Inequality and Voting Among Deprived Ethnic Groups: Evidence from India

H Zeynep Bulutgil[a] and Neeraj Prasad[b]

[a]University College London; [b]Tufts University

What are the conditions that determine the electoral success of parties that champion deprived ethnic groups? What is the impact of within group inequality on this outcome? Existing arguments focus on the role of institutions or the relationship between ethnicity and other social cleavages. This paper contributes to the second approach by studying the impact of within-group as well as between-group inequality on ethnic voting. We use elections to state legislatures within India to control for institutional and historical factors that may influence ethnic voting. Using data from the National Sample Survey we calculate inequality in consumption expenditure. We show that high within-group economic inequality among deprived ethnic groups hinders the electoral success of parties that champion these groups, whereas high between-group economic inequality has the opposite effect. Our findings also identify a potential causal mechanism (preference heterogeneity) that might link within-group inequality to ethnic voting.

What are the conditions that determine the electoral success of parties that champion deprived ethnic groups? How does economic inequality shape the electoral prospects of these parties? Theories of ethnic voting have typically focused on two types of explanations. The first type argues that institutions such as electoral and party systems, combined with the headcount and distribution of ethnic groups, provide the primary explanation for ethnic voting (Hechter 2004; Posner 2005; Chandra 2007). The second type places ethnicity in the broader socio-political context that includes the other competing social cleavages (Lijphart 1977; Dunning and Harrison 2010; Horowitz and Long 2016; Huber and Suryanarayan 2016; Huber 2017). Their starting point is the idea that where competing cleavages such as socio-economic ones cut-across rather than reinforce ethnic divisions, ethnicity becomes politically less salient and ethnic voting becomes less likely. This paper explores the relationship between two dimensions of cross-cuttingness and voting among subaltern ethnic groups in multi-ethnic democracies. The first dimension, between-group economic inequality, refers to relative differences in the mean wealth of ethnic groups; the second, within-group economic inequality, refers to the extent to which there is variation in the levels of wealth within each ethnic group. We show that both within-group and between-group inequality impact the electoral performance of an ethnic party but in different ways: high within-group inequality is detrimental to the ethnic party whereas high between group inequality is beneficial to the ethnic party.

Our paper makes theoretical contributions to the existing literature on cross-cutting cleavages and ethnic voting. We start from the well-established argument in comparative politics that overlapping identity and economic cleavages result in heightened probability of political mobilization (Taylor and Rae 1970; Lijphart 1977; Lipset and Rokkan 1967; Dunning and Harrision 2010). The causal logic of this argument, which goes through intervening mechanisms

such as social networks and policy preferences, predicts both the average income difference between groups and economic inequality within groups to shape political mobilization (Esteban and Ray 2008; Houle 2015; Huber 2017). From this perspective, it is surprising that between-group inequality has received much more attention as an underlying cause of ethnic mobilization than within-group inequality (Stewart 2008; Ostby, Nordas, and Rod 2009; Cederman, Weidmann, and Gleditsch 2011; Huber and Suryanarayan 2016; Bulutgil 2016). Our paper joins a small but growing number of studies that evaluate the impact of cross-cuttingness conceptualized as both within-group and between-group inequality on ethnic mobilization (Kuhn and Weidmann 2015; Houle, Kenny, and Park 2019; Higashijima and Houle 2018). Second, our findings provide critical evidence that speaks to an emerging theoretical debate in the literature. Currently, studies that take into account within-group inequality offer two types of findings. The first finding, based on studies that use data from India, is that within-group inequality does not have a significant impact on the extent to which political parties have unique ethnic bases of support (Huber and Suryanarayan 2016). The second finding, based on a study that uses cross-national surveys, is that within-group inequality does influence the success of ethnic parties (Houle, Kenny, and Park 2019). Our conclusions directly relate to this debate as they show that, based on extensive data on household consumption expenditures and elections in India, within-group inequality has a statistically and substantively significant impact on ethnic voting among underprivileged groups. This finding tilts the balance towards the idea that within-group inequality influences electoral performance of ethnic parties. Third, we also provide empirical evidence that supports a specific causal mechanism, preference heterogeneity measured as occupational heterogeneity, which links within-group inequality to ethnic voting.

Empirically, the paper makes several contributions. First, it uses the National Sample Survey (NSS), which in the context of political economy and ethnic voting is previously untapped, to measure inequality at the district level. This survey, the most comprehensive of its kind, has been conducted annually, with a thick round every five years, since 1950. The survey is designed to measure consumption, educational attainment, employment, and it directly reports household cash and non-cash consumption of more than half a million households. Comparable studies use asset ownership data from exit polls or surveys such as the Afrobarometer, World Values Survey, Latinobarometer, National Election Studies (Loknithi) to measure inequality (Houle, Kenny, and Park 2019; Huber and Suryanarayan 2016). Such opinion or exit polls are typically not designed to measure income and thus have a relatively limited coverage of asset ownership.[1] Use of household consumption enables a more accurate measurement of the key independent variables, inequality within and between ethnic groups. Second, the extensive survey also enables us to measure inequality at the district level and study elections to state legislative constituencies. Third, we improve upon the existing operationalization of between-group inequality. We measure this concept as the difference in consumption between members and non-members of a party's target ethnic group/s.[2] This represents an improvement over current methods which measure the average inequality between all ethnic groups and therefore are insensitive to the relative performance of individual ethnic groups (Huber and Suryanarayan 2016; Hero and Levy 2016).

The rest of the paper proceeds as follows. First, we introduce a model of the relationship between socio-economic inequality and "effective" headcount of economically deprived ethnic groups. Second, we describe our data and provide the conceptual and operational definition of our dependent and independent variables. We then move to the presentation and discussion of our results. In this section, we both test the primary predictions of the argument and evaluate a potential

---

[1] For example, Afrobarometer only records if the respondent has a radio, television, or a motor vehicle. Latinobarometer asks respondents to grade themselves on their perception of their relative standing in society on a scale of 1 to 10. Loknithi (NES) records ownership of 7 assets, 6 of which are related to means of transport or communication.

[2] For more detail on the measurement of between-group inequality, see the 'Data and Measurement' section.

causal mechanism (preference heterogeneity) that might link within-group inequality to ethnic voting. Finally, we conclude by explaining how our findings contribute to the literature on ethnic voting and mobilization.

Inequality and Ethnic Headcount

The literature on mobilization of ethnic groups by political parties in poor or developing democracies generally accounts for the salience of ethnicity by using two motivating frameworks. The first focuses on institutions such as electoral and party systems and emphasizes the prevalence of high information asymmetry in poor and less developed democracies (Chandra 2007; Hechter 2004; Birnir 2006; Posner 2005). In such contexts, ethnicity provides the most visible and easily identifiable marker for political parties to define 'in-groups' and 'out-groups'. Individual ethnic groups serve as fundamental units or building blocks. Whether a group is mobilized as a whole or in a coalition becomes a mere function of electoral arithmetic, composed of two determinants: an ethnic group's population share and the threshold needed to win elections (Chandra and Boulet 2012; Shertzer 2016; Posner 2005). The second framework places ethnicity in a broader socio-political context that includes the other competing social cleavages. This approach emphasizes the role of differentials in wealth, culture, social status, or political power between ethnic groups (Horowitz 1985; Huber and Suryanarayan 2016; Huber 2017; Dunning and Harrison 2010; Bulutgil 2016; Houle, Kenny, and Park 2019). Our paper argues that wealth differentials, both between and within ethnic groups, cause voter incentives to vary within a group. This variation causes the effective ethnic headcount, measured as those likely to vote for the corresponding ethnic party, to vary. In other words, we suggest that inequality affects voter-incentives, which in turn influences the effective ethnic headcount. Our framework applies particularly to subaltern or deprived ethnic groups.[3] This is because deprived ethnic groups mobilize to improve their social and economic standing. In order to improve a deprived group's socio-economic standing, a party will have to explicitly offer (or promise to offer) programmatic redistribution, large investments in public goods such as universal education, programs that promote socio-economic mobility, private transfers, or some form of patronage.[4] All of these mechanisms require some transfer of wealth from the rich to the poor.[5]

Figure 1(a) shows how an increase in inequality between members and non-members of an ethnic group increases ethnic headcount. In the hypothetical example, wealth is lognormally distributed, where mean wealth of deprived group is 5 units and the population mean wealth is 9 units.[6] The distribution of wealth is skewed so that the majority of the deprived group has lower than the group mean wealth endowment. Furthermore, a very small minority has more wealth than the average citizen. In the base scenario, 68% of those belonging to the deprived group have a wealth endowment lower than the population mean. Those with wealth greater than the population mean are more likely to be averse to programmatic redistribution, high taxation, or large investments in provision of public goods (Boix 2003; Campante and Ferreira 2007; Benabou 2000). Thus, if a party were to mobilize this hypothetical deprived group on the promise of redistribution or patronage, the 68% that are poorer than the average citizen are more likely to support the party; and the 32% who are richer than the average citizen are less likely to support the party. When inequality increases, the gap between the mean population wealth and the group mean population wealth will increase. If the gap were to increase such that the new mean population wealth is 15

---

[3] In another paper, we find that the impact of inequality on ethnic voting is different for privileged groups. See also Suryanarayan 2018.

[4] In contrast, privileged groups mobilize to maintain their status-quo social, economic, cultural, or political clout. As a result, wealth redistribution is unlikely to be a priority among such groups.

[5] Even patronage or private transfers, such as government jobs, are eventually financed by taxes paid by the elites of the privileged and the deprived groups.

[6] For more on skewed wealth distribution see Milanovic, Lindert, and Williamson 2011; Ansell and Samuels 2014.

units, the percentage of group population that has a wealth endowment lower than the population mean will increase from 68% to 78%. In this way, an increase in inequality between members and non-members of an ethnic group will bode well for an ethnic party due to the resultant increase in the effective ethnic headcount—defined as those likely to support redistributive measures.

(a) Between-Group Inequality.
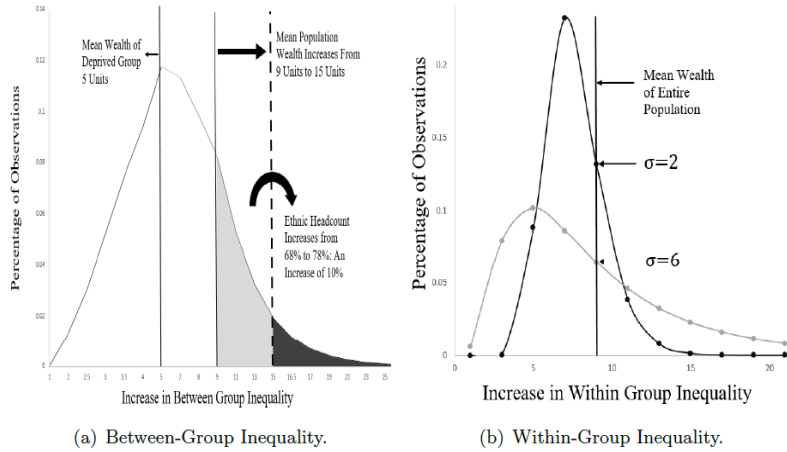
(b) Within-Group Inequality.

Figure 1: Effect of an Increase in Inequality on Ethnic Headcount

Figure 1(b) shows the effect of an increase in within-group inequality on effective ethnic headcount. It uses the same wealth distribution as used in the base case for Figure 1(a). The higher the standard deviation, the greater is the inequality within the group. When inequality is low and standard deviation equals 2 units, 78% of the ethnic population has wealth lower than the population mean. When inequality increases and standard deviation equals 6 units, 62% of the ethnic population has wealth lower than the population mean. Thus, when inequality within a deprived ethnic group increases, the percentage of co-ethnics who own more than the population mean increases, and effective ethnic headcount decreases.

Data and Measurement

In India, the lower house (or the sole house in unicameral states) of the state legislature is called the Vidhan Sabha or the Legislative Assembly. Members of the Vidhan Sabha are directly elected representatives with each member representing a state legislative constituency. Electoral institutions and electoral rules are uniform across all states. The elections are majoritarian, *first-past-the-post*, and are conducted by an independent nonpartisan body called the Election Commission of India (ECI). There are 4,210 state legislative constituencies in India.[7] Each state has a unique combination of national, state, and regional parties. By studying elections to state legislature we able to extract rich variation on the dependent variable and have a large sample size.

Inequality is measured using data from the National Sample Survey (NSS). The NSS data is often used to study well-being, poverty, consumption, unemployment, among others (Deaton 2003; Ravallion and Datt 2002). This paper uses data from three thick rounds: 61st NSS from 2004-2005, 66th NSS from 2009-2010, and 68th NSS from 2012. Besides consumption expenditure, the NSS also has data on the respondent's religion and caste category.

---

[7] Appendix A lists by each state the number of constituencies, districts, registered voters, national, state, and other regional political parties.

Unit of Analysis

We aggregate all electoral data from state legislatures to the district level. This is because consumption data from the National Sample Survey is only available at the district level. A large majority of state legislative constituencies are perfect subsets of a district, and districts are perfect subsets of a state.[8] Thus, we can make a comparison between the aggregate vote-share of party at the district level to inequality within and between ethnic groups at the district level. During the period of study, India had 640 districts in total, and the average population of each district was around 1.8 million.

The Dependent Variable: Vote-Share and Vote-Def

We intend to measure the effect of inequality on the electoral prospects of parties that champion deprived communities. In this direction, the paper uses two variants of the dependent variable: i) vote-share and ii) vote-def. Vote-Share measures a party's vote share at the district level. One disadvantage of this indicator is that it does not account for variation in the population-share of the party's target ethnic group. We tackle this problem in two ways. First, when using vote-share as the dependent variable, we include the group's population-share as a control variable. Second, we use a second indicator, vote-def, which incorporates the target groups' population-share into the dependent variable itself. The formula used to measure vote-def is given below;

$$\text{Vote-Def} = \text{Absolute (Vote-Share-Target Ethnic Group's Population Share)} \quad (1)$$

The underlying rationale is that: 1) When an ethnic party's vote-share is much smaller than its target groups' population-share, it implies that the party has not been able to get votes from its key supporters. 2) When an ethnic party's vote-share is much greater than its target groups' population-share, it implies that along with coethnics, people from other ethnic groups also vote for the party. Both (1) and (2) imply that citizens do not vote along their ethnic identity. 3) When an ethnic party's vote-share equals its target ethnic groups' population-share, it implies (or it is likely) that coethnics (or those belonging to the ethnic party's target group) vote along their ethnic identity.[9]

Two caveats are in order: First, we can only measure inequality by caste categories, such as Other Backward Caste, Scheduled Caste, Scheduled Tribe, or by religion, such as Muslims, Sikhs, Christians. Studies, such as Chandra (2007) and Thachil and Teitelbaum (2015), show that parties may mobilize along specific castes and sub-castes. Samajwadi Party, for instance, is known as a Yadav—Muslim alliance. This would introduce serious bias only if inequality among Yadavs is orthogonal to inequality among Other Backward Castes. Second, we cannot directly observe why individuals vote for co-ethnic parties but we do test for a potential motivating mechanism, preference heterogeneity, through its empirical implications. Summary statistics on vote-share and vote-def are provided in Appendix B.

---

[8] In contrast, a national parliamentary constituency may transcend a district, or in some instances, a district may include 2 or more parliamentary constituencies. This makes it difficult to attribute district data to parliamentary constituencies. In Jammu & Kashmir, and in some northeastern states, a few districts transcend state legislative constituencies. We have excluded such districts. Together, they account for less than 2% of India's population.

[9] When Vote-Share equals Target Groups Population-Share, there is of course the theoretical possibility that many non-coethnics voted for a given ethnic party. In India, particularly in the context of parties that champion deprived ethnic groups this is unlikely.

Key Independent Variables: Within and Between Group Inequality

We estimate within-group inequality using the Theil Index (Theil 1967; Mancini, Stewart, and Brown 2008; Hero and Levy 2016);

$$\text{WGI} - \text{Theil}_{j,d} = \frac{1}{n} \sum_i^n \frac{y_{i,j,d}}{\overline{y}_{d,j}} \log \frac{y_{i,j,d}}{\overline{y}_{d,j}} \tag{2}$$

where i is an individual or a household, d denotes a district, j stands for an ethnic group or an ethnic coalition, y is the consumption. For example, if a party targets two ethnic groups, e_1 and e_2, then $\overline{y}_{j,d}$ denotes the combined consumption of ethnic groups e_1 and e_2 in district d. In this way, we are able to measure inequality within multi-ethnic party's support base. Appendix B includes summary statistics for inequality by four groups, Muslim, Scheduled Tribe, Scheduled Caste, and Other Backward Caste.

We measure between-group inequality in two different ways. The first replicates the model described in Figure 1(a). In the model, between-group inequality is the difference in mean population-wealth and mean group-wealth. Likewise, we measure between- group inequality as the relative difference in mean district consumption and mean group consumption. Mathematically;

$$\text{BGI}_{j,d} = \text{abs} \frac{(\text{Mean Consumption}_d - \text{Group Mean Consumption}_{j,d})}{\text{Mean Consumption}_d} \tag{3}$$

where d denotes district and j denotes an ethnic group or an ethnic coalition. Additionally, we also use a modified version of Group Theil to measure inequality between members and non-members of an ethnic group or coalition. [10] The following adjusted formula is used:

$$\text{BGI} - \text{Theil}_{j,d} = P_j \frac{\overline{y}_{d,j}}{y_{i,d,j}} \log \frac{\overline{y}_{d,j}}{y_{i,d,j}} + P_{j'} \frac{\overline{y}_{d,j'}}{y_{i,d,j'}} \log \frac{\overline{y}_{d,j'}}{y_{i,d,j'}}$$

where p_j denotes group j's population-share.[11] Summary statistics for between-group inequality are included in Appendix B.

Control Variables

The following control variables are used: 1) Mean district consumption expenditure; 2) Percentage of population in a district that have not completed higher secondary education; 3) Percentage of households in a district that have access to electricity; 4) Estimated district-level population share of target ethnic group or coalition; 5) Indicator variable for regional party headquarters; 6) Margin of victory for constituency-level regression; and 7) and for some specifications an interaction term between within-group inequality and between-group inequality.

Identifying Parties That Champion Deprived Groups

To populate the list of ethnic parties, the paper starts with Thachil and Teitelbaum (2015) and Ziegfeld' (2018) categorization of parties into ethnic, non-ethnic, and regional. For each party we identify its target ethnic group. For example, the Bahujan Samaj Party is known to champion

---

[10] $GTheil = \sum_r^R P_r \frac{\overline{y}_r}{y} \log \frac{\overline{y}_r}{y}$, where Pr is the proportion of the population in group r, $\overline{y_r}$ is the average consumption of group r, and $\overline{y}$ is the average district consumption.

[11] A standard variance decomposition, such as the Theil Index, splits total inequality into within-group and between-group inequality. For this reason, we refrain from using both WGI-Theil and BGI-Theil in the same specification.

members of the Scheduled Caste (Chandra 2007). For each district, we estimate BSP's vote-share, population-share of the Scheduled Caste, inequality among members of the Scheduled Caste, inequality between members and non-members of the Scheduled Caste. In case of multi-ethnic parties, we measure inequality and population-share of all groups in the party's ethnic coalition. For example, Samajwadi Party (SP) is an example of an alliance between Other Backward Castes (OBC) and Muslims (Ziegfeld 2018). Accordingly, for Samajwadi Party, we measure in each district, inequality among the combined group 'OBC and Muslims', and inequality between members and non-members of the group 'OBC and Muslim'. A list of all parties with their target ethnic group and home state is included in Appendix C.

We use two potential definitions of parties that champion deprived groups in India. In the first and broader definition we i) include parties that champion OBC communities, Muslims as well as SC and ST communities but ii) exclude those that champion linguistic groups. The decision to exclude parties that champion linguistic groups is motivated by the reasoning that ethnic parties not only champion one or more ethnic groups but also exclude one or more ethnic groups (Chandra 2011). In India, state boundaries are drawn along spoken languages and most state legislative constituencies are linguistically homogeneous. Thus, a party that champions a linguistic group essentially appeals to the entire electorate and does not exclude any significant group. There is also an empirical basis to exclude parties that champion linguistic groups: without a well-defined 'out-group', between group ethnic inequality cannot be meaningfully measured.[12] In the second and narrower conceptualization, we only include parties that champion Scheduled Caste (SC) and Scheduled Tribe (ST) communities. This is because all other deprived communities tend to be relatively heterogeneous. For example, some OBC communities in southern India tend to be wealthier than the average citizen. To ensure that our results are not driven by variation among deprived communities, as a robustness check we also present our results including only SC and ST, which are consistently underprivileged across India.

Research Design, Results, and Analysis

This paper uses OLS with state and year fixed effects and robust standard errors, the following specification is used:

$$Y_{i,j,s,t} = \beta_0 + \beta_1 WGI_{i,j,s,t} + \beta_2 BGI_{i,j,s,t} + \beta'_3 \text{Control Variables}_{i,j,s,t} + \propto_s + \delta_t + \varepsilon_{i,j,s,t}$$

(5)

where, $Y_{i,j,s,t}$ indicates vote-share or vote-def for a party 'i' in district 'j' and in state 's' in year 't'. WGI stands for 'Within-group Inequality', BGI stands for 'Between-group Inequality.'

Table 1 shows the effect of within and between ethnic group inequality on the vote-shares of ethnic parties that champion economically deprived groups. WGI measures inequality within a party's target ethnic group/s. BGI measures inequality between members and non-members of a party's target ethnic group/s. Inequality and vote-share are measured at the district-level. Columns C1 to C6 use vote-share as the dependent variable and control for population share. Columns C7 and C8 use vote-def as the dependent variable and incorporate the target groups' population share in the measurement of the variable itself. Columns C4 includes an interaction term between within-group inequality and between-group inequality.

---

[12] With this definition we exclude some parties which had a history of OBC mobilization but over time started championing linguistic groups, such as Dravida Munnetra Kazhagam.

With vote-share as the dependent variable, the coefficient on within-group inequality is negative and statistically significant. This implies, all else equal, when inequality within a party's target ethnic group/s is low, the vote-share of the ethnic party is likely to be high. With vote-def as the dependent variable, the coefficient on within-group inequality is positive and statistically significant. This means when inequality within a party's target ethnic group/s is low, the difference between a party's vote-share and its target groups' population-share will be low. Figure 2(a) shows that as inequality increases from one standard deviation below the mean (Theil=0.03) to one standard deviation above the mean (Theil=0.19), the deficit between a party's vote-share and target population share increases from 23% to 29%.
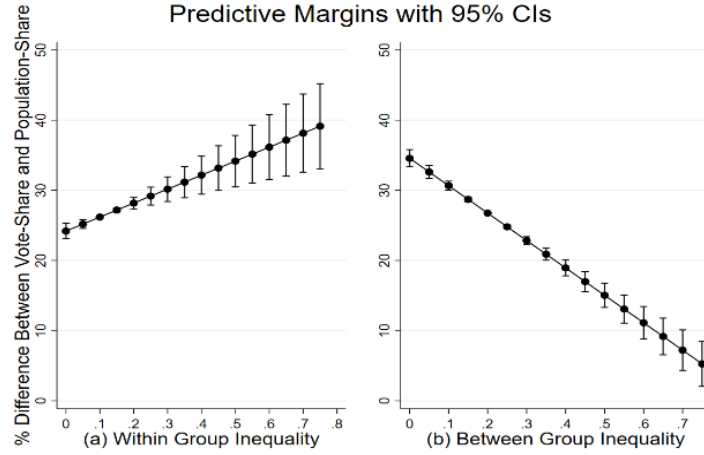

Figure 2: Ethnic Inequality and Performance of Ethnic Party (With Vote-Def as the Dependent Variable)

Table 1: Effect of Within and Between Group Inequality on Party Performance

Dependent Variable	Vote-Share						Vote-Def	
	C1	C2	C3	C4	C5	C6	C7	C8
	Theil	BGI	BGI-Theil	Theil & BGI	Vertical Inequality		Theil	BGI
Within-Group Inequality (WGI)	-0.029***			-0.050***		-0.037***	0.228***	
	(0.010)			(0.014)		(0.012)	(0.058)	
Between-Group Inequality (BGI)		0.011**	0.435***	0.004				-0.389***
		(0.005)	(0.135)	(0.005)				(0.030)
Interaction Term(WGI× BGI)				0.073**				
				(0.034)				
District Educational Attainment[a]	0.018**	0.022***	0.020**	0.018**	0.021**	0.018**	0.050	-0.009
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.034)	(0.033)
Mean District Consumption	-0.029	-0.029	-0.029	-0.034	-0.027	-0.024	0.161	0.208**
	(0.034)	(0.034)	(0.033)	(0.032)	(0.033)	(0.033)	(0.106)	(0.091)
Access to Electricity (% by district)	-0.011**	-0.012**	-0.016***	-0.011**	-0.012**	-0.012**	0.006	0.027
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.026)	(0.019)
Population Share (% by district)	0.027**	0.030**	0.025*	0.032**	0.026*	0.028**		
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)		
Regional Party HQ	0.192***	0.192***	0.192***	0.192***	0.192***	0.192***	-0.127***	-0.139***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.020)
Vertical Inequality (District)					-0.004	0.021		
					(0.014)	(0.017)		
R^2	0.520	0.520	0.523	0.521	0.520	0.520	0.048	0.186
N	4135	4135	4135	4135	4135	4135	4135	4135

Notes:* p<0.1, ** p<0.05, *** p<0.01. [a] measures the percentage that have *not* completed higher secondary education.

With vote-share as the dependent variable, the coefficient on between-group inequality is positive and statistically significant. In contrast, with vote-def, the coefficient on between-group inequality is negative and statistically significant. This implies, all else equal, when inequality

between members and non-members of a party's target group/s is high: vote-share of the corresponding ethnic party is likely to be high; and the difference between a party's vote-share and its target groups' population-share will be low. Figure 2(b) shows that as inequality increases from one standard deviation below the mean (BGI=0.01) to one standard deviation above the mean (BGI=0.39), the deficit between a party's target groups' population share and vote share decreases from 34% to 19%.
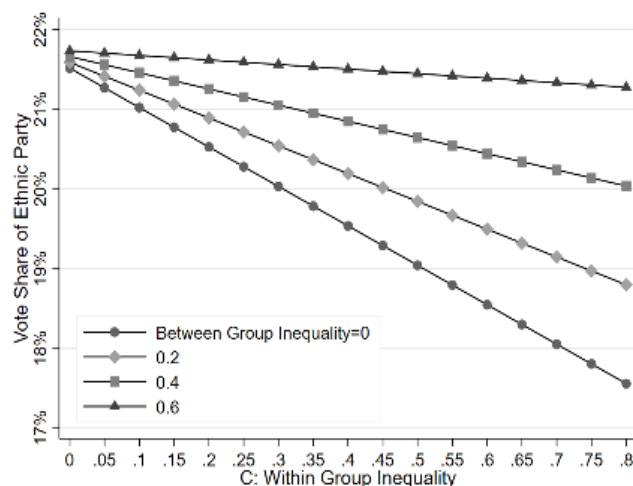

Figure 3: Ethnic Inequality and Performance of Ethnic Party (With Vote-Share as the Dependent Variable)

In Column C4 we include an interaction term between within-group inequality (Theil Index) and between group inequality (percentage difference in consumption). There will be some correlation between the two forms of inequality—either due to structural reasons, wherein the same underlying determinants cause both forms of inequality to exist; or due to measurement, wherein a standard variance decomposition breaks total inequality into two constituent components, within-group and between-group. Correlation due to the latter is mitigated by measuring within-group inequality using the Theil Index, and by measuring between-group inequality as the difference in district mean and group mean consumption. The joint-effect of within-group and between-group inequality can be understood Figure 3. For any value of between-group inequality, lower is the within-group inequality, higher is the expected vote-share of the ethnic party. For any given value of within-group inequality, higher is the between-group inequality, larger is the expected vote-share. A range of 0 to 0.6 for between group inequality covers 95% of the sample.

In Column C5, we include overall district-level inequality instead of inequality within a party's target ethnic group/s. We do not find a statistically significant relationship between district-inequality and vote-share. In Column C6, we include both overall district inequality and inequality within ethnic groups. We find that while inequality within ethnic groups has a statistically significant effect on vote-share, district level inequality has no effect. This suggests that the findings are indeed driven by inequality within a party's target ethnic group and not overall inequality within a district.

The national census, which measures population-share by caste-category and religion, is available only on a decennial basis. Furthermore, it does not report population numbers for an important caste category, Other Backward Caste (OBC). The thick rounds of the National Sample

Survey (NSS), on the other hand, are representative at the district-level, and include the caste category 'OBC.' We use the NSS data to construct estimates of ethnic population-shares at the district-level. We find positive correlation between population-share of target group and vote-share of corresponding ethnic party. We also include a binary variable to indicate when a party competes in its home state. In India, parties that champion deprived groups are generally regional parties and their organizational strength varies significantly from one state to the other. Therefore, we include a state-level indicator variable, 'Regional Party HQ.' This variable equals 1 when a party competes in its home state. We find that an ethnic party's vote-share is 15%-20% higher in its home state.

The coefficient on the variable measuring lack of education attainment is positive and statistically significant. A positive value indicates that parties that champion deprived groups are more successful when their target ethnic group's educational attainment is low. Furthermore, the coefficient on access to electricity is negative. In most parts of India, electricity is distributed by state owned utilities. Lack of access to electricity represents weak state-capacity to deliver goods and services (Demirguc-Kunt, Klapper, and Prasad 2017). In all, the finding that vote-share is high when 1) mean consumption is low, 2) educational attainment is low and 3) access to electricity is low, is consistent with the literature on clientelism and patronage in Indian politics. In poor democracies with high information asymmetries, it has been shown that citizens may vote for coethnics to gain patronage, access to government jobs, public goods and services (Chandra 2007; Ziegfeld 2018).

Table 2: Effect of Within and Between Group Inequality on Party Performance (Only SC/ST Parties)

Dependent Variable	Vote-Share					Vote-Def	
	C9	C10	C11	C12	C13	C14	C15
	Theil	BGI	Theil & BGI	Vertical Inequality		Theil	BGI
Within-Group Inequality (WGI)	-0.027***		-0.037**		-0.035***	0.216***	
	(0.009)		(0.015)		(0.012)	(0.057)	
Between-Group Inequality (BGI)		0.004	-0.000				-0.321***
		(0.005)	(0.006)				(0.029)
Interaction Term(WGI× BGI)			0.036				
			(0.035)				
District Educational Attainment[a]	0.014**	0.018**	0.014**	0.018**	0.015**	0.067**	0.012
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.026)	(0.027)
Mean District Consumption	-0.006	-0.002	-0.008	-0.001	0.000	0.185*	0.190**
	(0.030)	(0.030)	(0.030)	(0.030)	(0.029)	(0.096)	(0.088)
Access to Electricity (% by district)	-0.011*	-0.013*	-0.012*	-0.013**	-0.013**	0.002	0.018
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.024)	(0.017)
Population Share (% by district)	0.025*	0.026**	0.027**	0.024*	0.026**		
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)		
Regional Party HQ	0.169***	0.169***	0.169***	0.169***	0.169***	-0.110***	-0.125***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.023)	(0.016)
Vertical Inequality (District)				0.002	0.025		
				(0.016)	(0.021)		
R²	0.403	0.402	0.403	0.402	0.403	0.031	0.143
N	3181	3181	3181	3181	3181	3181	3181

Notes:* p<0.1, ** p<0.05, *** p<0.01. [a] measures the percentage that have *not* completed higher secondary education.

Table 2 uses the same specification as Table 1 but it applies only to a sub-sample consisting of political parties that champion Scheduled Caste or Scheduled Tribe communities. The results remain consistent. Specifically, for within-group inequality, our key empirical contribution, we find that WGI is negatively correlated with vote-share and positively correlated with vote-def. On between-group inequality, our findings are consistent when using vote-def as the dependent variable, but less so with vote-share. Figure 4a and 4b show the marginal effects of an increase in within-group and between-group inequality on the difference between vote-share of parties that champion Scheduled Caste or Scheduled Tribe communities and the population-share of the corresponding Scheduled Caste or Scheduled Tribe community. Figure 4, which applies to a

subsample of SC or ST parties, mirrors Figure 3, which applies to all ethnic parties that champion deprived groups. In Column C13, we find that overall district level inequality has no effect on vote-share of parties that champion SC or ST groups. In Column C14, we find that controlling for overall district inequality, inequality among SC or ST groups has a detrimental effect on vote-share of parties that champion SC or ST groups. Together they provide evidence to suggest that our findings are driven by inequality among SC/ST communities and not overall district level inequality.
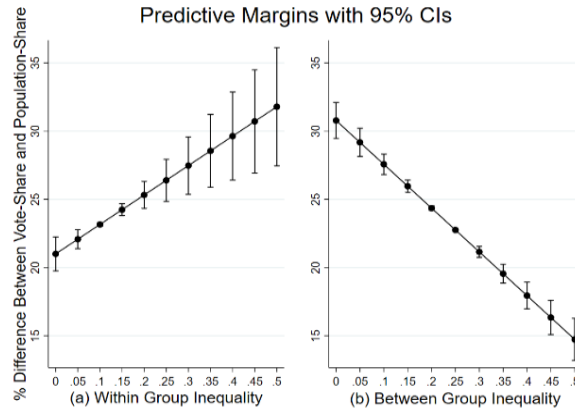

Figure 4: Performance of SC or ST Ethnic Party (With Vote-Def as the Dependent Variable)

Testing Preference Heterogeneity as a Causal Mechanism

Inequality may affect voting behavior through its impact on policy preferences of individual voters. When inequality within a group is high, members of that group would also be more diverse in terms of their policy preferences. For example, the rich may prefer lower taxes, while the poor may desire greater redistribution or investments in public goods. The resultant heterogeneity in preferences may stymie a party's ability to satisfy diverse interests within a group (Boix 2003; Campante and Ferreira 2007; Ansell 2010). Also, a social group provides a common setting for members to evaluate and form policy preferences (Dawson 2001; Lieberman and McClendon 2013). When inequality is high, ethnic social networks may be stratified by class. This may hamper development of homogenous preferences through group heuristics or through network effects (Houle, Kenny, and Park 2019). In this way inequality might affect a party's electoral performance through its impact on preference heterogeneity. In what follows we provide specific evidence to test preference heterogeneity as a mechanism.

We measure preference heterogeneity using heterogeneity of occupations within ethnic groups. When members of the same ethnic group are employed in different sectors and have varying skill levels, preference heterogeneity will result due to two reasons; i) economic inequality within groups leading to differences in policy preferences, ii) and stratification by occupational sectors, skill level, and economic class leading to weaker social networks. Figure 5 shows distribution of occupations among the Scheduled Caste in two states, Uttar Pradesh and Karnataka. Inequality in consumption among members of the Scheduled Caste is 50% higher in Karnataka as compared to Uttar Pradesh. The degree of heterogeneity in occupations is greater in Karnataka than in Uttar Pradesh. The Scheduled Caste in Uttar Pradesh are generally employed in occupations which do not require any formal schooling, such as being a domestic help or a farm labor (Category 9), a craftsman (Category 7), or a gardener (Category 6). In contrast, the Scheduled Caste in Karnataka are differentiated by both skill and occupational categories. Several among the Scheduled Caste in Karnataka belong to occupations that require university or college level education (Categories 1, 2, and 3). In this way, we use heterogeneity in occupations within an ethnic group to proxy for preference heterogeneity.
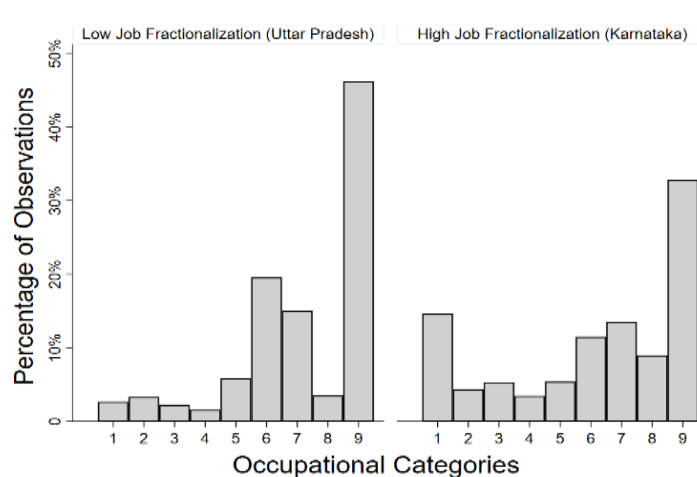
11

Figure 5: Heterogeneity in Occupations Among Members of the Scheduled Caste
Key: 1) Legislators, Senior Officials, and Managers; 2) Professionals; 3) Technicians and Associate Professionals; 4) Clerks; 5) Service Workers and Shop and Market Sales Workers; 6) Skilled Agricultures and Fisheries Workers; 7) Craft and Related Trade Workers; 8) Plant and Machine Operators and Assemblers; 9) Elementary Occupations.

Table 3 uses heterogeneity of occupations within a group as the key explanatory variable. It has two specifications where Table 3A applies to all ethnic parties that champion deprived groups and Table 3B applies to parties that champion SC or ST groups. The National Sample Survey codes each respondent's occupation using the 'National Classification of Occupations (NCO)'. This first level of NCO's classification includes the 9 sub-groups listed in Figure 5. We use this to estimate within-group fractionalization in job categories.[13] Columns C17 and C20 use vote-share as the dependent variables, while the rest use vote-def as the dependent variable.

Table 3: Within-Group Heterogeneity in Occupations

| Dependent Variable | A: All Parties | | | B: SC or ST Party | | |
| | Vote-Share | Vote-Def | | Vote-Share | Vote-Def | |
	C16	C17	C18	C19	C20	C21
Within-Group Fractionalization in Occupations	-0.012	0.161***	0.152***	-0.017**	0.146***	0.137***
	(0.007)	(0.026)	(0.027)	(0.006)	(0.025)	(0.026)
Between-Group Inequality (BGI)	0.010**	-0.349***	-0.349***	0.001	-0.289***	-0.288***
	(0.005)	(0.027)	(0.027)	(0.005)	(0.027)	(0.027)
District Educational Attainment[a]	0.019**	0.037	0.045	0.013*	0.051*	0.059**
	(0.008)	(0.034)	(0.033)	(0.008)	(0.028)	(0.028)
Mean District Consumption	-0.030	0.229**	0.232**	-0.005	0.213**	0.223**
	(0.034)	(0.091)	(0.091)	(0.032)	(0.087)	(0.087)
Access to Electricity (% by district)	-0.011**	0.006	0.004	-0.011	-0.000	-0.002
	(0.005)	(0.019)	(0.020)	(0.007)	(0.018)	(0.019)
Population Share (% by district)	0.032**			0.028**		
	(0.014)			(0.013)		
Regional Party HQ	0.193***	-0.138***	-0.138***	0.168***	-0.120***	-0.120***
	(0.024)	(0.019)	(0.019)	(0.037)	(0.019)	(0.019)
Within-Group Inequality (WGI)			0.073			0.078
			(0.045)			(0.048)
R^2	0.522	0.206	0.206	0.404	0.164	0.165
N	4130	4130	4130	3178	3178	3178

Notes:* p<0.1, ** p<0.05, *** p<0.01. [a] measures the percentage that have *not* completed higher secondary education.

---

[13] Within Group Fractionalization $= 1 - \sum_{i=1}^{9}$ Group Population Share$_i^2$, where i represents the nine job groups listed in Figure 3.

We find that the higher is the heterogeneity of occupations within a deprived group, i) the lower is the vote-share of their corresponding ethnic party, and ii) the higher is the difference between the vote-share and their target group's population-share.[14] Figure 6 shows that as heterogeneity of occupations within a group increases, the difference between a party's vote-share and the target groups' population-share also increases. This indicates that heterogeneity within a group is detrimental to the corresponding coethnic party's electoral performance. Columns C18 and C21 of Table 3 include both job fractionalization and inequality in consumption (WGI-Theil). When controlling for job fractionalization, inequality in consumption does not have a statistically significant relationship with the dependent variable. This finding suggests that preference heterogeneity measured as job fractionalization might partially work as an intervening mechanism that connects within group inequality to the performance of ethnic parties representing deprived groups.
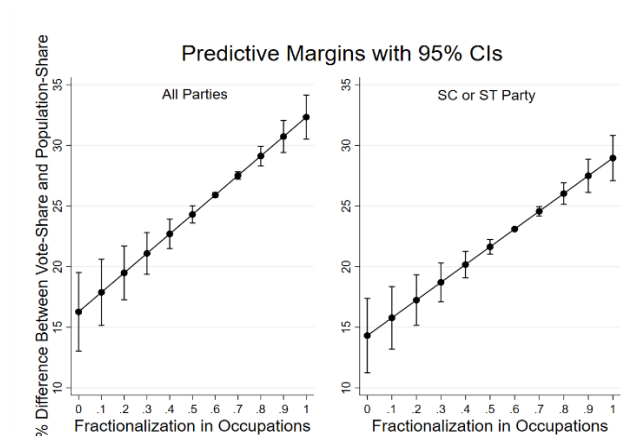

Figure 6: Performance of Ethnic Party and Within-Group Heterogeneity in Occupations

Additional Scenarios

Elections in India are majoritarian, *first-past-the-post*. Citizens would therefore use choices at the constituency-level as their frame of reference for voting. Aggregating vote-share at the district level does not capture heterogeneity between constituencies. In Online Appendix D, we use a multi-level mixed effects model to control for constituency-level heterogeneity.

Huber and Suryanarayan (2016) demonstrate that when ethnic inequality is high, the degree to which parties have unique ethnic bases of support increases. Applied to our sample, this would suggest that ethnic inequality may indirectly affect vote-share through the variable 'Regional Party HQ'—assuming that parties that champion deprived groups are more likely to be setup in states where they are likely to succeed. In Online Appendix E, we instrument for 'Regional Party HQ' with state-level ethnic population share and ethnic inequality using the 55th round of the National Sample Survey conducted in 1999. We find that both within-group inequality and between-group inequality shape the ethnification of party systems.

---

[14] These findings are statistically significant across all the specifications, except in C16, where it just falls short of conventional levels of significance at p=0.105.

Conclusion

The existing debate on ethnic voting largely takes place between the explanations that focus on institutional factors and the explanations that focus on the nature of and interaction between social cleavages. This paper joins a number of recent studies on ethnic voting by concentrating on the extent to which the interaction between socio-economic cleavages and ethnicity influences whether individuals vote on their ethnic identity. In doing so, the paper also follows the classical literature in comparative politics that argues that understanding the political implications of social cleavages requires one to locate these cleavages in the broader socio-political map of a given country (Dahl 1956; Lipset and Rokkan 1967; Taylor and Rae 1969; Lijphart 1977.)

Our findings have several noteworthy theoretical and empirical implications. First, we show that socio-economic cleavages within ethnic groups, which have generally not received as much scholarly attention as between group socio-economic differences, are in fact as important as between group economic inequality when it comes to explaining the electoral performance of parties that represent subaltern groups. Second, our findings provide critical evidence that speaks to an emerging debate in the literature between the studies that use cross-national evidence and find a relationship between within-group inequality and ethnic voting, and the studies that use data from India and find the opposite (Huber and Suryanarayan 2016; Houle, Kenny, and Park 2019). Our results, based on extensive data on India, support the argument that within-group inequality has a substantively and statistically important impact on ethnic voting. Third, we also provide evidence for "preference heterogeneity" as a causal mechanism that connects within-group inequality to voting behavior (Baldwin and Huber 2010; Dunning and Harrison 2010; Lieberman and McClendon 2013). Specifically, we demonstrate this mechanism by using heterogeneity in occupations as a proxy for diversity in within-group policy preferences and tightness of social networks.

Our paper also makes empirical contributions by using the National Sample Survey, which is more comprehensive, representative, and publicly accessible than the prevalently used National Electoral Survey, specifically for measuring inequality within and between-ethnic groups. Additionally, we measure between-group inequality while taking into account the ranking between different ethnic groups. Overall, our results suggest that the study of ethnic voting in particular and ethnic mobilization in general should take both between-group and within-group inequality into consideration. Our conclusions also open new avenues for future research. Specifically, to determine the portability of our findings, it would be desirable to conduct similar tests on within-group inequality in countries where ethnicity is not defined by historically existing caste differences.

References

Ansell, Ben W., and David J. Samuels. 2014. *Inequality and Democratization: An Elite- Competition Approach.* New York, NY: Cambridge University Press.

Baldwin, Kate, and John D. Huber. 2010. "Economic versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision." *American Political Science Review* 104, no. 4 (November): 644–662.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90, no. 1 (March): 96–129.

Birnir, Johanna Kristin. 2006. *Ethnicity and Electoral Politics.* Cambridge; New York: Cambridge University Press.

Boix, Carles. 2003. *Democracy and Redistribution.* Cambridge studies in comparative politics. Cambridge, UK; New York: Cambridge University Press.

Bulutgil, H. Zeynep. 2016. *The Roots of Ethnic Cleansing in Europe.* New York, NY: Cambridge University Press.

Campante, Filipe R., and Francisco H. G. Ferreira. 2007. "Inefficient lobbying, populism and oligarchy." *Journal of Public Economics* 91, no. 5 (June): 993–1021.

Cederman, Lars-Erik, Nils B. Weidmann, and Kristian Skrede Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Political Science Review* 105, no. 3 (August): 478–495.

Chandra, Kanchan. 2007. *Why Ethnic Parties Succeed: Patronage and Ethnic Head Counts in India.* Revised Edition. Cambridge: Cambridge University Press.

———. 2011. "What Is an Ethnic Party?" *Party Politics* 17 (2): 151–69.

Chandra, Kanchan, and Cilanne Boulet. 2012. "A Baseline Model of Change in an Activated Ethnic Demography." In Constructivist Theories of Ethnic Politics, edited by Kanchan Chandra, 229–276. Oxford University Press.

Chauchard, Simon. 2017. *Why Representation Matters: The Meaning of Ethnic Quotas in Rural India.* New York, NY: Cambridge University Press.

Dawson, Michael C. 2001. *Black Visions: The Roots of Contemporary African-American Political Ideologies.* Chicago: University of Chicago Press.

Dahl, Robert A. 1956. *A Preface to Democratic Theory.* Chicago, IL: University of Chicago Press.

Deaton, Angus. 2003. "Household Surveys, Consumption, and the Measurement of Poverty." *Economic Systems Research* 15 (2): 135–59.

Demirguc-Kunt, Asli, Leora Klapper, and Neeraj Prasad. 2017. "Measuring the Effectiveness of Service Delivery: Delivery of Government Provided Goods and Services in India." *World Bank Policy Research Working Paper No. 8207* (September): 1–55.

Dunning, Thad, and Lauren Harrison. 2010. "Cross-cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104, no. 1 (February): 21–39.

Dunning, Thad, and Janhavi Nilekani. 2013. "Ethnic Quotas and Political Mobilization: Caste, Parties, and Distribution in Indian Village Councils." *American Political Science Review* 107 (1): 35–56.

Esteban, Joan, and Debraj Ray. 2008. "On the Salience of Ethnic Conflict." *American Economic Review* 98, no. 5 (December): 2185–2202.

Hechter, Michael. 2004. "From Class to Culture." *American Journal of Sociology* 110 (2): 400– 445.

Hero, Rodney E., and Morris E. Levy. 2016. "The Racial Structure of Economic Inequality in the United States: Understanding Change and Continuity in an Era of Great Divergence." *Social Science Quarterly* 97, no. 3 (September): 491–505.

Higashijima, Masaaki, and Christian Houle. 2017. "Ethnic Inequality and the Strength of Ethnic Identities in Sub-Saharan Africa." *Political Behavior*.

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict.* Berkeley: University of California Press.

Horowitz, Jeremy, and James Long. 2016. "Strategic voting, information, and ethnicity in emerging democracies: Evidence from Kenya". *Electoral Studies* 44:351–361.

Houle, Christian. 2015. "Ethnic Inequality and the Dismantling of Democracy: A Global Analysis." *World Politics* 67 (3): 469–505.

Houle, Christian, Paul Kenny, and Chunho Park. 2019. "The Structure of Ethnic Inequality and Ethnic Voting." *Journal of Politics* 81(1):187-200.

Huber, John D. 2017. *Exclusion by Elections: Inequality, Ethnic Identity, and Democracy*. Cambridge University Press.

Huber, John D., and Pavithra Suryanarayan. 2016. "Ethnic Inequality and the Ethnification of Political Parties: Evidence from India." *World Politics* 68 (1): 149–188.

Jensenius, Francesca R. 2017. *Social Justice through Inclusion: The Consequences of Electoral Quotas in India*. Oxford, New York: Oxford University Press.

Kedar, Orit, and W. Phillips Shively. 2017. "Introduction to the Special Issue." *Political Analysis* 13 (4): 297–300.

Kuhn, Patrick M., and Nils B. Weidmann. 2015. "Unequal We Fight: Between- and Within-Group Inequality and Ethnic Civil War." *Political Science Research and Methods* 3 (3): 543–68.

Lieberman, Evan, and Gwyneth McClendon. 2013. "The Ethnicity Policy Preference Link in Sub-Saharan Africa." *Comparative Political Studies* 46, no. 5 (May): 574–602.

Lijphart, Arend. 1977. *Democracy in Plural Societies: A Comparative Exploration.* Yale University Press.

Lipset, Seymour Martin, and Stein Rokkan. 1967. "Cleavage Structures, Party Systems, and Voter Alignments." In *Party Systems and Voter Alignments. Cross National Perspectives,* 1–67. New York: Free Press.

Logan, John R., Wenquan Zhang, and Richard D. Alba. 2002. "Immigrant Enclaves and Ethnic Communities in New York and Los Angeles." *American Sociological Review* 67 (2): 299–322.

Mancini, Luca, Frances Stewart, and Graham K Brown. 2008. "Approaches to the Measurement of Horizontal Inequalities." In *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies,* 2008 edition, edited by Frances Stewart. Pal- grave Macmillan.

Milanovic, Branko, Peter H. Lindert, and Jeffrey G. Williamson. 2011. "Pre-Industrial Inequality." *The Economic Journal* 121, no. 551 (March): 255–272.

Ostby, Gudrun, Ragnhild Nordas, and Jan Ketil Rod. 2009. "Regional Inequalities and Civil Conflict in Sub-Saharan Africa." *International Studies Quarterly* 53, no. 2 (June): 301– 324.

Peng, Heng, and Ying Lu. 2012. "Model Selection in Linear Mixed Effect Models." *Journal of Multivariate Analysis* 109 (August): 109–29.

Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa.* Cambridge: Cambridge University Press.

Ravallion, Martin, and Gaurav Datt. 2002. "Why Has Economic Growth Been More Pro-Poor in Some States of India Than Others?" *Journal of Development Economics* 68 (2): 381–400.

Shertzer, Allison. 2016. "Immigrant group size and political mobilization: Evidence from European migration to the United States." *Journal of Public Economics* 139 (July): 1–12.

Stewart, Frances. 2008. *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* Palgrave Macmillan.

Suryanarayan, Pavithra. 2018. "When Do the Poor Vote for the Right Wing and Why: Status Hierarchy and Vote Choice in the Indian States." *Comparative Political Studies*.

Taylor, Michael, and Douglas Rae. 1969. "An Analysis of Crosscutting between Political Cleavages." *Comparative Politics* 1 (4): 534–547.

Thachil, Tariq. 2014. *Elite Parties, Poor Voters: How Social Services Win Votes in India.* New York: Cambridge University Press.

Thachil, Tariq, and Emmanuel Teitelbaum. 2015. "Ethnic Parties and Public Spending: New Theory and Evidence from the Indian States." *Comparative Political Studies* 48, no. 11 (September): 1389–1420.

Theil, Henri. 1967. *Economics and Information Theory.* North-Holland Publishing Company.

Trounstine, Jessica. 2016. "Segregation and Inequality in Public Goods." *American Journal of Political Science* 60 (3): 709–25.

Ziegfeld, Adam. 2018. *Why Regional Parties? Clientelism, Elites, and the Indian Party System.* Cambridge University Press.

Appendix A.

Table A1. Data On State Legislative Assembly (As of 2014)

State	Constituencies	Districts	Voters (Million)	Number of Parties		
				National	State	Other
Andhra Pradesh	294	13	57.9	6	2	43
Assam	60	33	0.7	5	3	16
Bihar	243	38	55.1	6	3	81
Chhattisgarh	90	27	16.9	6	0	28
Goa	40	2	1	5	1	11
Gujarat	182	33	38.1	6	0	34
Haryana	90	22	13.1	7	2	40
Himachal Pradesh	68	12	4.6	6	0	9
Jharkhand	87	24	18	7	3	53
Karnataka	224	30	43.7	6	1	52
Kerala	140	14	23.2	6	5	19
Madhya Pradesh	230	51	46.4	6	0	28
Maharashtra	288	36	76	7	1	69
Manipur	60	9	1.7	5	2	11
Meghalaya	60	11	1.5	4	2	8
Mizoram	40	8	0.7	3	3	2
Nagaland	60	11	1.2	3	1	3
Odisha	147	30	27.2	6	2	24
Punjab	117	22	17.8	6	1	30
Rajasthan	200	33	40.8	6	0	28
Sikkim	32	4	0.3	4	1	3
Tamil Nadu	234	32	47.1	5	3	41
Tripura	60	8	2.4	5	0	9
Uttarkhand	70	13	6.4	6	1	36
Uttar Pradesh	403	75	127.5	6	2	214
West Bengal	294	23	56.3	6	3	46
NCT of Delhi	70	11	11.9	6	0	28

Appendix B. Summary Statistics for Dependent and Key Independent Variables

Figure B1 shows the distribution of the two dependent variables. Figure B1(a) and B1(b) show the distribution of vote-share for parties that champion deprived groups when they contest in their home state and when they contest outside their home state. Figure B1(c) shows the distribution of vote-def for parties that champion deprived groups.
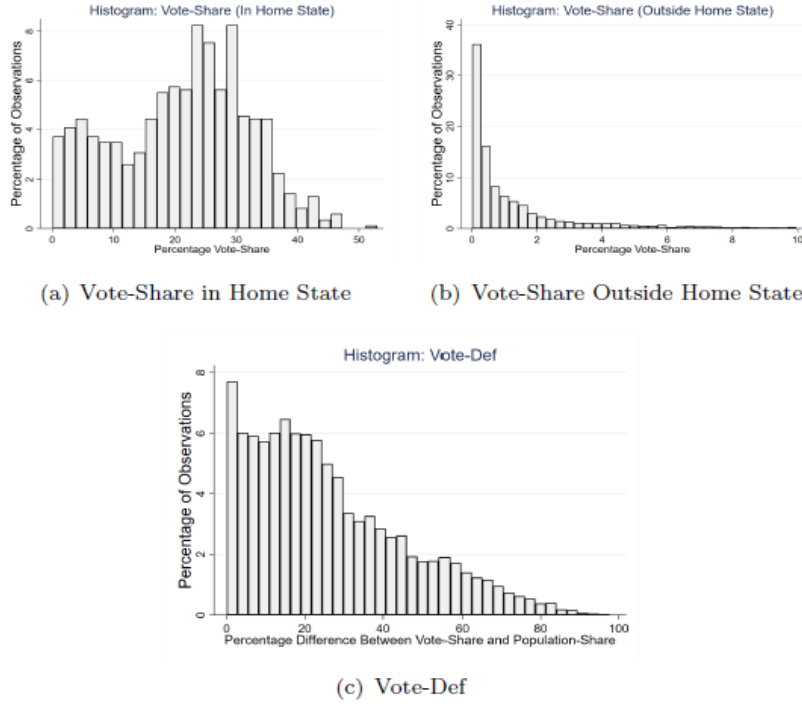

(a) Vote-Share in Home State

(b) Vote-Share Outside Home State

(c) Vote-Def

Figure B1. Histogram: Vote-Share and Vote-Def

Table B1. Summary Statistics

Variable	Mean	Standard Deviation
Vote-Share	0.058	0.097
Vote-Def	0.263	0.198
Within-Group Inequality	0.113	0.081
Between-Group Inequality (Consumption)	0.229	0.290
Between-Group Inequality (Theil)	0.065	0.011
District Educational Attainment[a]	0.91	0.1
Mean District Consumption	0.001	0.027
Access to Electricity (% by district)	0.665	0.306
Population Share (% by district)	0.313	0.209
Margin of Victory	0.105	0.092
Job Fractionalization	0.105	0.092

[a] measures the percentage that have *not* completed higher secondary education.

Appendix C.

Table C1. List of Ethnic Parties that Champion Deprived Ethnic Groups

Name of Ethnic Party	Target Population	Home State
All India Majlis-e-Ittehad-ul Muslimeen	Muslim	Andhra Pradesh
Sikkim Democratic Front	Scheduled Tribe	Sikkim
Bahujan Samaj Party[†]	Scheduled Caste	Uttar Pradesh
Janata Dal (United)[†]	OBC Caste (Hindu)	Bihar
Rashtriya Janata Dal[†]	OBC Caste (Hindu) and Muslims	Bihar
Samajwadi Party[†]	OBC (Hindu) And Muslims	Uttar Pradesh
Indian Union Muslim League	Muslims	Kerala
Jharkhan Mukti Morcha	Scheduled Tribe	Jharkhand
Lok Janshakti Party	Scheduled Caste	Bihar
Jharkhand Vikas Morcha (Prajatantrik)	Scheduled Tribe	Jharkhand

Online Appendix D: Controlling for Heterogeneity Between Constituencies

In Table D1, instead of one general random effect that captures how each observation deviates from the predicted fixed effects, the mixed model generates multiple random effects that capture how observations deviate within a constituency, and how each constituency deviates from the overall group.[15] We use three levels of nesting, wherein each constituency belongs to a district and each district is a part of a state. We also report the Interclass Correlation (ICC). If ICC approaches 0 then there is no variance to explain at the constituency-level, and if the ICC approaches 1 then there is no variance to explain at the district or state level. Table D1 includes 'margin of victory' as a control variable.[16] The purpose is to differentiate between competitive and non-competitive races. The average margin in our sample is 10.5%. We do not find a statistically significant relationship between margin of victory and vote-share.

Table D1: Controlling for Heterogeneity Between Constituencies

Dependent Variable	Vote-Share	Vote-Share	Vote-Share
	AP1	AP2	AP3
Within-Group Inequality (WGI)	-0.026**	-0.023*	
	(0.012)	(0.013)	
Between-Group Inequality (BGI)	-0.010***	-0.008**	
	(0.003)	(0.004)	
Between-group Inequality (Theil)			0.200***
			(0.068)
Interaction Term (WGI×BGI)	0.066**	0.057**	
	(0.026)	(0.028)	
District Educational Attainment[a]	0.016**	0.021***	0.021***
	(0.007)	(0.008)	(0.008)
Mean District Consumption	-0.022	0.016	0.019
	(0.027)	(0.029)	(0.028)
Access to Electricity (% by district)	-0.006	-0.002	-0.003
	(0.005)	(0.005)	(0.005)
Population Share (% by district)	0.038***	0.037***	0.023***
	(0.004)	(0.004)	(0.004)
Regional Party HQ	0.253***	0.253***	0.224***
	(0.002)	(0.002)	(0.053)
Margin of Victory	-0.008	0.004	0.002
	(0.007)	(0.007)	(0.007)
ICC	0.31	0.28	0.11
N	15687	12038	12038

Notes:* $p<0.1$, ** $p<0.05$, *** $p<0.01$. [a] measures the percentage that have *not* completed higher secondary education.

Column AP1 of Table 3 applies to all state legislative constituencies. Column AP 2 excludes those constitutes that are reserved for members of Scheduled Caste or Scheduled Tribe. Such constituencies have higher than average population share of Scheduled Caste or Scheduled Tribe. This may be of advantage to parties that champion the two subaltern groups. We therefore present results for both scenarios, where AP1 includes and AP2 excludes such constituencies. We find that even after controlling for constituency-level effects, within-group inequality is negatively correlated with vote-share.[17] While the coefficient for between-group inequality is negative, the coefficient for the interaction term is positive. The joint-effects can be understood using Figure B1 constructed using the specification in Column AP2. It shows that vote-share increases with an increase in between-group inequality and a decrease in within-group inequality. However, at extreme values of between-group inequality, we find the relationship between within-group

---

[15] Kedar and Shively 2017, 2: "All comparative politics is multilevel." For examples see: Peng and Lu 2012.
[16] Margin of Victory = (Difference between 1st and 2nd highest votes) / total number of votes cast.
[17] We cannot estimate effects for the dependent variable vote-def. This is because constituency-level data on population-share is not available.

inequality and vote-share may turn positive. However, this change occurs when between-group inequality is greater than 0.6. This region covers less 5% of the observations and hence we are unable to verify if the relationship is statistically significant. Second, we believe that the multi-level model does not exactly replicate the Tables 1 and 2 because data on inequality is measured at the district level while vote-share is measured at the constituency level. On an average, each district has 6-7 state legislative constituencies. Since, National Sample Survey reports household consumption only at the district level, we are unable to extend our analysis to the level of constituencies. In Column AP12 we drop within-group inequality and use the Theil Index to measure between-group inequality. We find a positive correlation reaffirming the expectation that high inequality between ethnic groups boosts the vote-share of ethnic parties.

Online Appendix E: Ethnification of Party Systems

        Huber and Suryanarayan (2016) demonstrate that when ethnic inequality is high, the degree to which parties have unique ethnic bases of support increases. Their findings applied to our sample would suggest that parties that champion deprived groups are more likely to succeed in states where inequality between the members and non-members of the group is high. This implies, ethnic inequality not only impacts vote-share but also the variable 'Regional Party HQ.' The underlying rationale being that parties that favor deprived groups are more likely to be found in states where such parties are likely to succeed. In order to separate the direct effects of inequality on vote-share from the indirect effects, captured through 'Regional Party HQ', we use instrumental variables. In Table E1 we instrument 'Regional Party HQ' using three variables, 1) inequality among members of a party's target ethnic group, 2) inequality between members and non-members of party's target ethnic group, 3) and population-share of a party's target ethnic group. All three variables are measured at the state-level and so is 'Regional-Party HQ.' Furthermore, the three variables are estimated using the 55th round of National Sample Survey conducted in 1999. Table F1 shows that ethnic inequality not only manifests through ethnic voting but also through *ethnification* of party systems. The first stage regressions, show that an ethnic party that champions deprived groups is more likely to be found when coethnic population-share is high, inequality within target ethnic groups is low, and inequality between members and non-members of target group/s is high. This also implies that the estimated effects of inequality on vote-share are downward biased in Table 2. Or alternately, ethnic inequality manifests through ethnic voting in the short-run and *ethnification* of party systems in the long run.

Table E1: Ethnification of Party Systems

Dependent Variable	Vote-Share		
	AP4	AP5	AP6
Within-group Inequality (WGI)	-0.016*	-0.041***	
	(0.009)	(0.013)	
Between-group Inequality (BGI)	0.011***	0.000	
	(0.004)	(0.006)	
Between-group Inequality (Theil)			0.401***
			(0.087)
Interaction Term (WGI×BGI)		0.103***	
		(0.039)	
Mean District Consumption (log)	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
District Educational Attainment[a]	0.013**	0.012**	0.013**
	(0.006)	(0.006)	(0.006)
Access to Electricity (% by district)	-0.067***	-0.066***	-0.069***
	(0.005)	(0.005)	(0.005)
Population Share (% by district)	0.067***	0.066***	0.060***
	(0.006)	(0.006)	(0.005)
Margin of Victory	-0.015*	-0.016*	-0.016*
	(0.008)	(0.008)	(0.008)
Regional Party HQ (Instrumented)	0.160***	0.161***	0.159***
	(0.007)	(0.007)	(0.007)
		First Stage	
State Within-group Inequality (1999)		1.419***	
		(0.070)	
State Between-group Inequality (1999)		-0.409***	
		(0.040)	
State Ethnic Population-Share (1999)		0.344***	
		(0.041)	
Wooldridge's Score Test	285.39	283.22	296.10
(p values)	0	0	0
First Stage Regression-Adjusted R^2	0.40	0.40	0.41
N	14228	14228	14228

Notes:* $p<0.1$, ** $p<0.05$, *** $p<0.01$. [a] measures the percentage that have *not* completed higher secondary education. The Wooldridge's Score tests the null hypothesis H_0: 'Regional Party HQ' is exogenous. With a p value of 0, we can reject the null that 'Regional Party HQ' is exogenous.