Department of Economics

DISCUSSION PAPER SERIES IN ECONOMICS

DP No. 04

Group Size and Political Representation Under Alternate Electoral Systems

August 2018

Sugat Chaturvedi
Sabyasachi Das

https://www.ashoka.edu.in/ecodp

Group Size and Political Representation Under Alternate Electoral Systems*

Sugat Chaturvedi — Indian Statistical Institute
Sabyasachi Das — Ashoka University

August, 2018

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 first establish a robust empirical regularity using an ethnicity-country level panel data comprising 438 ethno-country minority groups across 102 democracies spanning the period 1946–2013. We show that a minority group's population share has no relation with its absolute representation in the national executive under PR but has an inverted-U shaped relation under MR. The pattern is stable over time and robust to alternate specifications. The developmental outcomes for a group proxied using stable nightlight emissions in a group's settlement area follow the same pattern. We reproduce the main results by two separate identification strategies—(i) instrumenting colony's voting system by that of the primary colonial ruler and, (ii) comparing the same ethnicity across countries within a continent. We argue that existing theoretical framework with a two group set up is not able to explain this pattern. Our proposed model incorporates the spatial distribution of *multiple* minority groups in a probabilistic voting model and justifies these patterns as equilibrium behavior. The data further validate a critical assumption of the model and its additional comparative static results. Our work highlights that electoral systems can have important effects on power inequality across minorities, and consequently, their well-being.

JEL Classification: D72, D78, H11, J15
Keywords: Electoral systems, minorities, political representation, settlement patterns.

*Sugat Chaturvedi: Economics and Planning Unit, Indian Statistical Institute, Delhi. Email: `sugatc15r@isid.ac.in`. Sabyasachi Das: Economics Department, Ashoka University, NCR, India. Email: `sabyasachi.das@ashoka.edu.in`. We thank Sourav Bhattacharya, Diane Coffey, Dean Spears, Abhiroop Mukhopadhyay, Bharat Ramaswami, Tridip Ray along with conference participants at the 13th Annual Conference on Growth and Development at ISI Delhi (2017) and the Royal Economic Society Annual Conference in Sussex (2018) 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 minorities as they are understandably more vulnerable to exclusion. Sustained exclusion from power often breeds resentment among minority group members against the government which may, in turn, destabilize a democracy (Cederman, Wimmer and Min, 2010). Political representation, on the other hand, creates an institutional arrangement for the minorities to voice their interests and desires to the government.[1] Therefore, political representation of groups in general, and minorities specifically, may facilitate a more peaceful, stable and competitive democracy.

Importantly, political representation has always been unequal across minorities, even within a democracy. Our data show that only a third of minority groups have any representation in democracies during the post–World War II period and only about half of the variation in political representation can be explained by differences across countries. In contrast, the "majority" group in a country is almost always represented.[2]

In this context it may be useful to ask whether different electoral institutions create different incentives for political parties to represent some minorities but not others, and to what extent the size of a group matters for this consideration. In this paper we examine this issue by looking at how population share of minorities affects their probability of being represented in the national government, and how this relation 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.

We answer the research question in three steps. First, we use a recently released ethnicity level panel dataset comprising over a 100 democracies and establish a robust

[1]Political representation, for example, has been linked to various measures of positive political outcomes for minority groups. Previous works show that representation fosters trust and approval in government decision-making (Banducci, Donovan and Karp, 2004), engenders greater political participation among the 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).

[2]We define the largest group in any country to be the majority group in that country. This definition allows us to include countries which do not have a group with absolute majority. More than 80% of the majority groups in our sample indeed have absolute majority in their respective countries. Our results, both empirical and theoretical, do not change if we restrict attention to countries where the largest group has absolute majority.

causal relationship between group size and political representation of minorities. 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. Importantly, we use nightlight emissions per unit area in a group's settlement area as a proxy for (per capita) developmental outcome of the group to show that it also follows the same pattern. In the second step, we build a theoretical framework which models spatial distribution of *multiple* minority groups in a two party probabilistic voting model and justifies these patterns as equilibrium outcome. Finally, we go back to the data and provide evidence validating our model. We do this by verifying one critical assumption of the model and testing some of its additional comparative static results.

The aforementioned result is in sharp 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 U.S. cities and compare the welfare levels across the two electoral systems for minorities of varying sizes. In their model access to power for minorities never falls with population share within any electoral system and eventually increases in PR. We show that this apparently intuitive result gets modified in presence of multiple minorities. Since we are concerned with representation in national governments, the assumption of multiple minorities seems reasonable in our context. Our results imply that when more than two groups are present in a MR country, there is an "optimal" size of a minority group above which its political representation begins to fall. On the other hand, group size has no bearing on political representation in a PR system. Our contribution lies in showing the generality of this result across space and over time and providing a theoretical understanding that undergirds this empirical finding.

We now discuss the dataset and the empirical methods we use to establish the results before moving on to providing the explanation for it using our model.

For the empirical analysis, we combine several datasets, including the Ethnic Power Relations (EPR) dataset, to create a group-country level panel dataset. The final data we use spans the period 1946–2013 and comprises 438 ethno-country minority groups in 102 democratic countries. It contains various group level political outcomes and demographic details along with information about political systems of their respective countries for each year in our sample period. Most importantly for us, the EPR dataset provides a power status variable that codes each group's level of access to the national executive of the country in each year. There are six primary power statuses for any group, as coded by the data, indicating the degree of representation enjoyed by the group in the government. These are, in the descending order of power, *monopoly*, *dominant*, *senior partner*, *junior partner*, *powerless* and *discriminated by the state*. We define a group to be *politically included* if its power status is not coded as *powerless* or *discriminated*.[3] As indicated above, minorities on average are politically included only in about one-third of the cases compared to 94% for the majorities. Evidently, political inclusion in the government is an important marker of power for the minorities. We, therefore, use this indicator as our main dependent variable.

We do the analysis using a linear probability model and compare groups *within* a country-year observation.[4] This, we believe, is a strong specification which controls for all the potentially time varying observable and unobservable factors that are specific to a country (or a region or the world). This includes the fractionalization of groups, their political alliances, voter attitudes towards various groups and any political or economic crisis which may change the electoral strategies of the political parties. 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 0.260. The result survives a number of robustness checks, namely (i) doing the analysis across both halves of the time period separately (indicating that the relationship has not changed much over time) and only for the year 2013 which is the last year in the sample; (ii) using population share as a fraction of the population share of the largest group in the country-year observation as the main explanatory variable; (iii) restricting sample to countries where the largest group has an absolute majority; (iv) restricting sample only to parliamentary democracies; (v) restricting sample to election years only; (vi) restricting the sample only to countries that are classified in the Polity IV project as full democracies i.e. countries with a polity score of at least 7; (vii) using the original ordered power rank variable as the dependent variable.[5]

We then test if the developmental outcomes of the groups follow the same pattern. Since data on income or allocation of public resources is not available at the level of groups within countries, we use (logarithm of) nightlight luminosity per unit area in a group's settlement area as a proxy for per capita level of development of the group.

[3]Stated otherwise, a group is included if its power status is one of the following: monopoly, dominant, senior partner, or junior partner.

[4]The analysis therefore drops countries with one minority group.

[5]Power rank is coded as an integer from 1 to 6 where 1 corresponds to being being discriminated, 2 to being powerless and so on.

Using this as our dependent variable we show that it replicates the same pattern. This gives support to our claim that electoral systems do have a real bearing on how group size affects political representation of minorities, and consequently, their well-being.

However, interpreting this pattern as a causal relationship can be problematic. Firstly, the electoral system of a country is not exogenously given. Political actors in positions of power may strategically choose electoral systems that maximize their chances of winning, as Trebbi, Aghion and Alesina (2008) show in the context of US municipalites. This means that the electoral system at the time of democratization of a country, and even changes in it later may depend on existing distribution of power across the groups (Colomer, 2004; Persson and Tabellini, 2003). We attempt to address this endogeneity issue by looking at countries which were once colonized. Consistent with Reynolds, Reilly and Ellis (2008), we show that electoral systems of the former colonial rulers systematically predicts electoral system of the colonies post independence.[6] We, therefore, use this as an instrument for the electoral system of a colony. The exclusion restriction requires that the electoral system of the colonial ruler does not have a direct effect on the group politics in the colonies post-independence. The two-stage-least-squares estimates replicate our results for both political representation and nightlight luminosity.

One potential criticism of the IV specification is that it takes the group sizes within a country as exogenous. However, there might be unobserved characteristics of groups that can affect their population share as well as access to power in the national executive. For example, there might be cultural and geographic factors which could make a group economically successful, and affect its size and political power at the same time. To address this we use an alternative strategy where we compare a group present in more than one country within a continent and exploit the plausibly exogenous variation in its population share by using group-region-year fixed effects.[7] In this strategy, the variation comes primarily from a group falling unequally on the two sides of the national boundary.[8] This strategy restricts our sample to only those groups which are located in more than one country and consequently, our sample size falls drastically, by more than 80%. Even in the reduced sample, however, we find a statistically significant

[6] We restrict our sample to colonies which democratized not too long after gaining independence. We use a maximum lag of 30 and 50 years between gaining independence and democratization for our analysis. We do this to improve the predictive power of the first stage. See Sections 4.1 and 5.2 for a detailed discussion about this.

[7] A region as defined in our data is essentially a continent.

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

inverted U-shaped relation in MR and no relation in PR for political representation. The predicted optimal population share estimates under MR in both the identification strategies remain virtually identical. Importantly, the nightlight regressions have statistically significant coefficients in this specification and mirror the pattern observed for political representation.

The existing theoretical framework is unable to explain our empirical findings, as we have indicated above. We, therefore, propose a model of electoral competition between two parties in a probabilistic voting setup to contrast PR and MR elections. Importantly, we allow for multiple minorities in our model. Political parties promise representation to each group as platforms during elections. This determines the per capita private transfer of government resources targeted towards group members. This readily implies that in PR, where parties essentially maximize votes, all minorities irrespective of their size are equally represented. There are two opposing forces in action that deliver this result; though offering higher representation to the larger group gets a party more votes, it is cheaper for a party to attract a higher *share* of voters from a smaller group. The result follows from the observation that when representations are equal, these two forces balance each other out across groups.

In MR, on the other hand, parties want to win electoral districts and hence, they have to consider settlement patterns of groups across districts, i.e., over space. We postulate that area occupied by a minority group has a *concave* relationship with its population share.[9] This implies that for a majority group of 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. This is at the core of the inverted U-shaped result that emerges as the equilibrium in our model. We show evidence in favor of our concavity assumption in the data and test some additional comparative static results that the model delivers regarding the exact shape of the inverted-U relationship.

Our work is related to the large literature that examines 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. Empirically, some of the important outcome variables that have been studied with regard to effects of electoral systems are corruption (Kunicova and

[9]

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 their variation over space and time in specific countries (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 when electoral systems changed. Our analysis also highlights this concern and points out the exact nature of this influence, both theoretically and empirically. Moreover, while these papers are interested in the level of power enjoyed by minorities, our paper additionally 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 minority groups of differing sizes *within* a system. Our result, consequently, has important implications for power inequality across minorities. It suggests that PR creates lower power inequality between minority groups, and their welfare inequality presumably is also minimal as a consequence. 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 in MR countries while the mid-sized groups enjoy a greater access to the government.

The rest of the paper is organized as follows: Section 2 discusses the two electoral institutions that we consider and their evolution during the post-war period, section 3 elaborates on the various datasets used and summarizes the main variables, section 4 explains the empirical methodology and the identification strategy, and section 5 discusses the results. We then develop the model in section 6 and verify its assumptions and additional predictions in section 7. Finally, section 8 concludes.

2 Electoral Systems in Democracies

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 the MR system 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 and the party with majority of districts forms the government. Proponents of this system claim that it helps in formation of a strong and

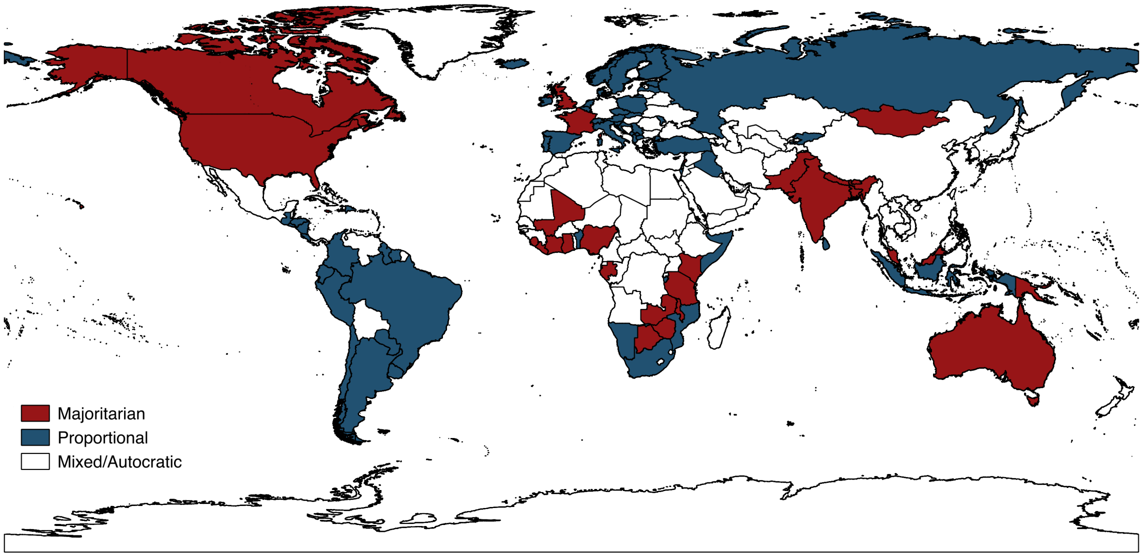


Figure 1: Electoral system distribution in 2013

accountable government (Norris, 1997). 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 elections in countries such as India, Nigeria and United Kingdom among others. Around 63% of country-year observations that follow MR have this system in our dataset.[10] In contrast, in the PR system, parties present list of candidates and seats are allocated to parties in proportion to their vote share in multimember districts. This reduces the disparity in vote share at the national level and the 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.[11]

Colonialism has played a major role in the choice of electoral system. Most of the countries that were once British and French colonies adopted the MR system while those that had been colonized by Belgium, Netherlands, Portugal and Spain adopted PR. Patterns of colonization and the effect of influential neighbors have resulted in a regional clustering of the systems as shown in figure 1.

[10] 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.

[11] 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.

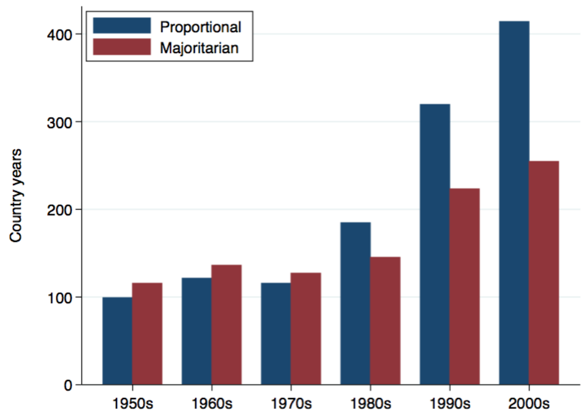


Figure 2: Electoral systems by decade

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 2, 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.[12] 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.[13]

The pattern of transition towards PR across the world has not been accompanied by a substantive political inclusion of minorities. This is shown in figure 3, in which we plot the proportion of minority groups in democracies in each power status category during 1946–2013. As figure 3 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 those who rule virtually alone (dominant and

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

[13]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.

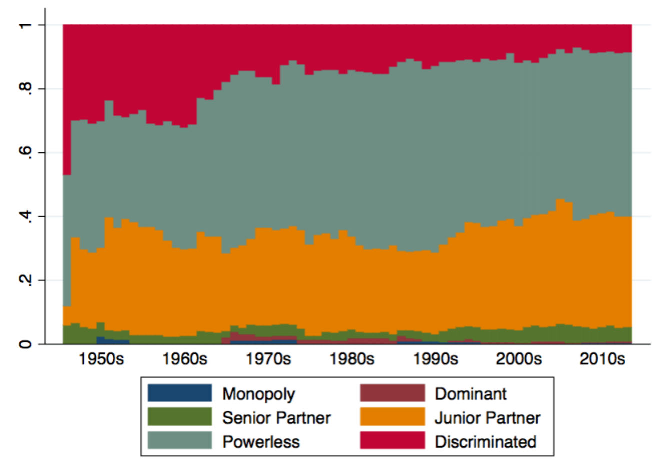


Figure 3: Minority power status over time

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.

3 Data Description

In this section, we describe the various data sources that we have put together for this project and discuss the main variables that concern us.

3.1 Data Sources

3.1.1 EPR Dataset

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."[14] (Cederman, Wimmer and Min, 2010) The dataset is concerned with

[14] 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.

groups that are politically relevant; a group is politically relevant if at least one political organization 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 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 the 20 groups - the second highest in our sample.[15] 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). Few other countries also have a relatively large number of groups - such as Indonesia, Lebanon, Botswana etc.[16] 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 between 2 to 14, with the average number of groups in the total sample being 4.6. We list in Appendix C the samples of countries used in our empirical exercises along with their number of minority groups and number of years in the sample, i.e., having a democratic regime.

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*).[17] The power ranking of the groups follows a fairly objective method, given by the degree

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

[16] Indonesia has 11 groups - Acehnese, Amboinese, Balinese, Bataks, Chinese (Han), Dayak, East Timorese, Gorontalos, Javanese, Madura, Makassarese and Bugis, Malay, Minahasa, Minangkabaus, Papuans, Sundanese and Ternate. Lebanon has 10 groups - Alawites, Armenian Catholics, Armenian Orthodox, Druze, Greek Catholics, Greek Orthodox, Maronite Christians, Palestinians, Protestants, Shi'as and Sunnis. Botswana has 9 groups - Birwa, Herero/Mbanderu, Kalanga, Kgalagadi, Mbukushu, San, Tswana, Tswapong, White and Yeyi.

[17] 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% of our sample and we do not consider them in our analysis.

and nature of presence of members of the group in the most important political positions in the national government. The details about group demarcation of the countries and the justification of the power rankings of each group is fully described in the official website of 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.[18]

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 by taking inputs from about one hundred country experts. This consultation period lasted about two years through multiple workshops. It was then followed by a final workshop where the final coding of attributes was decided after taking into account the inputs provided by the experts and accumulated knowledge available for the countries.

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 the datasets 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.

3.1.2 Electoral systems and polity characteristics

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

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

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 recently held election. We restrict our analysis to Majoritarian and Proportional systems. 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 the standard categorization based on the Polity IV score because we wish to look at all the countries which have competitive elections and have one of the two electoral systems of our interest. Our measure is a component in creating the Polity IV score, but there are other aspects of a regime such as existence of free and fair media 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.

3.1.3 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. 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 instrumental variable analysis which we describe later.

3.2 Summary statistics

Appendix table A1 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 system, whereas 56.13 percent have PR system. The countries with the MR system 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 the PR system. These differences, however, are not statistically significant at 10% level. On an average, the largest group comprises of 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 per cent). 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.

4 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 the population share of a group has any relationship with its probability of being included in the national executive and whether the relationship is different across the two electoral systems. The following is our preferred specification:

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

where I_{ict} is a dummy indicating whether the 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 the proportional electoral 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 relationship.

Given this specification, we compare groups *within* a country-year observation. This specification is therefore able to control for a host of observable and unobservable factors that vary at the level of country-year observations and may affect the relationship that we wish to estimate. We argue that two groups of the same size across two different countries or in the 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, that are present in the country, their explicit or implicit political alliances, electoral strategies of political parties, voters' attitudes towards the groups and any political, economic or social contingency which may affect all these factors in complex and unpredictable ways. It may also 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, therefore, 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 of interest 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 compare the same minority over time, by exploiting its temporal change in population share and political inclusion status. The specification could be written as:

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

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 β_1–β_4 are likely to have noisier estimates. Since this is not our preferred empirical specification, we discuss the results in the Appendix section B.

4.1 Identification I: IV Strategy

The baseline specification treats the electoral system of a country as exogenous. However, Trebbi, Aghion and Alesina (2008) show in the context of US municipalities that choice of electoral system can be endogenous to the size distribution of groups. If such concerns are true at the level of countries then our interaction terms in specification (1) would be misidentified. There is a small number of 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. We, therefore, propose to look at a subset of countries which had once been colonies. We use the electoral system of their primary colonial ruler at the time of the colony's independence as an instrument for the colony's current electoral system. Reynolds, Reilly and Ellis (2008) argue that a lot of the colonies adopted the electoral system of their colonial ruler. Therefore, this could potentially work as an instrument for our purpose. The exclusion restriction for this specification requires that the electoral system of the colonialists did not directly differentially affect the political power of groups of different sizes.

We keep in the sample only those colonies which democratized not too long after gaining independence from their colonial ruler. Some countries, such as Indonesia and Brazil, became dictatorships after gaining independence and remained so for many decades before becoming democracies. In such cases the colonial ruler's electoral system is going to matter much less for a country. For example, there are 7 countries which democratized at least 50 years after becoming independent.[19] Only one of them have the MR system 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 getting independent.[20] 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 the proportional system in the colony's year of independence. We then get the estimates of β_1–β_4 from specification (1) in the second stage regression.

4.2 Identification II: Boundary Design

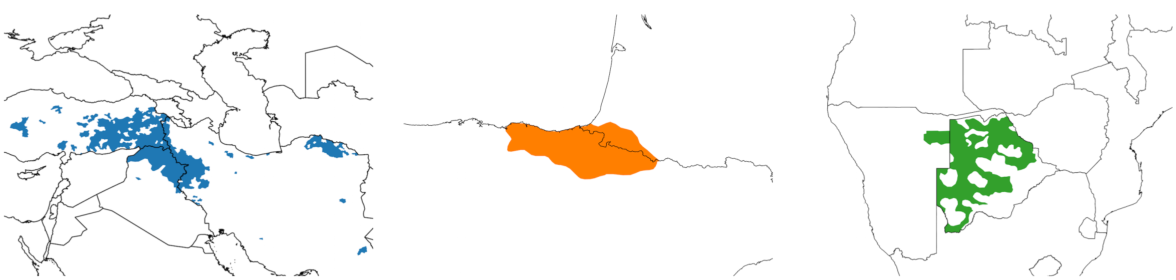
The IV strategy treats the population shares of groups as exogenous. However, there could be unobservable cultural and geographic factors which may affect the level of economic development of some groups which may impact both its size as well as its access to state power. In such cases the regression would suffer from omitted variable bias. Also, the IV specification allows us to identify the slopes of the relationship between group size and political representation across two types of countries. It, however, doesn't identify the intercept of the relationship.

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

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

We adopt a second identification strategy which attempts address the endogeneity issue and identifies both the slope as well as the intercept. For this identification, we notice that sometimes a group is present in more than one country and often those countries are neighbors. Examples include the Kurds who are present in both Turkey and Iran (panel A in figure 4), the Basques present in France and Spain (panel B) and the San group present in Botswana and Namibia (panel C) etc. Therefore, we exploit the differences in the sizes of the *same* group across those neighboring 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 neighboring countries comes from the group falling unequally across the two sides of a national boundary, and therefore, can be treated exogenously. To implement this strategy we restrict our sample to groups living in more than one country in the same region. This is because in our sample if two or more countries within a region share one group then those countries are always neighbors. We therefore estimate the following model:

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



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

Figure 4: 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.

where δ_{irt} denotes ethnicity-region-year fixed effects, error term ϵ_{ic} is double clustered at the level of group and country to adjust standard errors against potential autocorrelation within group and country. The coefficient θ now is the intercept of the relationship and β_1–β_4 are our other coefficients of interest, as before. The ethnicity-region-year fixed effect ensures that we compare the same group across countries within a region-year observation. This specification accounts for any region specific historical factor, including the prevalent political power of the group at the time of the creation of the countries, that may have been important for the differences in its sizes. It further controls for any time varying political factor in the region, observable or otherwise, that may affect the relationship.

5 Results

5.1 Baseline results

Table 1, column (4) shows the results from our baseline specification. The coefficient of population share is positive and significant at 1% level and coefficient of population share-squared is negative and significant at 5% level. The magnitudes of the coefficients imply that for the countries with MR system 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 when the population share is 0.260. The interactions of population share and its square with the proportional system dummy are statistically significant (at 5% 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 .325 and .960 respectively. This indicates that there is no relation between population share and political inclusion under the PR system. The results reported in columns (1) and (2) are with weaker specifications and include only the linear term for population share. Column (1) includes country and year fixed effects separately and doesn't include any control at the level of groups or country-year. Column (2) reports the same coefficients when these controls are added to the regression. In both cases we see that the relationship between population share and political inclusion is much weaker in PR compared to MR. However, we do find that with the linear specification there is a statistically significant positive relationship in PR system. This result, however, goes away once the squared terms are included to allow for non-linearity in the relationship, as we see in column (3). Importantly, in column (3) the dummy for proportional system has a positive and marginally significant coefficient. This suggests that very small minority groups presumably enjoy higher political representation under PR compared to the MR system.

The coefficients of ethnicity level controls in the same regressions are reported in appendix table A2. These coefficients are of the expected sign. The coefficient of peace years is positive and statistically significant at 1% level. An additional decade without

Table 1: Inverted U-shaped relation under MR and no relation under PR

	Political inclusion			
	(1)	(2)	(3)	(4)
β_1: Population share	2.839***	2.198***	4.405***	4.825***
	(0.450)	(0.279)	(1.239)	(1.227)
β_2: Population share - squared			-7.884**	-9.276**
			(3.883)	(3.955)
β_3: Proportional*Population share	-1.503***	-1.205**	-3.011*	-3.661**
	(0.559)	(0.489)	(1.687)	(1.721)
β_4: Proportional*Population share - squared			6.903	9.106*
			(5.159)	(5.313)
Proportional	0.216*	0.195	0.247*	
	(0.126)	(0.126)	(0.144)	
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	.000	.022	.231	.325
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	–	–	.774	.960
Predicted optimal size	–	–	0.279	0.260
Mean inclusion	0.367	0.367	0.366	0.366
Observations	9,304	9,294	9,294	8,706
R-squared	0.591	0.645	0.652	0.687
Group-year controls	NO	YES	YES	YES
Country-year controls	NO	YES	YES	NO
Country FE	YES	YES	YES	NO
Year FE	YES	YES	YES	NO
Country-year FE	NO	NO	NO	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. There are 438 ethno-country groups in 102 countries and 87 countries and 421 ethno-country groups for the period 1946–2013 in column (4). 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. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

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 in the group. 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. The coefficient is significant at 1% level. This shows that our measure of political representation, though based on subjective evaluation of experts, is nonetheless meaningful.

Table A3 reports the results of various robustness exercises we carry out to ensure that the result is not driven by any specific subsample 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. 2013. The coefficients here are quite similar to the column (1) coefficients. In column (4) we replace the main explanatory variable to the relative population share, i.e., the ratio of population share to the population share of the largest 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.

One may argue that our measure of political inclusion is subjective in nature, and therefore, any pattern observed in it may not reflect the actual well-being of groups. It is, therefore, important for us to show whether the same pattern is replicated when we look at an objective measure of developmental outcome of groups. However, data on developmental outcomes or allocation of public resources at the level of ethnic groups across countries is hard to get. We get around this problem by using nightlight intensity as a proxy for the level of economic development for groups which are settled in a geographically well demarcated region within a country.[21] Nightlight luminosity is now a well-documented and widely used proxy for the level of economic development of any geographic region, especially for subnational regions for which income data is not readily available across a wide range of countries.[22] Further, electricity in most countries

[21] The GeoEPR database provides GIS maps of the settlement areas for these groups (see Wucherpfennig, 2011).

[22] 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

is publicly provided and is an essential public good for any region within a country. Therefore, nightlight luminosity could also be thought of as a direct proxy for government allocation of resources in an area. We use (logarithm of) nightlight intensity per unit area as our dependent variable to test the specification (1).[23] The use of nightlight luminosity as our measure 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 (2) reports the results. Column (4) shows the results when the group population share is replaced by the relative population share as defined above. Both the columns show that the result for political inclusion is replicated with nightlight as outcome variable. The estimated population share with peak nightlight intensity is 0.21 which is similar to what we estimated for political inclusion. Moreover, we see in columns (1) and (3) that even with just linear terms we find that group size strongly predicts nightlight per unit area in MR countries, but there is very weak and statistically insignificant relationship between them in PR countries. This suggests that the patterns of political inclusion indeed have implications for the level of economic development of the groups.

5.2 Identification results

The IV results are reported in table 3. Panel B 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% level. Panel A reports the second stage results using political inclusion dummy and log of nightlight intensity per unit area as the 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 the 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 table shows that the relationships

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.

[23] We add 0.01 as a constant to nightlight intensity per area measure before taking the logarithm to include observations with zeros.

Table 2: Nightlight emissions follow the same patterns

	ln(Nightlight per area)			
	(1)	(2)	(3)	(4)
β_1: Population share	3.471**	11.02***		
	(1.490)	(3.840)		
β_2: Population share - squared		-24.49**		
		(11.30)		
β_3: Proportional*Population share	-2.763	-10.11		
	(1.787)	(6.103)		
β_4: Proportional*Population share - squared		24.29		
		(16.17)		
(β_1): Relative population share			1.438**	6.046***
			(0.662)	(1.808)
(β_2): Relative population share-squared				-6.653**
				(2.601)
(β_3): Proportional*relative population share			-1.167	-4.405
			(0.824)	(2.855)
(β_4): Proportional*relative population share-squared				4.954
				(3.804)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	0.53	0.86	0.59	0.52
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	–	0.99	–	0.58
Predicted optimal size	–	0.214	–	–
Observations	3,469	3,469	3,469	3,469
R-squared	0.812	0.816	0.811	0.818
Ethnicity-year controls	YES	YES	YES	YES
Country-year FE	YES	YES	YES	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. The dependent variable is logarithm of nightlight luminosity per unit area of groups which have well-demarcated settlement areas. Relative population share is the ratio of population share of the group and the population share of the largest group in the country-year observation. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

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 1, column (4)). Importantly, the Kleibergen-Paap rk LM statistic for the first stage regressions are high in all specifications, alleviating concerns related to underidentification. The F statistics for the two first stage regressions are also very large in magnitudes in each of the cases. Finally, for the sake of transparency, we report in appendix table A4 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 sign. The column (2) coefficients are also noisy. Importantly, the regressions don't pass the underidentification tests as the Kleibergen-Paap rk LM

Table 3: IV replicates main results

Panel A: Second stage	Democracy lag < 30 years		Democracy lag < 50 years	
	Political inclusion	ln(Nightlight per area)	Political inclusion	ln(Nightlight per area)
	(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*Population share	-4.332*	-41.69**	-4.049**	-36.53**
	(2.421)	(17.67)	(1.962)	(16.37)
β_4: Proportional*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: 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. 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. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

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*Population share	-8.269***	-58.72
	(2.686)	(35.96)
β_4: Proportional*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 are double clustered at the group and country level. *** p<0.01, ** p<0.05, * p<0.1.

statistics are low. This suggests that our sample restrictions are indeed useful in making our specification stronger.

We employ the second identification strategy as described in section 4.2 to test the robustness of our results. Table 4 reports the coefficients with political inclusion (column 1) and log nightlight intensity per unit area (column 2) as the dependent variables. Dimico (2016) 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. However, we 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. 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 the ones 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 the values 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 Proportional system dummy is positive and significant for political inclusion, suggesting that minorities of very small size get better represented in the PR system relative to the MR system. This is also consistent with the baseline result in column (3) of table 1. We plot the marginal effect of population share on political inclusion for the two identification methods in the appendix figure A9. The figures imply that mid-sized groups enjoy higher level of political inclusion under MR compared to PR.

6 Model

We now attempt to understand the rationale behind our empirical results. In this section we develop a probabilistic voting model of electoral competition, based on Persson and Tabellini (2002), and try to determine the conditions under which the patterns observed in the data will emerge as equilibrium outcomes.

6.1 Basic Setup

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 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 as:

$$U_j = U(f_j)$$

where f_j denotes per capita private transfers to the group j. The utility function is strictly increasing and strictly concave i.e. $U'(f_j) > 0$ and $U''(f_j) < 0$. To ensure interior solution we further take that $U'(f_j) \to \infty$ as $f_j \to 0$. 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 a 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. A group's promised representation in the government, G_j^h, determines how much per capita transfer voters of group j will get if party h comes to power. We denote it by

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

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 instead of G_j^h. 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 $\sigma_{i,j} \sim U[\frac{-1}{2\phi_j}, \frac{1}{2\phi_j}]$.

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 ideological bias of a member i of group j towards party B. ϕ_j is a measure of responsiveness of group j voters to private transfers determined through promised political representation by a party. Minority groups 1 and 2 are identical in their political responsiveness to transfers i.e. $\phi_1 = \phi_2 = \phi$ and group 3 is more responsive to transfers compared to the minorities i.e. $\phi_3 > \phi$. This is an important assumption and 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.[24] 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

[24] The African Americans in the USA, for example, are more attached to the Democratic Party and would presumably respond less to promises of transfer of resources by the Republican Party. The muslims in India, similarly, are historically associated with the Indian National Congress party and the BJP, the other national party, would find it difficult to attract them using promises of public resource allocation towards them.

the majoritarian system it is the probability of obtaining more than half of the electoral districts. We assume that in majoritarian system there are K electoral districts with equal population size. We denote by n_j^k the population share of group j relative to population in district k. Therefore,

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

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

6.2 Equilibrium Characterization

Since the parties are symmetric, we have policy convergence in equilibrium, i.e., parties choose the same equilibrium policy under both systems. 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:*

$$\phi_j U'(f_j^*) = \phi_l U'(f_l^*) \quad \forall\, j \neq l. \tag{4}$$

Proof: See Appendix D.1. ■

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 in 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 the MR system:

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:*

$$\phi_j U'(f_j^*) \sum_{k=1}^{K} \frac{n_j^k / n_j}{\sum\limits_{j'=1}^{3} \phi_{j'} n_{j'}^k} = \phi_l U'(f_l^*) \sum_{k=1}^{K} \frac{n_l^k / n_l}{\sum\limits_{j'=1}^{3} \phi_{j'} n_{j'}^k} \quad \forall\, j \neq l \tag{5}$$

Proof: See Appendix D.2. ■

We emphasize two aspects of the result above. Firstly, the characterization evidently implies that the equilibrium representation and transfer to groups under the MR system would generally depend on the population shares. Importantly, the transfer would depend on distribution of groups across electoral districts, suggesting that *settlement patterns* of groups across districts or over space would be important in determining the exact nature of the relationship between group size and transfer. Moreover, if all groups have the same responsiveness to transfers, i.e., if $\phi_1 = \phi_2 = \phi_3$, then equation (5) collapses to equation (4). Therefore, heterogeneity in responsiveness across groups, especially across majority and minority groups is critical for group size to matter in MR systems.

We can rewrite equation (5) 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.

Therefore, the proposition above states that in majoritarian system 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 the more responsive one, 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 a MR system wish to win electoral districts (as opposed to votes). Therefore, if a minority group 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 is going to play an important role in determining the nature of the comparative static exercise we perform in the following section.

6.3 Comparative Statics

Our empirical exercise estimated the relationship between representation and group size within a country-year observation, i.e., it compared multiple minorities within a country (in a given year) and exploited the variation in their group sizes to generate the result. Keeping parity with it, in this section we study the behavior of equilibrium representation and transfer in MR for minorities of differing group sizes. Specifically, we see how the equilibrium outcome variables change when we change the composition of n_1 and n_2 keeping the population share of the majority, n_3 fixed. Our main comparative static exercise will therefore look at the effect of changing n_1 by keeping n_3 constant. Now, any change in the composition of population shares of minorities at the national level would necessarily change 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 would be hard to comment on the nature of the comparative static result without specifying how changes in the population shares of groups relates to the consequent changes 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 we postulate that $A_j = n_j^{\alpha_j}$ for some $\alpha_j \geq 0$.[25] We assume that for group 3 (i.e., majority group) $\alpha_3 = 0$, or $A_3 = 1$, i.e., the majority group is dispersed all over the space in the country. For the groups 1

[25]Note that the same space can have presence of multiple groups, and therefore, the sum of A_js need not be one. If groups overlap over space, the sum of A_js would in fact be larger than one.

and 2, we consider two possibilities. In one case, we assume that $\alpha_1 = \alpha_2 = \alpha > 0$, 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 $\alpha_2 = 0$.[26] Importantly, for groups which are geographically concentrated, we have $\alpha < 1$, i.e., the area of settlement of a group has a concave relationship with its population share. This assumption will turn out to be important for the result we derive below. For mathematical simplicity we assume that group population is uniformly distributed across its area of settlement.

Now we consider dividing the country in K equal sized electoral districts. We note that in the case where both minorities are geographically concentrated, we will have three types of districts: (i) group 3 is present with only one minority group in the district, (ii) all the three groups present, and (iii) only group 3 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. Since the majority group is present everywhere, the proportion of this type of district will be 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 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 $n_1^* = (1-\alpha)^{\frac{1}{\alpha}}$.*

2. *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 $O^* \in (0, 1)$.*

[26] 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.

Proof: See Appendix D.3. ■

The result implies that when both groups are concentrated, the equilibrium representation and transfers of 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 getting larger 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.

An alternative way to think about it is to notice the fact that the concave relationship between population share and area occupied also 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 then they become less important everywhere. If they become too concentrated then their importance remain 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. It is apparent from our discussion that our main result for the MR system is critically dependent on the concavity assumption and the inverted-U shaped relationship is observed only for the minorities which are geographically concentrated. We now go back to our data to verify whether this indeed is true.

7 Validation of the Model

In this section, we first empirically verify one key parameter restriction of the model that we need for our main result. Proposition 3 requires the minority groups' settlement area to be inelastically related to their population shares. To test this assumption we run the following specification:

$$\ln S_{ict} = \alpha \ln n_{ict} + \gamma X_{ict} + \delta_{ct} + \epsilon_c \tag{6}$$

where S_{ict} is the settlement area of a group i which is geographically concentrated in country c in year t and n_{ict} is the population share that group. α therefore 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. Therefore, we can estimate the equation (6). The results are reported in appendix table A8. Column (1) reports the main estimate of α to be 0.625. It is statistically significant and significantly lower than one at 1% level. This confirms our hypothesis. Further, 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.

The primary aim of the model is to justify the empirical pattern established in the Section 4 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:

$$Y_{ict} = \delta_{ct} + \eta_1 n_{ict} + \eta_2 n_{ict}^2 + \eta_3 C_{ict} * n_{ict} + \eta_4 C_{ict} * n_{ict}^2 + \gamma X_{ict} + \epsilon_c \tag{7}$$

Table 5: 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*Population share	-3.661**		
	(1.721)		
Proportional*Population share - squared	9.106*		
	(5.313)		
Concentrated*population share		4.811***	-0.987
		(1.610)	(3.290)
Concentrated*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. 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) replicates the baseline result of column (4) in table 1. 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 are clustered at the country level and reported in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

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 5 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 (1) on various sub-samples of the data where we vary the size of the majority group. The results are reported in table 6. 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.

Table 6: Predicted optimal minority size is 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*Population share	-2.607	-4.385*	-7.838***
	(1.787)	(2.220)	(2.553)
β_4: Proportional*Population share - squared	5.324	9.334	21.95***
	(5.160)	(6.619)	(7.421)
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	.377	.559	.857
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	.991	.640	.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. 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. Largest group size in column (1) ≥ 0.3, in column (2) ≥ 0.5, and in column (3) ≥ 0.7. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

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.

8 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. We empirically show robust causal evidence that in countries with the PR system, population share of a minority has no effect on its political representation, while in countries with MR the relationship is inverted U-shaped. We then provide a theoretical framework with a multiple minority group set up that generates the same equilibrium predictions. We finally validate the model by confirming a critical assumption that delivers the desired result and then verifying the model's additional comparative static results. Our results imply that under PR, group size inequality does not translate into inequality in the political representation of minorities and consequently, their welfare inequality would also be minimal. On the other hand, power inequality among minorities in countries with the MR system 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 the MR system. 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

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.

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

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.

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, A. 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.

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.

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.

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.

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.

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.

A Additional Figures and Tables

Table A1: 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. ***p<0.01, **p<0.05, *p<0.1.

Table A2: Inverted U-shaped relation under MR and no relation under PR

	Political inclusion			
	(1)	(2)	(3)	(4)
β_1: Population share	2.839***	2.198***	4.405***	4.825***
	(0.450)	(0.279)	(1.239)	(1.227)
β_2: Population share - squared			-7.884**	-9.276**
			(3.883)	(3.955)
β_3: Proportional*Population share	-1.503***	-1.205**	-3.011*	-3.661**
	(0.559)	(0.489)	(1.687)	(1.721)
β_4: Proportional*Population share - squared			6.903	9.106*
			(5.159)	(5.313)
Proportional	0.216*	0.195	0.247*	
	(0.126)	(0.126)	(0.144)	
Years peace		0.00437***	0.00409***	0.00415***
		(0.00154)	(0.00135)	(0.00130)
Aggregate settlement		0.556***	0.549***	0.541***
		(0.0997)	(0.110)	(0.114)
Statewide settlement		0.329	0.294	0.139
		(0.333)	(0.375)	(0.352)
Regional and urban settlement		0.195***	0.174**	0.170**
		(0.0740)	(0.0784)	(0.0789)
Urban settlement		-0.00516	0.0180	0.00905
		(0.0653)	(0.0663)	(0.0650)
Regional settlement		-0.0143	-0.0105	-0.00942
		(0.0517)	(0.0488)	(0.0483)
Migrant settlement		-0.146	-0.140	-0.150
		(0.197)	(0.195)	(0.195)
Transethnic-kin inclusion		0.00805	0.00421	0.000118
		(0.0434)	(0.0446)	(0.0477)
Transethnic-kin exclusion		0.103***	0.0897**	0.103***
		(0.0380)	(0.0347)	(0.0348)
Fraction largest religion		-0.145	-0.125	-0.108
		(0.109)	(0.109)	(0.105)
Fraction largest language		0.155**	0.193**	0.210***
		(0.0627)	(0.0737)	(0.0748)
Ethnic fractionalization		0.0282	0.0203	
		(0.0239)	(0.0251)	
Number of relevant groups		0.0146	0.0123	
		(0.0197)	(0.0197)	
Competitiveness of participation		0.00705	0.00848	
		(0.0158)	(0.0166)	
Constraints chief executive		-0.0149	-0.0169	
		(0.00960)	(0.0104)	
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	.000	.022	.231	.325
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	–	–	.774	.960
Predicted optimal size	–	–	0.279	0.260
Mean inclusion	0.367	0.367	0.366	0.366
Observations	9,304	9,294	9,294	8,706
R-squared	0.591	0.645	0.652	0.687
Country FE	YES	YES	YES	NO
Year FE	YES	YES	YES	NO
Country-year FE	NO	NO	NO	YES

Notes: Data is at the level of ethnicity-country-year. Only minorities are part of the sample. There are 438 ethno-country groups in 102 countries and 87 countries and 87 countries and 421 ethno-country groups for the period 1946–2013 in column (4). 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. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: 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***	6.063***	3.732***	6.903***	5.543***
	(2.016)	(1.027)	(0.928)		(1.814)	(2.174)	(1.367)	(1.979)	(1.784)
β_2: Population share - squared	-14.14*	-8.141***	-12.31***		-7.732	-14.79*	-5.714	-17.21**	-7.910
	(7.334)	(2.919)	(2.643)		(5.362)	(7.800)	(4.473)	(7.771)	(5.773)
β_3: Proportional*Population share	-6.674**	-2.745*	-4.838***		-4.385*	-7.941***	-3.949**	-6.577***	-4.972*
	(2.453)	(1.526)	(1.586)		(2.220)	(2.848)	(1.915)	(2.429)	(2.797)
β_4: Proportional*Population share - squared	17.33**	6.322	10.79**		9.334	23.86**	9.625	20.29**	11.81
	(8.306)	(4.472)	(4.404)		(6.619)	(9.429)	(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*relative population share				-1.574***					
				(0.582)					
(β_4): Proportional*relative population share-squared				1.815***					
				(0.675)					
$H_0 : \beta_1 + \beta_3 = 0$ (p-value)	.717	.161	.259	.087	.559	.358	.853	.770	.808
$H_0 : \beta_2 + \beta_4 = 0$ (p-value)	.277	.611	.663	.600	.640	.139	.240	.298	.605
Predicted optimal size	0.221	0.275	0.257	–	0.332	0.205	0.327	0.201	0.350
Mean dependent	0.332	0.378	0.403	0.366	0.214	0.419	0.320	0.363	2.276
Observations	2,295	6,411	303	8,706	5,750	3,967	1,773	5,832	8,706
R-squared	0.669	0.704	0.735	0.693	0.675	0.696	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) has 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 observation. 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 are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: IV results: Full sample

	Political inclusion	ln(Nightlight per area)
Panel A: Second stage		
	(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*Population share	-6.262**	19.23
	(2.482)	(16.39)
β_4: Proportional*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. 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 dependent variable in column (3) of Panel A is logarithm of nightlight luminosity per unit area of groups which have well-demarcated settlement areas. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: 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. p<0.01, ** p<0.05, * p<0.1.

Table A7: Weighting Replicates Main Results

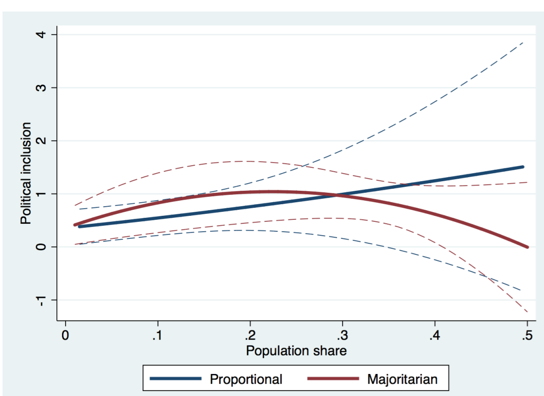
	Political Inclusion	ln(Nightlight per area)
	(1)	(2)
β_1: Population share	3.756***	9.300**
	(1.143)	(3.741)
β_2: Population share - squared	-5.087	-18.75**
	(3.161)	(8.889)
β_3: Proportional*Population share	-3.474**	-13.42*
	(1.584)	(7.062)
β_4: Proportional*Population share - squared	7.032	31.17*
	(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. 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. 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 are clustered at the country level in column (1). Standard errors are double clustered in column (2). *** p<0.01, ** p<0.05, * p<0.1.

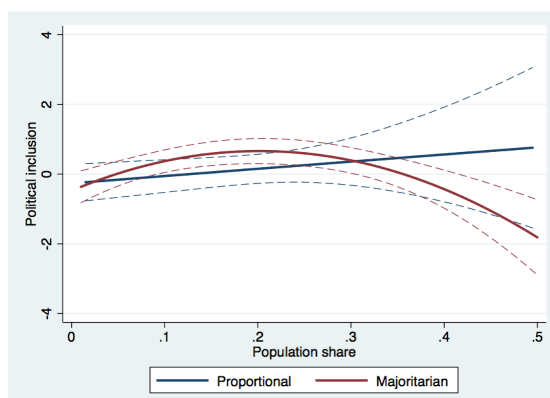
Table A8: Settlement Area Expands Inelastically: $\alpha < 1$

	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)	.001	.007	.005
Mean dependent	10.140	10.006	9.783
Observations	6,665	5,946	4,357
R-squared	0.792	0.779	0.742
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 are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.



(a) IV Estimation



(b) Same Group Across Countries

Figure A9: Marginal Effect of Group Size on Political Inclusion

B Panel Analysis

We report the results of specification (2) in table B1. We take relative population share as the independent variable to control for change in population share of the majority group as a consequence of change in population share of a minority. 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 in section 4.

Table B1: 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*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*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. 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. Standard errors are clustered at the country level. *** p<0.01, ** p<0.05, * p<0.1.

C List of Countries

S.no.	Country	Years	Minorities	Baseline	IV Strategy	FE Strategy
1.	Albania	6	2	✓		✓
2.	Argentina	43	1			✓
3.	Australia	17	2	✓	✓	
4.	Bangladesh	21	3	✓	✓	
5.	Belarus	1	2	✓	✓	✓
6.	Belgium	41	2	✓		✓
7.	Benin	23	3	✓	✓	
8.	Bhutan	6	2	✓		
9.	Bolivia	15	3	✓		
10.	Botswana	48	9	✓	✓	✓
11.	Brazil	36	2	✓		✓
12.	Bulgaria	18	3	✓		✓
13.	Cambodia	4	4	✓	✓	
14.	Canada	65	2	✓	✓	
15.	Central African Republic	10	3	✓	✓	
16.	Chile	49	2	✓		
17.	Colombia	41	2	✓		✓
18.	Congo	5	4	✓	✓	
19.	Costa Rica	66	2	✓	✓	✓
20.	Cote d'Ivoire	3	4	✓	✓	
21.	Croatia	14	5	✓	✓	✓
22.	Czechoslovakia	3	3	✓	✓	✓
23.	Djibouti	6	1			✓
24.	Ecuador	44	3	✓		
25.	El Salvador	29	1			✓
26.	Estonia	22	3	✓		✓
27.	Ethiopia	10	8	✓		✓
28.	France	61	3	✓		✓
29.	Gabon	5	3	✓	✓	
30.	Ghana	15	4	✓	✓	
31.	Greece	51	3	✓		✓
32.	Guatemala	18	3	✓	✓	
33.	Guinea-Bissau	10	2	✓	✓	
34.	Guyana	17	2	✓		✓
35.	Honduras	32	2	✓		
36.	India	63	19	✓	✓	
37.	Indonesia	15	11	✓		

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

81.	Switzerland	67	2	✓		
82.	Tanzania	19	4	✓	✓	
83.	Thailand	23	3	✓		
84.	Turkey	45	2	✓		✓
85.	Uganda	5	5	✓	✓	
86.	Ukraine	11	4	✓	✓	✓
87.	United Kingdom	68	6	✓		
88.	United States	68	5	✓		
89.	Venezuela	20	1			✓
90.	Yugoslavia	7	5	✓		✓
91.	Zambia	18	6	✓	✓	
92.	Zimbabwe	5	2	✓	✓	

D Proofs of Propositions

D.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).

D.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:

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

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.

D.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:

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

For simplicity, let $U(f_j) = \log(f_j)$. 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:

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

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

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

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

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,

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

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:

$$\begin{aligned}
x(0,\alpha) &= -\alpha < 0 \\
x(1,\alpha) &= -3\alpha < 0
\end{aligned}$$

At critical point q^*:

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

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

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

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:

$$K^1 = Kn_1^{\alpha}(n_1^{1-\alpha} + n_2 + n_3) \qquad K^2 = K(1 - n_1^{\alpha})(n_2 + n_3)$$

Group proportions in each constituency type:

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

Again, taking $U(f_j) = ln(f_j)$, we get first order conditions. 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 of proposition 3, 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 $n_1^* = (1-\alpha)^{\frac{1}{\alpha}}$.