specification (5) using relative population share in place of population share. The relative population share is defined as the ratio of the population share of the group in country-year observation to the population share of the majority group in the same country-year pair.

We report the results in table F9. Columns (1) and (4) report the results for our two main dependent variables using the full sample. We see that the coefficients β_3 and β_4 for column (1) do not have the expected signs and all the coefficients are noisily estimated. The coefficients for the nightlight regression (column 4) do have the expected signs. The magnitudes of β_1 and β_2 imply that group size has an inverted-U shaped relationship with nightlight intensity in MR countries, though the standard errors of the coefficients are high. The coefficients β_3 and β_4 have the opposite signs, implying that the relationship is flatter for PR. Since annual variations in population share would not immediately translate to changes in representation or material welfare, we keep in sample every third (columns (2) and (5)) and fifth (columns (3) and (6)) year that a group is present in the data. We see that the all coefficients for political inclusion have the expected signs in column (3), though the magnitude of β_3 is smaller than β_1. The coefficients for the nightlight regressions in column (5) and (6) maintain their correct signs. The coefficients for the interaction terms are, however, smaller in magnitudes. The panel results indicate that the relationship observed for minorities within a country-year becomes less precise when we follow the same minority over the years. This is expected given our discussion above.

D Robustness of the Baseline Result across Sub-samples

In table F3 we run specification (3) on various sub-samples of the data. Columns (1) and (2) show results for two time periods 1946–1979 and 1980–2013, respectively. The broad patterns depicted in our baseline specification continue to hold over time, though the coefficients are larger for the earlier period, indicating a more pronounced inverted-U relationship for MR countries in the first half of the post–war period. Column (3) shows the cross-sectional result for the latest year in our sample, i.e., for 2013. The coefficients here are quite similar to the column (1) coefficients. In column (4) we replace the main explanatory variable by the relative population share, i.e., the ratio of population share to the population share of the majority group in the country-year observation. Columns (5) restricts the sample to countries with an absolute majority and column (6) restricts the sample to parliamentary democracies only. In column (7) we only include election years in the sample and column (8) includes countries which are full democracies according to the Polity IV dataset (i.e., countries with a polity score of at least 7). Finally, in column (9) we use the power rank variable as our dependent variable. The variable takes value 1 through 6 with 1 being discriminated, 2 powerless and so on. In all specifications we fail to reject that $\beta_1 + \beta_3 = 0$ and $\beta_2 + \beta_4 = 0$. Therefore, in all specifications we get that there is no relation between population share and political inclusion in a PR system. Similarly, in all specifications we get that the relationship is inverted-U shaped in the MR system, though the coefficient β_2 is noisily estimated in some specifications. The consistency of the pattern across various sub-samples of the data strongly suggests that the result is a general phenomenon observed across democracies.

Further, we rerun the baseline regressions for political inclusion and log night-light per area by reweighing the observations by the (inverse of) the number of minority groups in the country-year observations. We do this to ensure that our results are not driven by countries with large number of groups. We report the results in appendix table F6. Coefficients in both columns suggest that our result remains the same with this specification.

E Validating Additional Comparative Static Results of the Model

The primary aim of the model is to justify the empirical pattern established in the Section 5 of the paper. The model, however, generates some additional predictions regarding the exact nature of the relationship between group size and access to political power. It is, therefore, important to test if these additional comparative static results hold in order to verify if the proposed model is indeed valid. We now turn to that discussion in the following paragraphs.

Proposition 3 states that we should observe the inverted-U shaped relationship between group size and power status under the MR system only for groups which are geographically concentrated. Also, a group's geographic concentration should not matter for the result of the PR system. We verify this by running the following specification for the samples of MR and PR country-year observations separately:

$$Y_{ict} = \delta_{ct} + \eta_1 n_{ict} + \eta_2 n_{ict}^2 + \eta_3 C_{ict} * n_{ict} + \eta_4 C_{ict} * n_{ict}^2 + \gamma X_{ict} + \epsilon_c \tag{6}$$

where C_{ict} is a dummy indicating whether the group i is geographically concentrated in country c in year t. Proposition 3 implies that for the sample of MR countries, η_1 and η_2 should be zero and we should have $\eta_3 > 0$ and $\eta_4 < 0$. For the set of PR countries all the coefficients η_1–η_4 should be zero. Table F7 reports the results and the predictions are verified. Column (1) reproduces the main result, and columns (2) and (3) provides the estimates of η_1–η_4 for MR and PR countries, respectively. As is evident, for the MR countries the relationship is only true for geographically concentrated groups. For PR countries, none of the coefficients are statistically significant.

Proposition 3 further specifies that under the MR system, the peak political representation is achieved when the population share of the group equals $\frac{1-n_3}{2}$ when the group is geographically concentrated, where n_3 is the population share of the majority group. Therefore, for larger values of the majority group's share, the peak is achieved at lower values of the minority group's size. We test this prediction by running specification (3) on various sub-samples of the data where we vary the size of the majority group. The results are reported in table F8. Columns (1)–(3) report the results for sub-samples where the majority group's population share is larger than 0.3, 0.5, and 0.7, respectively. The table also reports the population shares at which the peak inclusion is achieved. We see that the population share at which the peak inclusion is achieved declines as we move to countries with larger majority groups.

F Figures and Tables

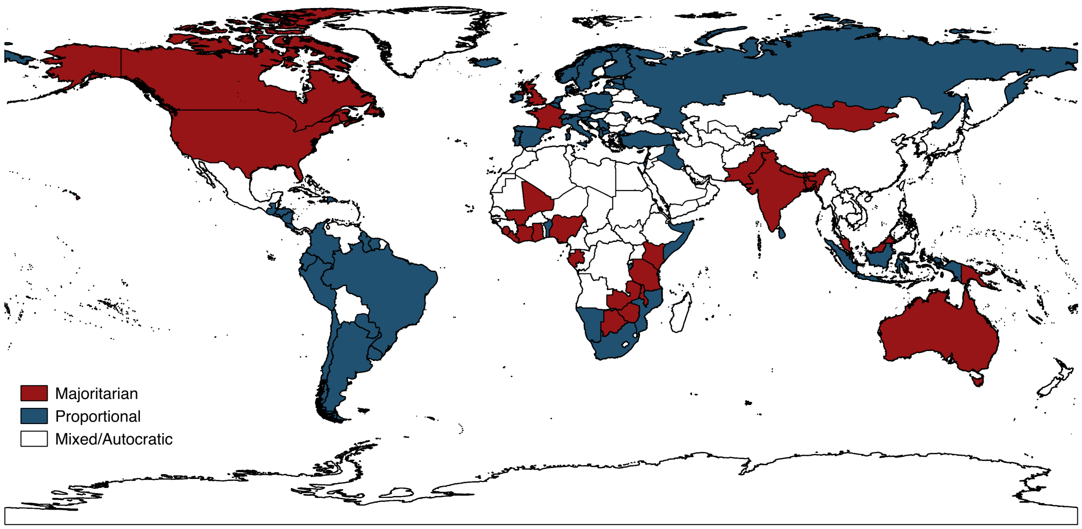


Figure F1. Electoral system distribution in 2013

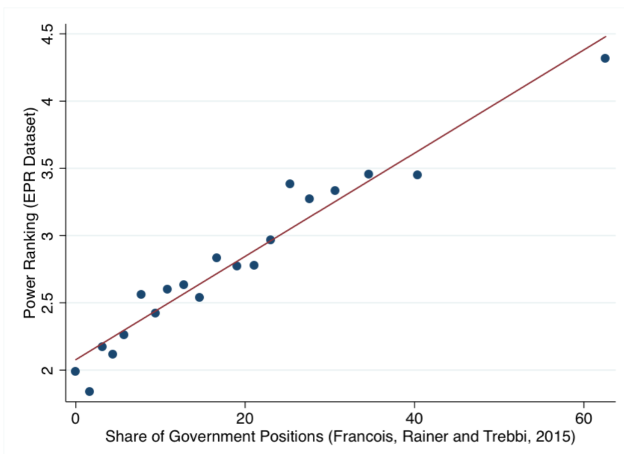


Figure F2. Comparing Power Rank with Cabinet Shares

Table F1—Descriptive statistics

	All data	Majoritarian system	Proportional system	Difference
	(1)	(2)	(3)	(4)
Panel A: Ethnicity level				
Political inclusion	0.366	0.444	0.275	0.169
	(0.482)	(0.497)	(0.446)	(0.112)
Power rank	2.294	2.391	2.180	0.211
	(0.793)	(0.770)	(0.806)	(0.188)
Population share	0.074	0.070	0.079	-0.009
	(0.099)	(0.090)	(0.108)	(0.024)
Years peace	31.418	29.223	34.029	-4.806
	(20.285)	(19.178)	(21.236)	(4.162)
Aggregate settlement	0.002	0.001	0.004	-0.003
	(0.046)	(0.031)	(0.059)	(0.005)
Statewide settlement	0.032	0.026	0.040	-0.014
	(0.176)	(0.158)	(0.195)	(0.045)
Regional and urban settlement	0.381	0.416	0.339	0.077
	(0.486)	(0.493)	(0.474)	(0.114)
Urban settlement	0.087	0.103	0.067	0.036
	(0.282)	(0.305)	(0.251)	(0.061)
Regional settlement	0.369	0.325	0.421	-0.096
	(0.482)	(0.468)	(0.494)	(0.106)
Dispersed settlement	0.109	0.118	0.098	0.020
	(0.312)	(0.323)	(0.298)	(0.074)
Migrant settlement	0.020	0.011	0.031	-0.020
	(0.140)	(0.103)	(0.174)	(0.028)
Transethnic-kin inclusion	0.417	0.402	0.435	-0.033
	(0.493)	(0.490)	(0.496)	(0.103)
Transethnic-kin exclusion	0.521	0.460	0.594	-0.135
	(0.500)	(0.498)	(0.491)	(0.105)
Fraction largest religion	0.719	0.750	0.682	0.069
	(0.209)	(0.222)	(0.186)	(0.053)
Fraction largest language	0.879	0.889	0.867	0.023
	(0.223)	(0.214)	(0.232)	(0.045)
Observations	9,294	5,049	4,245	9,294
Panel B: Country level				
Ethnic fractionalization	2.433	2.885	2.079	0.806
	(1.989)	(2.201)	(1.723)	(0.494)
Number of relevant groups	4.596	5.470	3.913	1.557
	(3.772)	(4.221)	(3.221)	(0.944)
Largest group size	0.735	0.687	0.772	-0.086
	(0.219)	(0.238)	(0.195)	(0.054)
Absolute majority	0.849	0.753	0.923	-0.170*
	(0.359)	(0.432)	(0.266)	(0.086)
Competitiveness of participation	3.989	3.873	4.079	-0.207
	(1.056)	(1.252)	(0.962)	(0.232)
Constraints chief executive	6.121	5.978	6.233	-0.256
	(1.291)	(1.370)	(1.497)	(0.270)
Observations	2,601	1,141	1,460	2,601

Notes: The data is at the ethnicity-country-year level for 438 ethno-country groups in Panel A and country-year level for 102 countries in Panel B for the period 1946–2013. Standard deviation in parenthesis in columns (1), (2) and (3). Standard errors clustered at the country level in parenthesis in the last column.

Table F2—Inverted-U shaped relation under MR and no relation under PR

	Political inclusion			
	(1)	(2)	(3)	(4)
β_1: Population share	4.405***	4.825***	10.77***	11.02***
	(1.239)	(1.227)	(3.741)	(3.840)
β_2: Population share - squared	-7.884**	-9.276**	-24.48**	-24.49**
	(3.883)	(3.955)	(10.67)	(11.30)
β_3: Proportional*Population share	-3.011*	-3.661**	-9.729*	-10.11
	(1.687)	(1.721)	(5.840)	(6.103)
β_4: Proportional*Population share - squared	6.903	9.106*	24.03	24.29
	(5.159)	(5.313)	(15.27)	(16.17)
Proportional	0.247*		0.328	
	(0.144)		(0.382)	
Years peace	0.00409***	0.00415***	0.00378	0.00361
	(0.00135)	(0.00130)	(0.00360)	(0.00357)
Aggregate settlement	0.549***	0.541***	-3.023***	-3.003***
	(0.110)	(0.114)	(0.432)	(0.430)
Statewide settlement	0.294	0.139		
	(0.375)	(0.352)		
Regional and urban settlement	0.174**	0.170**	0.0107	0.0264
	(0.0784)	(0.0789)	(0.419)	(0.410)
Urban settlement	0.0180	0.00905		
	(0.0663)	(0.0650)		
Regional settlement	-0.0105	-0.00942	-1.029**	-1.004**
	(0.0488)	(0.0483)	(0.399)	(0.392)
Migrant settlement	-0.140	-0.150		
	(0.195)	(0.195)		
Transethnic-kin inclusion	0.00421	0.000118	0.0502	0.0355
	(0.0446)	(0.0477)	(0.152)	(0.159)
Transethnic-kin exclusion	0.0897**	0.103***	0.114	0.134
	(0.0347)	(0.0348)	(0.132)	(0.142)
Fraction largest religion	-0.125	-0.108	-0.0609	-0.0201
	(0.109)	(0.105)	(0.537)	(0.530)
Fraction largest language	0.193**	0.210***	1.437***	1.441***
	(0.0737)	(0.0748)	(0.387)	(0.388)
Ethnic fractionalization	0.0203		-0.0680	
	(0.0251)		(0.0806)	
Number of relevant groups	0.0123		0.0134	
	(0.0197)		(0.0776)	
Competitiveness of participation	0.00848		0.0201	
	(0.0166)		(0.0310)	
Constraints chief executive	-0.0169		0.0107	
	(0.0104)		(0.0298)	
Observations	9,294	8,706	3,756	3,469
R-squared	0.652	0.687	0.821	0.816
Country FE	YES	NO	YES	NO
Year FE	YES	NO	YES	NO
Country-year FE	NO	YES	NO	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. The dependent variable for columns (1) and (2)–political inclusion—is a dummy variable that takes value one if the group in a country in a given year is neither powerless nor discriminated by the state. The sample for column (1) includes 438 ethno-country groups in 102 countries, and for column (2) includes 421 ethno-country groups in 87 countries the period 1946–2013. The dependent variable for columns (3) and (4) is logarithm of nightlight luminosity per unit area of groups which have well-demarcated settlement areas. Standard errors clustered at the country level are reported in parentheses.

Table F3—Main results are robust

	Political inclusion								Power rank
	1946-1979	1980-2013	2013						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_1: Population share	6.259	4.470	6.333		5.130	6.816	3.732	6.903	5.543
	(2.016)	(1.027)	(0.928)		(1.814)	(1.902)	(1.367)	(1.979)	(1.784)
β_2: Population share - squared	-14.14	-8.141	-12.31		-7.732	-17.18	-5.714	-17.21	-7.910
	(7.334)	(2.919)	(2.643)		(5.362)	(6.820)	(4.473)	(7.771)	(5.773)
β_3: Proportional x Population share	-6.674	-2.745	-4.838		-4.385	-7.080	-3.949	-6.577	-4.972
	(2.453)	(1.526)	(1.586)		(2.220)	(2.362)	(1.915)	(2.429)	(2.797)
β_4: Proportional x Population share - squared	17.33	6.322	10.79		9.334	22.30	9.625	20.29	11.81
	(8.306)	(4.472)	(4.404)		(6.619)	(8.043)	(5.958)	(8.821)	(9.201)
(β_1): Relative population share				2.381					
				(0.402)					
(β_2): Relative population share-squared				-2.108					
				(0.459)					
(β_3): Proportional x relative population share				-1.574					
				(0.582)					
(β_4): Proportional x relative population share-squared				1.815					
				(0.675)					
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.717	0.161	0.259	0.087	0.559	0.862	0.853	0.770	0.808
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.277	0.611	0.663	0.600	0.640	0.277	0.240	0.298	0.605
Predicted optimal size	0.221	0.275	0.257	–	0.332	0.198	0.327	0.201	0.350
Mean dependent	0.332	0.378	0.403	0.366	0.214	0.428	0.320	0.363	2.276
Observations	2,295	6,411	303	8,706	5,750	4,854	1,773	5,832	8,706
R-squared	0.669	0.704	0.735	0.693	0.675	0.681	0.702	0.728	0.675
Ethnicity-year Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country FE	NO	NO	YES	NO	NO	NO	NO	NO	NO
Country-year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. Political inclusion is a dummy variable that takes value one if the group in a country in a given year is neither powerless nor discriminated by the state. Columns (1) and (2) have sample for the periods 1946–1979 and 1980–2013, respectively. Column (3) runs the specification for the year 2013 only. Column (4) uses relative population share as the main explanatory variable. Relative population share is the ratio of population share of the group and the population share of the largest group in the country-year. Column (5) restricts the sample only to countries where the largest group is absolute majority. Column (6) restricts the sample to parliamentary democracies. Column (7) restricts the sample to only election years. Column (8) restricts the sample only to full democracies i.e. countries with a polity score ≥ 7. Column (9) uses power rank of a group as the dependent variable. Standard errors clustered at the country level are reported in parentheses.

Table F4—IV results: Full sample

Panel A: Second stage	Political inclusion	ln(Nightlight per area)
	(1)	(2)
β_1: Population share	5.823	5.307
	(1.660)	(8.759)
β_2: Population share - squared	-11.79	-8.388
	(4.994)	(20.48)
β_3: Proportional x Population share	-6.262	19.23
	(2.482)	(16.39)
β_4: Proportional x Population share - squared	18.88	-73.32
	(9.990)	(47.51)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.76	0.02
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.37	0.03
Predicted optimal size	0.247	0.316
Observations	5,047	2,226
R-squared	0.702	0.765
Ethnicity-year controls	YES	YES
Country-year FE	YES	YES
Kleibergen-Paap rk LM stat	2.42	1.89
Cragg-Donald Wald F stat	432.12	183.47
F stat (Proportional*Population share)	193.93	106.45
F stat (Proportional*Population share - squared)	543.95	325.80

Panel B: Country level	Proportional
Colonialist proportional	0.463
	(0.118)
Mean dependent	.450
Observations	1,309
R-squared	0.388
Region-year FE	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. Standard errors clustered at the country level are reported in parentheses.

Table F5—Group size distribution is not correlated with colonialist's system

	Colonialist Proportional			
	(1)	(2)	(3)	(4)
Minority Fractionalization	0.0220			
	(0.0215)			
Number of relevant minorities		0.00772		
		(0.0125)		
Largest group size			0.0789	
			(0.187)	
Absolute majority				0.0607
				(0.0908)
Observations	95	95	95	95
R-squared	0.220	0.214	0.212	0.215
Region-year FE	YES	YES	YES	YES

Notes: Country level data for 95 countries. Earliest year for which group size data is available is taken for each country.

Table F6—Weighting Replicates Main Results

	Political Inclusion	ln(Nightlight per area)
	(1)	(2)
β_1: Population share	3.756	9.300
	(1.143)	(3.741)
β_2: Population share - squared	-5.087	-18.75
	(3.161)	(8.889)
β_3: Proportional x Population share	-3.474	-13.42
	(1.584)	(7.062)
β_4: Proportional x Population share - squared	7.032	31.17
	(4.717)	(17.39)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.80	0.51
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.55	0.41
Predicted optimal size	0.369	0.248
Observations	8,706	3,469
R-squared	0.737	0.863
Country-year FE	YES	YES
Ethnicity-year controls	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. All the observations are weighted by the inverse of the number of relevant minorities used in each regression in the given country-year. Standard errors clustered at the country level are reported in parentheses.

Table F7—The pattern in MR is explained by geographical concentration

	Political inclusion		
	(1)	(2)	(3)
Population share	4.825	1.910	3.324
	(1.227)	(1.609)	(3.122)
Population share - squared	-9.276	-1.864	-4.437
	(3.955)	(5.917)	(6.917)
Proportional x Population share	-3.661		
	(1.721)		
Proportional x Population share - squared	9.106		
	(5.313)		
Concentrated x population share		4.811	-0.987
		(1.610)	(3.290)
Concentrated x population share - squared		-11.67	1.054
		(5.589)	(7.651)
Mean inclusion	0.366	0.447	0.265
Observations	8,706	4,830	3,876
R-squared	0.687	0.648	0.734
Ethnicity-year controls	YES	YES	YES
Country-year FE	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample.Column (1) shows the baseline result of column (4) in table 2. Column (2) uses only MR countries and column (3) uses only PR countries. Concentrated is a dummy variable that takes value one if the group has a well-demarcated settlement area in a country. Standard errors clustered at the country level are reported in parentheses.

Table F8—Optimal minority size smaller in countries with larger majority

	Political inclusion		
	(1)	(2)	(3)
β_1: Population share	3.741	5.130	7.531
	(1.297)	(1.814)	(2.159)
β_2: Population share - squared	-5.365	-7.732	-17.93
	(3.650)	(5.362)	(5.977)
β_3: Proportional x Population share	-2.607	-4.385	-7.838
	(1.787)	(2.220)	(2.553)
β_4: Proportional x Population share - squared	5.324	9.334	21.95
	(5.160)	(6.619)	(7.421)
$H_0: \beta_1 + \beta_3 = 0$ (p-value)	0.377	0.559	0.857
$H_0: \beta_2 + \beta_4 = 0$ (p-value)	0.991	0.640	0.540
Predicted optimal size	0.349	0.332	0.210
Mean inclusion	0.286	0.214	0.156
Observations	6,917	5,750	3,871
R-squared	0.685	0.675	0.732
Ethnicity-year controls	YES	YES	YES
Country-year FE	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Largest group size in column (1) ≥ 0.3, in column (2) ≥ 0.5, and in column (3) ≥ 0.7. Standard errors clustered at the country level are reported in parentheses.

Table F9—Panel Analysis Produces Similar Patterns

	Political Inclusion			ln(Nightlight per area)		
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1): Relative population share	1.547	1.513	2.127	9.835	31.62	18.96
	(1.749)	(1.290)	(1.741)	(9.883)	(15.07)	(12.96)
(β_2): Relative population share-squared	-1.020	-1.068	-1.487	-6.674	-21.80	-11.14
	(1.271)	(0.963)	(1.252)	(6.865)	(10.15)	(9.354)
(β_3): Proportional x relative population share	0.420	0.0847	-0.199	-3.760	-4.417	-4.998
	(0.925)	(0.725)	(0.762)	(1.650)	(1.527)	(2.091)
(β_4): Proportional x relative population share-squared	-0.207	0.367	0.933	6.624	9.091	9.362
	(0.825)	(0.651)	(0.781)	(4.671)	(3.856)	(5.922)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.19	0.19	0.28	0.52	0.06	0.27
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	0.17	0.37	0.66	0.99	0.13	0.85
Observations	9,289	2,979	1,695	3,748	1,194	648
R-squared	0.918	0.921	0.930	0.990	0.992	0.993
Ethnicity-country FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. Standard errors clustered at the country level are reported in parentheses. In columns (2) and (5), we keep every third year in which a group is present in the data, and in columns (3) and (6) we keep every fifth year in which a group is present in the data.

G List of Countries

S.no.	Country	Years	Minorities	Baseline	IV Strategy
1.	Albania	6	2	✓	
2.	Australia	17	2	✓	✓
3.	Bangladesh	21	3	✓	✓
4.	Belarus	1	2	✓	✓
5.	Belgium	41	2	✓	
6.	Benin	23	3	✓	✓
7.	Bhutan	6	2	✓	
8.	Bolivia	15	3	✓	
9.	Botswana	48	9	✓	✓
10.	Brazil	36	2	✓	
11.	Bulgaria	18	3	✓	
12.	Cambodia	4	4	✓	✓
13.	Canada	65	2	✓	✓
14.	Central African Republic	10	3	✓	✓
15.	Chile	49	2	✓	
16.	Colombia	41	2	✓	
17.	Congo	5	4	✓	✓
18.	Costa Rica	66	2	✓	✓
19.	Cote d'Ivoire	3	4	✓	✓
20.	Croatia	14	5	✓	✓
21.	Czechoslovakia	3	3	✓	✓
22.	Ecuador	44	3	✓	
23.	Estonia	22	3	✓	
24.	Ethiopia	10	8	✓	
25.	France	61	3	✓	
26.	Gabon	5	3	✓	✓
27.	Ghana	15	4	✓	✓
28.	Greece	51	3	✓	
29.	Guatemala	18	3	✓	✓
30.	Guinea-Bissau	10	2	✓	✓
31.	Guyana	17	2	✓	
32.	Honduras	32	2	✓	
33.	India	63	19	✓	✓
34.	Indonesia	15	11	✓	
35.	Iran	4	10	✓	
36.	Iraq	4	2	✓	

37.	Israel	47	4	✓	
38.	Italy	49	5	✓	
39.	Japan	24	3	✓	
40.	Kenya	12	7	✓	✓
41.	Kosovo	4	5	✓	✓
42.	Kyrgyzstan	8	3	✓	✓
43.	Laos	2	5	✓	✓
44.	Latvia	21	3	✓	
45.	Lebanon	37	10	✓	
46.	Liberia	14	5	✓	
47.	Macedonia	16	4	✓	✓
48.	Malawi	20	2	✓	✓
49.	Malaysia	15	4	✓	✓
50.	Mali	21	2	✓	✓
51.	Mauritania	1	2	✓	✓
52.	Mauritius	38	6	✓	✓
53.	Moldova	20	3	✓	✓
54.	Montenegro	8	5	✓	✓
55.	Mozambique	15	2	✓	✓
56.	Myanmar	11	10	✓	✓
57.	Namibia	15	11	✓	✓
58.	Nepal	19	4	✓	
59.	New Zealand	6	2	✓	✓
60.	Nicaragua	24	3	✓	
61.	Nigeria	22	5	✓	✓
62.	Pakistan	17	7	✓	✓
63.	Panama	13	4	✓	
64.	Peru	44	3	✓	
65.	Philippines	36	3	✓	
66.	Poland	23	4	✓	✓
67.	Romania	18	3	✓	
68.	Russia	7	38	✓	
69.	Serbia	7	6	✓	✓
70.	Sierra Leone	20	3	✓	✓
71.	Singapore	17	3	✓	✓
72.	Slovenia	22	7	✓	✓
73.	South Africa	20	13	✓	✓
74.	Spain	36	4	✓	
75.	Sri Lanka	62	3	✓	✓
76.	Sudan	7	12	✓	✓
77.	Switzerland	67	2	✓	

78.	Tanzania	19	4	✓	✓
79.	Thailand	23	3	✓	
80.	Turkey	45	2	✓	
81.	Uganda	5	5	✓	✓
82.	Ukraine	11	4	✓	✓
83.	United Kingdom	68	6	✓	
84.	United States	68	5	✓	
85.	Yugoslavia	7	5	✓	
86.	Zambia	18	6	✓	✓
87.	Zimbabwe	5	2	✓	✓

H Proofs of Propositions

H.1 Proof of Proposition 1

Consider the case of party A. Vote share of party A among members of group j is given by:

$$\pi_{A,j} = Pr[U(f_j^A) > U(f_j^B) + \delta + \sigma_{i,j}]$$

Assuming that $\psi \geq \phi_j$ for all j, we get:

$$\pi_{A,j} = \frac{1}{2} + \phi_j[U(f_j^A) - U(f_j^B) - \delta]$$

Party A will win elections if more than half the population votes for it. Probability of winning for party A is given by:

$$p_A = Pr[\frac{\sum_{j=1}^{3} n_j \pi_{A,j}}{\sum_{j=1}^{3} n_j} > \frac{1}{2}]$$

This can simply be written as:

$$p_A = \frac{1}{2} + \frac{\psi \sum_{j=1}^{3} \phi_j n_j (U(f_j^A) - U(f_j^B))}{\sum_{j=1}^{3} \phi_j n_j}$$

Thus, party A solves:

$$\max_{f_j^A \geq 0} p_A = \frac{1}{2} + \frac{\psi \sum_{j=1}^{3} \phi_j n_j (U(f_j^A) - U(f_j^B))}{\sum_{j=1}^{3} \phi_j n_j}$$

$$s.t. \quad \sum_{j=1}^{3} n_j f_j^A \leq S$$

Solving the above optimization problem gives the equilibrium condition in 1.

H.2 Proof of Proposition 2

In a K district majoritarian election, probability of winning for party A in constituency k, as can be seen from the result under proportional electoral system, is given by:

$$p_A^k = \frac{1}{2} + \frac{\psi \sum_{j=1}^{3} \phi_j n_j^k (U(f_j^A) - U(f_j^B))}{\sum_{j=1}^{3} \phi_j n_j^k}$$

Party A will win the election if it wins more than half the votes in more than half the districts. If both parties win in equal number of districts, then the winner will be chosen randomly. Party A solves the following optimization problem under majoritarian elections:

$$\max_{f_j^A \geq 0} p_A \quad s.t. \quad \sum_{j=1}^{3} n_j f_j^A \leq S$$

Since the parties are symmetric, in equilibrium, $p_A^k = \frac{1}{2}$ for all districts. Thus, given a district k, we denote the probability of winning in any other given district, with a slight abuse of notation, as p_A^{-k}. When K=2, Probability of winning can be written as:

$$p_A = p_A^k p_A^{-k} + \frac{1}{2}[p_A^k(1 - p_A^{-k}) + p_A^{-k}(1 - p_A^k)]$$

This can be simplified to:

$$= \frac{1}{2} p_A^k + \frac{1}{4}$$

And when K>2, probability of winning is:

$$
\begin{aligned}
p_A = & \sum_{i=\lfloor K/2 \rfloor}^{K-1} \binom{K-1}{i} p_A^k (p_A^{-k})^i (1-p_A^{-k})^{K-1-i} \\
& + \sum_{i=\lfloor K/2 \rfloor+1}^{K-1} \binom{K-1}{i} (1-p_A^k)(p_A^{-k})^i (1-p_A^{-k})^{K-1-i} \\
& + \frac{1}{2}[\frac{1+(-1)^K}{2}][\binom{K-1}{\lfloor K/2 \rfloor -1} p_A^k (p_A^{-k})^{(K/2)-1} (1-p_A^{-k})^{K/2} \\
& + \binom{K-1}{\lfloor K/2 \rfloor} (p_A^{-k})^{K/2} (1-p_A^{-k})^{(K/2)-1} (1-p_A^k)]
\end{aligned}
$$

This can be simplified to:

$$
\begin{aligned}
p_A = & \frac{1}{2^{K-1}}[\binom{K-1}{\lfloor K/2 \rfloor} p_A^k + \sum_{i=\lfloor K/2 \rfloor+1}^{K-1} \binom{K-1}{i}] \\
& + \frac{1}{2^K}[\frac{1+(-1)^K}{2}][(\binom{K-1}{\lfloor K/2 \rfloor -1} - \binom{K-1}{\lfloor K/2 \rfloor}) p_A^k + \binom{K-1}{\lfloor K/2 \rfloor}]
\end{aligned}
$$

Using this, we calculate:

$$
\frac{dp_A}{dp_A^k} = C(K) = (\frac{1+(-1)^{K-1}}{2}) \binom{K-1}{\lfloor K/2 \rfloor} \frac{1}{2^{K-1}} + (\frac{1+(-1)^K}{2}) \binom{K}{\lfloor K/2 \rfloor} \frac{1}{2^K}
$$

For the first order condition to the optimization problem, we need to calculate:

$$
\frac{dp_A}{df_j^A} = \sum_{k=1}^{K} \frac{dp_A}{dp_A^k} \frac{dp_A^k}{df_j^A}
$$

Substituting the expression for dp_A/dp_A^k, we can write this as:

$$
\frac{dp_A}{df_j^A} = C(K) \sum_{k=1}^{K} \frac{dp_A^k}{df_j^A}
$$

We can now easily solve the optimization problem to give the equilibrium condition given in 2. Consider the case where all groups are equally responsive to electoral promises i.e. $\phi_j = \phi$ for all j. Since $\sum_{j=1}^{3} n_j^k = 1$ for all k and $\sum_{k=1}^{K} n_j^k / n_j = K$

for all j, 2 can be simplified to:

$$U'(f_i^*) = U'(f_l^*) \quad \forall i, l$$

Now, consider the case where $n_j^k = n_j$ for all k. In this case, 2 can be simplified to:

$$\phi_i U'(f_i^*) = \phi_l U'(f_l^*) \quad \forall i, l$$

Both the above special cases indicate that when groups are evenly distributed across districts or when all groups are equally responsive to electoral promises, majoritarian elections give the same equilibrium political representation and per capita transfers as the proportional representation system.

H.3 Proof of Proposition 3

(a) When group 2 is concentrated, we have four types of constituencies based on the identity of groups residing in them: (1) Only group 1 and 3 reside (2) Only group 2 and 3 reside (3) Group 1, 2 and 3 all reside (4) Only group 3 resides. Densities D^m of constituency type m are:

$$D^1 = n_1^{1-\alpha} + n_3 \qquad D^2 = n_2^{1-\alpha} + n_3 \qquad D^3 = n_1^{1-\alpha} + n_2^{1-\alpha} + n_3 \qquad D^4 = n_3$$

Since constituencies have equal populations:

$$D^m a^m = \frac{1}{K} \quad \forall m$$

Where a^m is the area per consituency for each type m. Using this we get:

$$a^1 = \frac{1}{K(n_1^{1-\alpha} + n_3)} \quad a^2 = \frac{1}{K(n_2^{1-\alpha} + n_3)} \quad a^3 = \frac{1}{K(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)} \quad a^4 = \frac{1}{K(n_3)}$$

Number of consituencies K^m of each type can be calculated by dividing total area of occupied by all constituencies of a given type by a^m:

$$\begin{aligned}
K^1 &= K(n_1^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_3) \\
K^2 &= K(n_2^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_2^{1-\alpha} + n_3) \\
K^3 &= K(O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3) \\
K^4 &= K(1 - n_1^\alpha - n_2^\alpha + O \cdot min(n_1, n_2)^\alpha)(n_3)
\end{aligned}$$

Proportion of group i in constituency of type m n_i^m:

$$n_1^1 = \frac{n_1^{1-\alpha}}{n_1^{1-\alpha} + n_3} \quad n_1^2 = 0 \quad n_1^3 = \frac{n_1^{1-\alpha}}{n_1^{1-\alpha} + n_2^{1-\alpha} + n_3} \quad n_1^4 = 0$$
$$n_2^1 = 0 \quad n_2^2 = \frac{n_2^{1-\alpha}}{n_2^{1-\alpha} + n_3} \quad n_2^3 = \frac{n_2^{1-\alpha}}{n_1^{1-\alpha} + n_2^{1-\alpha} + n_3} \quad n_2^4 = 0$$
$$n_3^1 = \frac{n_3}{n_1^{1-\alpha} + n_3} \quad n_3^2 = \frac{n_3}{n_2^{1-\alpha} + n_3} \quad n_3^3 = \frac{n_3}{n_1^{1-\alpha} + n_2^{1-\alpha} + n_3} \quad n_3^4 = 1$$

Since, $U(f_j) = \log(f_j)$. Therefore, $U'(f_j) = \frac{1}{f_j}$. Similar to the proof of proposition 2, we can obtain the first order conditions at equilibrium as:

$$\begin{aligned}\gamma f_1 =& K\phi(n_1^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_3)(\frac{n_1^{-\alpha}}{\phi n_1^{1-\alpha} + \phi_3 n_3}) \\ &+ K\phi(O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)(\frac{n_1^{-\alpha}}{\phi(n_1^{1-\alpha} + n_2^{1-\alpha}) + \phi_3 n_3})\end{aligned}$$

$$\begin{aligned}\gamma f_2 =& K\phi(n_2^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_2^{1-\alpha} + n_3)(\frac{n_2^{-\alpha}}{\phi n_2^{1-\alpha} + \phi_3 n_3}) \\ &+ K\phi(O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)(\frac{n_2^{-\alpha}}{\phi(n_1^{1-\alpha} + n_2^{1-\alpha}) + \phi_3 n_3})\end{aligned}$$

$$\begin{aligned}\gamma f_3 =& K\phi_3(n_1^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_3)(\frac{1}{\phi n_1^{1-\alpha} + \phi_3 n_3}) \\ &+ K\phi_3(n_2^\alpha - O \cdot min(n_1, n_2)^\alpha)(n_2^{1-\alpha} + n_3)(\frac{1}{\phi n_2^{1-\alpha} + \phi_3 n_3}) \\ &+ K\phi_3(O \cdot min(n_1, n_2)^\alpha)(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)(\frac{1}{\phi(n_1^{1-\alpha} + n_2^{1-\alpha}) + \phi_3 n_3}) \\ &+ K\phi_3(1 - n_1^\alpha - n_2^\alpha + O \cdot min(n_1, n_2)^\alpha)(\frac{1}{\phi_3})\end{aligned}$$

$$n_1 f_1 + n_2 f_2 + n_3 f_3 = S$$

The equilibrium value of per capita private transfers to group 1:

$$f_1 = \frac{S\gamma f_1}{n_1\gamma f_1 + n_2\gamma f_2 + n_3\gamma f_3}$$

Calculating the denominator of the above expression using the first order conditions we get:

$$\begin{aligned}
n_1\gamma f_1 + n_2\gamma f_2 + n_3\gamma f_3 =& K(n_1^{\alpha} - O \cdot min(n_1,n_2)^{\alpha})(n_1^{1-\alpha} + n_3)(\frac{\phi n_1^{1-\alpha} + \phi_3 n_3}{\phi n_1^{1-\alpha} + \phi_3 n_3}) \\
&+ K(n_2^{\alpha} - O \cdot min(n_1,n_2)^{\alpha})(n_2^{1-\alpha} + n_3)(\frac{\phi n_2^{1-\alpha} + \phi_3 n_3}{\phi n_2^{1-\alpha} + \phi_3 n_3}) \\
&+ K(O \cdot min(n_1,n_2)^{\alpha})(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)(\frac{\phi(n_1^{1-\alpha} + n_2^{1-\alpha}) + \phi_3 n_3}{\phi(n_1^{1-\alpha} + n_2^{1-\alpha}) + \phi_3 n_3}) \\
&+ K(1 - n_1^{\alpha} - n_2^{\alpha} + O \cdot min(n_1,n_2)^{\alpha})(n_3)(\frac{\phi_3 n_3}{\phi_3 n_3}) \\
=& K(n_1 + n_2 + n_3) = K
\end{aligned}$$

When $n_1 < n_2$, we get from first order condition:

$$\frac{f_1}{S\phi} = \frac{\gamma f_1}{K\phi} = \frac{1-O}{w_1} + \frac{O}{w_3}$$

Where,

$$w_1 = \phi + \frac{(\phi_3 - \phi)(n_3)}{n_1^{1-\alpha} + n_3} \qquad w_3 = \phi + \frac{(\phi_3 - \phi)(n_3)}{n_1^{1-\alpha} + n_2^{1-\alpha} + n_3}$$

Derivative of w_1 and w_3 w.r.t. n_1:

$$w_1' = -\frac{(1-\alpha)(\phi_3 - \phi)n_3 n_1^{-\alpha}}{(n_1^{1-\alpha} + n_3)^2} \qquad w_3' = -\frac{(1-\alpha)(\phi_3 - \phi)n_3(n_1^{-\alpha} - n_2^{-\alpha})}{(n_1^{1-\alpha} + n_2^{1-\alpha} + n_3)^2}$$

As we can see $w_1' < 0$ and $w_3' < 0$ when $n_1 < n_2$. Therefore, $\frac{df_1}{dn_1} < 0$ in this case. When $n_1 \geq n_2$, we can rewrite the first order condition as:

$$\frac{f_1}{S\phi} = \frac{\gamma f_1}{K\phi} = \frac{1-Or}{w_1} + \frac{Or}{w_3}$$

Where,

$$r = (n_2/n_1)^{\alpha}, \quad r' = -\alpha r(\frac{1}{n_1} + \frac{1}{n_2}), \quad r \in [0,1]$$

Differentiating:

$$\frac{1}{S\phi}\frac{df_1}{dn_1} = \frac{-(1-Or)w_1'}{w_1^2} + Or'(\frac{1}{w_3} - \frac{1}{w_1}) + \frac{-(Or)w_3'}{w_3^2}$$

The first additive term on the R.H.S. is positive and the second and third terms are negative. It can be seen that $\frac{df_1}{dn_1}$ is strictly decreasing in O and is positive as O tends to 0. Therefore, to prove that the expression $\frac{df_1}{dn_1} < 0$ when $O > O^*$ for some $O^* \in (0,1)$, it is sufficient to show tha $\frac{df_1}{dn_1} < 0$ when $O = 1$. Substituting O =1 and rearranging the above expression, we need to show:

$$-\frac{(1-r)w_1'}{w_1^2} < -r'(\frac{1}{w_3} - \frac{1}{w_1}) + \frac{rw_3'}{w_3^2}$$

Substituting the values of w_1, w_2, w_1', w_3', r, r' and simplifying, our expression is reduced to:

$$z - \frac{1}{z} < \frac{\alpha(n_2/n_1+1)}{(1-\alpha)(1-(n_2/n_1)^\alpha)}$$

Where $z = 1 + \frac{\phi n_2^{1-\alpha}}{\phi n_1^{1-\alpha}+n_3}$

$$\implies \phi n_2^{1-\alpha}(2 + \frac{\phi n_2^{1-\alpha}}{\phi n_1^{1-\alpha} + \phi_3 n_3}) < \frac{\alpha(n_2/n_1+1)(\phi(n_1^{1-\alpha}+n_2^{1-\alpha}) + \phi_3 n_3)}{(1-\alpha)(1-(n_2/n_1)^\alpha)}$$

As the ratio $\frac{\phi_3}{\phi}$ increases, the above inequality will be satisfied more easily. Therefore, it is sufficient to show that weak inequality holds in the above expression when $\phi_3 = \phi$. Using this and rearranging, we now need to show:

$$(n_1^{1-\alpha}n_2^{1-\alpha})(2 + \frac{n_2^{1-\alpha}}{n_1^{1-\alpha}+n_3}) \leq \frac{\alpha(n_1+n_2)(n_1^{1-\alpha}+n_2^{1-\alpha}+n_3)}{(1-\alpha)(n_1^\alpha - n_2^\alpha)}$$

This can be rearranged to give:

$$n_1^{3-2\alpha}X + n_1^{2-\alpha}n_3 Y \leq 0$$

Where,

$$X = (2-3\alpha)q^{1-\alpha} - (2-\alpha)q - \alpha - \alpha q^{2-\alpha}$$
$$Y = (2-3\alpha)q^{1-\alpha} - (2-\alpha)q - \alpha - \alpha q^{2-\alpha} - \alpha(1+q+\frac{n_3}{n_1^{1-\alpha}}(1+q))$$
$$q = \frac{n_2}{n_1}, \qquad q \in [0,1]$$

As we can see, $Y < X$ and n_3 can take any value in $(0,1)$, therefore it is both necessary and sufficient to show that $X \leq 0$. In fact, it is sufficient to show that:

$$x(q,\alpha) = (2-3\alpha)q^{1-\alpha} - (2-\alpha)q - \alpha \leq 0 \qquad \forall q \in [0,1], \quad \alpha \in (0,1)$$

Since x is continuous in q, the above condition will hold if it can be shown to hold at the boundaries and at each critical point in (0,1). At the boundaries:

$$x(0,\alpha) = -\alpha < 0$$
$$x(1,\alpha) = -3\alpha < 0$$

At critical point q^*:

$$\frac{dx(q,\alpha)}{dq} = (1-\alpha)(2-3\alpha)q^{-\alpha} - 2 + \alpha = 0$$
$$\implies q^* = (\frac{(1-\alpha)(2-3\alpha)}{2-\alpha})$$

$\therefore q^* \in (0,1)$ only when $\alpha \in (0, \frac{2}{3})$. Substituting the value of q^* and simplifying we need to show:

$$x(q^*,\alpha) = \alpha((\frac{1-\alpha}{2-\alpha})^{\frac{1-\alpha}{\alpha}}(2-3\alpha)^{\frac{1}{\alpha}} - 1) \leq 0$$
$$\implies (\frac{2-\alpha}{1-\alpha})^{1-\alpha} \geq 2-3\alpha$$

Let $t = 1-\alpha$. Now we need to show:

$$y(t) = (1+\frac{1}{t})^t - 3t + 1 \geq 0 \qquad \forall t \in (\frac{1}{3}, 1)$$

Again, since y(t) is continuous in t, we only need to show that the above condition is true at the boundary points and at each critical point in $(\frac{1}{3}, 1)$. At the boundaries:

$$y(\frac{1}{3}) = 4^{\frac{1}{3}} > 0$$
$$y(1) = 0$$

At the critical point:

$$\frac{dy(t)}{dt} = (1 + \frac{1}{t})^t (ln(1 + \frac{1}{t}) - \frac{1}{1+t}) - 3 = 0$$

Substituting the value of $(1 + \frac{1}{t})^t$ in $y(t)$ and rearranging sides, we now need to show:

$$(3t - 1)(ln(1 + \frac{1}{t}) - \frac{1}{1+t}) \leq 3$$

Since $t \in (\frac{1}{3}, 1)$, therefore:

$$3t - 1 < 2 \qquad ln(1 + \frac{1}{t}) < ln(4) \qquad \frac{1}{1+t} > \frac{1}{2}$$
$$\therefore (3t - 1)(ln(1 + \frac{1}{t}) - \frac{1}{1+t}) < 2(ln(4) - \frac{1}{2}) = 1.77 < 3$$

This implies that $x(q^*, \alpha) \leq 0$. Thus, $x(q, t) \leq 0$. Therefore, when $n_1 \geq n_2$, $\frac{df_1}{dn_1} < 0$ if and only if $O > O^*$ for some $O^* \in (0, 1)$.

(b) When group 2 is dispersed, settlement areas of each group are:

$$A_1 = n_1^{\alpha} \qquad A_2 = 1 \qquad A_3 = 1$$

In this case, there are two types of constituencies: (1) Group 1, 2 and 3 all reside and (2) Only group 2 and 3 reside. Densities of constituencies are:

$$D^1 = n_1^{1-\alpha} + n_2 + n_3 \qquad D^2 = n_2 + n_3$$

Since the populations across the K constituency are equal, we can calculate area per constituency:

$$a^1 = \frac{1}{K(n_1^{1-\alpha} + n_2 + n_3)} \qquad a^2 = \frac{1}{K(n_1 + n_2)}$$

Number of constituencies of each type:

$$K^1 = K n_1^{\alpha}(n_1^{1-\alpha} + n_2 + n_3) \qquad K^2 = K(1 - n_1^{\alpha})(n_2 + n_3)$$

Group proportions in each constituency type:

$$\begin{aligned} n_1^1 &= \frac{n_1^{1-\alpha}}{n_1^{1-\alpha} + n_2 + n_3} & n_1^2 &= 0 \\ n_2^1 &= \frac{n_2}{n_1^{1-\alpha} + n_2 + n_3} & n_2^2 &= \frac{n_2}{n_2 + n_3} \\ n_3^1 &= \frac{n_3}{n_1^{1-\alpha} + n_2 + n_3} & n_3^2 &= \frac{n_3}{n_2 + n_3} \end{aligned}$$

We get first order conditions similar to (a). At equilibrium:

$$\gamma f_1 = K\phi(n_1^{\alpha})(n_1^{1-\alpha} + n_2 + n_3)\frac{n_1^{-\alpha}}{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3}$$

$$\begin{aligned} \gamma f_2 =& K\phi(n_1^{\alpha})(n_1^{1-\alpha} + n_2 + n_3)\frac{1}{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3} \\ &+ K\phi(1 - n_1^{\alpha})(n_2 + n_3)\frac{1}{\phi n_2 + \phi_3 n_3} \end{aligned}$$

$$\begin{aligned} \gamma f_3 =& K\phi_3(n_1^{\alpha})(n_1^{1-\alpha} + n_2 + n_3)\frac{1}{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3} \\ &+ K\phi_3(1 - n_1^{\alpha})(n_2 + n_3)\frac{1}{\phi n_2 + \phi_3 n_3} \end{aligned}$$

$$n_1 f_1 + n_2 f_2 + n_3 f_3 = S$$

Similar to the proof in (a), equilibrium per capita transfer to group 2 are:

$$f_1 = \frac{S\gamma f_1}{n_1\gamma f_1 + n_2\gamma f_2 + n_3\gamma f_3}$$

Calculating the denominator by substituting values from first order condition:

$$
\begin{aligned}
n_1\gamma f_1 + n_2\gamma f_2 + n_3\gamma f_3 =& K(n_1^{\alpha})(n_1^{1-\alpha} + n_2 + n_3)\frac{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3}{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3} + \\
& K(1 - n_1^{\alpha})(n_2 + n_3)\frac{\phi n_2 + \phi_3 n_3}{\phi n_2 + \phi_3 n_3} \\
=& K(n_1 + n_2 + n_3) = K
\end{aligned}
$$

Using this and the first order condition:

$$
\frac{f_1}{S\phi} = \frac{\gamma f_1}{K\phi} = \frac{n_1^{1-\alpha} + n_2 + n_3}{\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3}
$$

Differentiating and simplifying:

$$
\frac{1}{S\phi}\frac{df_1}{dn_1} = \frac{(\phi_3 - \phi)n_3((1-\alpha)n_1^{-\alpha} - 1)}{(\phi(n_1^{1-\alpha} + n_2) + \phi_3 n_3)^2}
$$

Since, $\phi_3 > \phi$, it follows:

$$
\begin{aligned}
\frac{df_1}{dn_1} > 0 \quad \text{if} \quad n_1 < (1-\alpha)^{\frac{1}{\alpha}} \\
\frac{df_1}{dn_1} < 0 \quad \text{if} \quad n_1 > (1-\alpha)^{\frac{1}{\alpha}}
\end{aligned}
$$

$\therefore$ There is an inverted-U shaped relation between n_1 and f_1^* and hence between n_1 and G_1^* with peak at $n_1^* = (1-\alpha)^{\frac{1}{\alpha}}$.