Ethnic inequality and the ethnification of political parties: Evidence from India

John D. Huber[*]

Pavithra Suryanarayan[†‡]

April 4, 2014

Abstract

Why does ethnicity become a salient element of electoral politics in some political systems but not others? We argue that in majoritarian political systems, inequality between groups facilitates the emergence of strong ethnic bases of support for political parties. Using survey data from the Indian states and OLS and fixed effects models, the paper finds a strong positive relationship between inter-group economic inequality and party system ethnification. A similar pattern exists across majoritarian countries. An examination of the voting behavior of ethnic groups shows that as income differences between groups increase, the groups tend to support different parties. Thus, the analysis reveals a strong class component of identity politics in India, and it underlines the importance of disentangling the effect of group identity from that of economic well-being when studying identity politics.

[*]Professor, Department of Political Science, Columbia University, 7th Floor IAB, West 118th street, (212) 854-3646, jdh39@columbia.edu.

[†]Ph.D student, Department of Political Science, Columbia University, 7th Floor IAB, West 118th street, (212) 854-3646, ps2550@columbia.edu.

1 Introduction

In ethnically divided democratic systems, political parties often seek to build support from specific groups, leading to the "ethnification" of party systems and electoral politics. The degree to which this occurs, however, varies considerably across political systems, with some characterized by political parties with distinctive group orientations and others by parties that lack any clear ethnic base. Why does ethnicity become a salient element of electoral politics in some systems but not others?

Most existing research on ethnic politics pays little attention to the differences in economic endowments of groups, focusing instead on the intrinsic importance of ethnic categories.[1] Ethnicity provides a marker that elites can easily use to target voters and form electoral coalitions (e.g., Bates 1983, Horowitz 1985, Chandra 2004 and Posner 2004 and 2005). Electoral politics is therefore akin to an "ethnic head count" where the contest is to form a "minimum winning coalition." Parties seek to represent winning coalitions of groups by strategically employing appeals to particular ethnic identities, and voters in turn strategically invoke the identities that are most advantageous to them in efforts to become part of a winning coalition. The stakes of such ethnic competition are high, as the parties that win power can distribute government resources to the groups that support them. From this perspective, the specific attributes of group members are secondary. Ethnic identity creates a largely exogenous marker that simplifies targeting, coalition building, and exclusion from government resources. It does not matter if group members have a particular set of policy preferences or ideologies, but rather what is important is that individuals can be identified for inclusion in government patronage (or exclusion from it) based on their group membership.

In contrast, this paper considers the relationship between group identity and economic well-being, which we will call "class." If ethnicity and class are related, the salience

---

[1]An exception to this Horowitz's (1985) categorization of ethnic societies into ranked and unranked systems with the former being systems where ethnicity and class location overlap and latter where groups are economically heterogeneous.

of ethnicity in electoral politics should be stronger for two reasons. First, when economic well-being and group identity are related, political cleavages will be reinforcing, reducing cross-pressures on voters and thus making it easier for parties to use group identities as they endeavor to assemble the broad-based electoral coalitions that majoritarian systems require. Second, when groups vary in their levels of economic well-being, they should have different policy preferences, making it possible to appeal to groups with general policy appeals rather than through ethnic appeals. A party can appeal to poor ethnic groups, for example, by championing policies that benefit the poor.

This paper develops this argument and tests it by comparing party system ethnification across the states in India's majoritarian electoral system. Focusing on one case allows us to test our argument while holding constant many of the factors that can make cross-national comparisons difficult. Across the Indian states, the electoral rules and the menu and meaning of ethnic identities are largely the same, and while the states differ in their economic development, individuals across the states find themselves in essentially the same political and economic framework. Using survey data from two National Election studies (NES) in India (1999 and 2004) and a metric of party system ethnification defined in Huber (2012), we measure the salience of ethnicity in state party systems for the three central definitions of identity in India: religion, caste and subcaste (or *jati*). We find that party system ethnification is highest using *jati* and lowest using religion. We therefore focus primarily on the *jati* definition of groups to test our argument empirically.

We also use the surveys to measure between-group inequality ("BGI"), a component of the Gini decomposition that measures differences in the average income of groups. When BGI is low, and group average incomes are therefore similar, there is low correlation between group identity and economic well-being. When it is high, the opposite is true. We find a very strong and robust positive relationship between BGI and the ethnification of party systems – the more inequality there is between ethnic groups, the higher the levels of ethnification, suggesting perhaps that ethnic politics in India has a stronger economic

or "class" orientation than has been previously recognized. Using fixed-effects models that take advantage of the two time periods in our data, we also provide evidence of a causal effect of ethnic inequality: those states with the largest increases in ethnic inequality between 1999 and 2004 are also the states where there is the largest increase in party system ethnification in the 2004 election.

The relationship between BGI and party system ethnification does not, however, imply a strong class dimension of ethnic politics in India: it may be that groups with similar levels of economic well-being vote together, but it is also possible that electoral coalitions form between richer and poorer groups. We therefore probe the relationship between group income differences and group voting patterns in India using a group-level analysis where we measure the voting differences between each pair of groups in a state. We regress this measure of voting differences on income differences between the groups and find that members of one *jati* group tend to vote for different parties than members of another *jati* group when income differences between the two groups is large. This suggests that voting differences across *jati* groups are associated with economic factors.

Some might worry that the Indian case is unique and thus a difficult one from which to generalize. We therefore situate India comparatively using cross-national data from majoritarian political systems. The analysis shows a strong association between between-group inequality and party system ethnification across countries, and it also indicates that the level of party system ethnification in India is about what one would expect given India's level of between-group inequality.

The paper is organized as follows. The next section elaborates our argument about group-based economic inequality and the ethnification of party systems. Section **??** then presents the measures of party system ethnification and between group inequality that will be used in the empirical analysis, and section **??** describes relevant features of ethnic politics in India. Four different empirical analyses follow: section **??** explores which definition of group identity in India is associated with the highest level of party system ethnification;

4

section **??** presents our state-level analysis in India; section **??** presents our group-level analysis in India, and section **??** presents our cross-national analysis. We conclude by discussing the implications of our results.

2 Party ethnification and inequality between groups in majoritarian systems

Our central concern is with the "ethnification of political parties," a term we use to describe situations where parties have a strong ethnic basis of support. If it is relatively easy to know a voter's ethnic identity by knowing the party they supported, the ethnic basis of support is strong compared with situations where one learns little about a voter's ethnic identity from his or her vote choice. As Chandra (2012) describes, focusing on the ethnic basis of support for parties is but one way to study the role of ethnicity in electoral politics, but it is an important one because it provides a clear indicator of the extent to which group-based identities are relevant to the outcomes of party strategies and electoral behavior.

Why should inequality between groups be associated with party ethnificiation in majoritarian systems like India's? A key feature of majoritarian elections is that they force parties to build electoral coalitions by appealing to diverse sets of voters. Parties or candidates can only win with broad-based support, a challenge given that broad-based coalitions invariably have voters with diverse interests. The level of inequality between groups should be important to the role of "ethnic politics" in efforts to build such electoral coalitions for two related reasons.

First, economic inequality between groups should lead to systematic differences in groups' public policy preferences. Baldwin and Huber (2010) emphasize that economic differences across groups "can lead to different group needs with respect to public goods, feelings of alienation or discrimination by some groups, different attitudes toward redis-

5

tribution across groups, and different 'class' identities by different groups" (pp. 644-45).[2] Consequently, parties can make general policy appeals – such as commitments to provide public goods like clean water, education, health care – that transcend groups, but that are particularly attractive to members of specific groups, such as relatively poor ones. This will lead to a strong relationship between group identity and political behavior, even when there exist few overt appeals for votes based on group identity. In addition, if group identity and income are related, parties can more easily make commitments to specific groups – such as minority groups that are crucial to the success of the party – without alienating members of the broader coalition. If a party, for example, commits to building new schools or water treatment plants, or to hiring more public sector workers in a region dominated by a specific group, this is less likely to cause tension within the broader electoral coalition if the party is generally committed to policies related to schools, clean water and public sector employment. In the US, for example, the left-wing Democrats can make appeals for votes from the blacks by emphasizing redistributive issues that are of importance to non-black members of the electoral coalition they seek to build, and they can stress affirmative action policies without alienating many white supporters, who have historically supported the broader redistributive and social justice agenda implicit in affirmative action policies.

The second reason why inequality between groups should be associated with party system ethnification concerns cross-cutting cleavages, which create cross-pressures on voters. As Dunning and Harrison (2010) demonstrate in their study of Mali, cross-cutting cleavages can reduce incentives for ethnic voting because they force voters to make trade-offs on different dimensions of importance. Income and ethnicity are obvious potential sources of cross-cutting cleavages. If a party is trying to build a broad-based electoral coalition of different groups, then when poor individuals tend to be in different ethnic groups than rich individuals, a party can make appeals on the economic and non-economic dimension (ethnicity) without dividing a group against itself. If group identity is not cor-

---

[2]Baldwin and Huber (2010) link between group inequality to lower public goods provision, and Alesina et al 2013) link it to lower economic development.

related with income, a party that appeals to the poor and to particular groups based on their non-economic policies will force the rich in these groups to choose between their economic interests and the group-specific policies. Thus, parties attempting to build majorities will inevitably create cross-pressures on voters from the same group, diminishing the importance of ethnicity in vote choice.

Some might expect that these two related arguments might not do a good job of explaining ethnic political behavior because ethnic politics are often held to revolve around clientelistic or patronage-based appeals, where parties compete for votes by committing to group-specific payoffs. But even in situations where patronage or clientelism are important, majoritarian systems make it very difficult to sustain purely distributive electoral strategies because of the inherent instability of any given electoral coalition of groups. That is, in an ethnically diverse society, any majority electoral coalition of groups that a party attempts to build based on patronage payoffs can be defeated by another coalition that makes at least one group (in the first coalition) better off. This inherent instability makes it more difficult for parties to credibly commit to electoral coalitions of specific groups, and invites the use of some sort of policy to help cement patronage-oriented electoral coalitions. In general, the role of income-related policy preferences – in contrast to the patronage demands of group members – has perhaps been under-appreciated in studies of ethnic politics, and the inequality-based perspective described here could be relevant even in contexts where patronage and clientelism are important.

In sum, in majoritarian systems where parties attempt to build electoral coalitions that are necessarily diverse, the role of ethnicity in this dynamic should be strongest when ethnicity is correlated with income. It is therefore important to examine empirically whether patterns of party system ethnification are related to inequality between groups.

3 Measuring Party System Ethnification and Ethnic Inequality

To quantify the degree to which the group basis of political support for parties varies across India's states, we draw on Huber's (2012) measure of *Party Voting Polarization* (PVP), which is designed to make possible comparisons across political systems that vary in the definition of groups or in the number of parties. We briefly describe the measure here, and further details are found in Huber (2012).

The measure is defined as follows. First, for each pair of political parties i and j in a state, we calculate the distance between their electoral bases of support, $\tilde{r}_{ij}$. Formally,

$$\tilde{r}_{ij} = \sqrt{\frac{1}{2} \sum_{g=1}^{G} (P_g^i - P_g^j)^2}, \tag{1}$$

where P_g^i and P_g^j are the proportion of supporters of parties i and j who come from group g, and there are G groups. The maximum $\tilde{r}_{ij}$ is 1, which occurs when all of party i's support comes from one group, and all of party j's support comes from another group. The minimum distance is zero, which occurs when the proportion of support that comes from each group for party i is the same as the proportion of support that comes from each group for party j. In general, as the distribution of groups supporting any two parties becomes more similar, $\tilde{r}_{ij}$ will decline.

Second, we use the vote shares of parties to aggregate the $\tilde{r}_{ij}$'s, yielding the system-level measure of PVP:

$$PVP = 4 \sum_{i=1}^{N} \sum_{j=1}^{N} p_i p_j^2 \tilde{r}_{ij}, \tag{2}$$

where p_i is the proportion of the vote received by party i and there are N parties.

The $p_i p_j^2$ term in the definition of PVP ensures that holding the ethnic distance between parties constant (and assuming it is non-zero), the measure increases as the party

system moves toward two parties of equal size. This "polarization perspective" emerged in large part from Horowitz's (1985) argument that many ethnic groups is less problematic for stable governance than two equal-sized groups. The polarization perspective has gained the upper hand in studies of ethnic politics and civil conflict: the intuitive logic provided by Horowitz has considerable appeal and empirical tests have been supportive.[3] The polarization perspective applied to party systems is appropriate in the majoritarian electoral systems found across the Indian states. The first-past-the-post electoral systems will disproportionately reward the top two parties, so the impact of ethnically-based support for parties becomes less clear in situations where this vote support exists across a range of parties. In addition, in majoritarian systems, we should also expect that when parties have an ethnic basis in politics, electoral stakes will be largest when two competing parties are close to the threshold of victory.

The PVP measure used here therefore has the following two central properties. First, holding the number and size of parties constant, the measure increases with increasing distance in the ethnic basis of support for parties (i.e., increases with $\tilde{r}_{ij}$). Second, holding constant the distances in ethnic bases of support, PVP increases as the party system moves toward two parties, each receiving 50 percent of the vote, and each receiving their support from a single (different) ethnic group.

To explore whether the measure of party system ethnification varies with the level of inequality between groups, we use India's NES (see details below) to calculate between-group inequality (BGI). BGI is a component of the Gini index that is based on the average income differences between groups, weighted by group size. When BGI is large, the average incomes of group will vary and there will be a relatively strong correlation of group identity and income. When BGI is small, there is very little correlation of income and group because the groups' average incomes are roughly the same. The formula for BGI is

---

[3]See e.g., Reynol Querol 2002 and Montalvo and Reyno Querol 2005. See Esteban, Mayoral and Ray (2012) for evidence regarding the circumstances under which fractionalization and polarization are relevant to explaining ethnic violence.

given by

$$BGI = \frac{1}{2\bar{y}}(\sum_{m=1}^{k}\sum_{n=1}^{k} p_m p_n \mid \bar{y}_m - \bar{y}_n \mid),\tag{3}$$

where m and n index groups, p_m is the proportion of the population in group m, $\bar{y}_m$ is the average income of group m, $\bar{y}$ is the average income in the society, and there are k groups in society. BGI is one of three components of the Gini coefficient. The other two are within-group inequality (WGI), which measures inequality that is strictly within groups by aggregating the Gini coefficient for each group, and overlap (O), which has been interpreted as a measure of income stratification.[4]

To calculate BGI, we need a measure of individual "income." Given that a large proportion of individuals do not have meaningful cash incomes in a developing country like India, it is not feasible to accurately measure inequality by using standard income variables. We therefore follow the strategy used in a number of previous studies in emerging economies that employ various asset indicators to gauge economic well-being (see, for example, Filmer and Pritchett (2001) and McKenzie (2005)). Specifically, the Indian election surveys of 1999 and 2004 ask individuals if they own particular items. Both surveys contain the following 7 variables that are associated with economic well-being in India:

- Car/Jeep/Van
- Tractor
- Television/Color Television/Cable Television
- Scooter/Motorcycle/Moped
- Telephone/Mobile telephone
- Bicycle
- Pumping set

---

[4]See Yitzhaki and Lerman (1991) for a formal definition of the three components of the Gini decomposition and a useful discussion of their substantive meaning.

Respondents were given a score of 1 for each asset listed above. To measure a respondent's "income" we conduct a factor analysis on these assets in each state. The resulting factor scores describe the degree to which the various assets distinguish the well-being of citizens, and thus are used to weight the assets which are aggregated (using the factor weights) to determine an individual's "income." Respondents' "income," based on their asset ownership and factor scores, are then rescaled to their percentile rank (thus ranging from 0 and 100), which gives all individuals a non-zero "income," allowing us to perform the Gini decomposition on the "income" values. Although this approach is standard for measuring inequality in countries where cash incomes are insignificant for large proportions of the population, it is important to recognize that these measures underestimate total inequality because they do not capture the differences in income that exist among the relatively well-off (who tend to have all of the relevant asset items). While it is not feasible to meaningfully combine asset information with income information, an alternative way to test our arguments is to measure inequality using educational attainment rather than assets. This is an additional strategy we follow below to explore the robustness of our results.

4   Ethnic politics across states in India

Individuals in India typically have multiple "ethnic" identities, three of which have become salient political cleavages and the focus of much scholarship: religion (e.g., Brass 2003, Rudolph and Rudolph 1993, Varshney 2003, Wilkinson 2004, Jaffrelot 2005b and Tachil 2010), caste (e.g., Srinivas 1962, Rudolph 1965, Krishna 2003, Jaffrelot 2005a and Chandra 2004) and subcaste, or *jati* (e.g., Chhibber 1999, Chandra 2004 and Dunning and Nilekani 2013).[5] These three ethnic categories are related to each other, and are in fact

---

[5]Some have also focused on language (e.g., Fearon (2003), as part of a large cross-national study of language differences across groups), although in studies of India, language differences are seldom viewed as politically salient any more.

nested. The broadest categorization is religion and the Hindu-Muslim divide is the most salient religious cleavage in the country.[6] Hindu-Muslim politics has taken on a heightened salience with the rising fortunes of the right-wing Hindu nationalist party, the Bharatiya Janata Party (BJP). Brass (2003), for example, argues that the vote for the BJP is directly linked to the proportion of muslims in an area. Rudolph and Rudolph (1993) argue that economic competition between hindus and muslims (as determined by their occupational patterns in the area) leads to greater salience of religion in politics. Wilkinson (2004) notes that hindu-muslim riots tend to occur in closely contested elections. We categorize respondents in our data into Hindu, Muslim, Christian, Sikh, Buddhist, Jain and Parsi.

"Caste" is nested within religion. The role of caste in Indian politics is considered to be so central that Chandra (2004, 212) writes: "In India, people do not cast their votes, they vote their caste." While scholars use the term differently, with some using "caste" to refer to what we call "*jati*" and others to refer to the *"varna"* system, we use caste to refer to seven broad and politically salient umbrella categories that represent social-status ranking (Heath 2005, Chhibber, Jensenius and Suryanarayan 2013 and Dunning and Nilekani 2013).[7] Thus individuals in our surveys were sorted into the following caste categories: upper castes, peasant castes, upper backward castes, lower backward castes, schedules caste, scheduled tribes and muslims.

Finally, we focus on sub-castes, or *jati*, which are nested within the broad caste categories listed above. We use this term to refer to a "hereditary, endogamous, usually localized group" (Srinivas 1962, 3) which reflects historical occupational categories and which to this day, guide religious and marriage customs. To categorize respondents into their sub-castes we use a survey question asking: "What is your Caste/Jati-biradari/Tribe name?" To give an example of how these categories are nested within each other, a re-

---

[6]The Hindu-Muslim cleavage emerged early on in electoral politics in colonial India and eventually led to a violent partition and the creation of Pakistan and present-day Bangladesh. Sporadic violence in the form of religious riots continue to occur in the Indian states, particularly during election years.

[7]In the the *varna* system, caste groups are categorized into five broad categories of Brahmins (priests), Kshatriya (warrior), Vaishya (merchant), Shudras (backward castes) and the caste groups that lie outside the varna system (the former untouchables).

spondent from the state of Andhra Pradesh who has a "*kapu*" sub-caste is categorized as "Peasant" caste and "Hindu" religion.

We examine variation in party system ethnification using these three categories of ethnicity across the Indian states. We do this because state-level factors are central in national-level voting. The Indian states have constitutionally mandated autonomy over key policy areas such as education, healthcare, law and order and agriculture amongst others, and their role in economic policymaking has increased after major economic changes in 1991(Chhibber and Nooruddin 2004). Even when the central government legislates on key developmental projects such as the Mahatma Gandhi National Rural Employment Guarantee Act that was enacted in 2006, there is a high level of discretion given to states in implementation. The introduction of village-level government institutions through the 73rd constitutional amendment in 1992 further served to strengthen state-level political parties and intensified efforts by national parties to embark on state-centric organization building.

The Congress Party, which had dominated politics for many years, evolved into a collection of regionally diverse party organizations under one banner, with little control by central elites over local party messages. Another major party, the Bharatiya Janata Party (BJP), is also an organization that is vertically divided. Although the party was ideologically driven, cohesive and boasted a centralized machinery in the past, by the 1990s, the ideological wing of the party was sidelined and state-level opportunists and moneyed candidates were given a central role in the party with a view to winning elections. In addition, recent changes in regional party systems have seen the rise of parties that compete in only one or two states and that garner significant national parliamentary seats in those states such as the DMK and AIADMK in Tamil Nadu, the TDP in Andhra Pradesh, the Akali Dal in Punjab, the BSP in Uttar Pradesh and the TMC in West Bengal, amongst others.

Consequently, parties and voters tend to focus on state-level political and economic

factors during national elections. We can see this by examining a question in the 2004 NES survey which asks respondents to state the importance of state government versus national government performance when casting their national vote.[8] Across India, the weight respondents put on state government performance was equal to the weight put on national government performance.

5 Party ethnification by group identity in the Indian states

We use the Party Voting Polarization (PVP) measure to explore patterns of party system ethnification across the Indian states. The core data for our analysis come from the National Election Studies of 1999 and 2004, each conducted in the aftermath of national parliamentary elections.[9] The NES surveys are distinctive for their large samples of voters across the Indian states and are conducted face-to-face in the local language using a structured questionnaire.[10] NES 2004, for instance, was conducted in 22 Indian languages and had 27,189 respondents.[11]

Two states, Jammu and Kashmir and Manipur, are excluded from the analysis. These states have complex separatist movements dominating their politics. The decision to exclude the two states, however, does not affect the substantive results.

The measures of PVP in the states are based on the three most salient definitions of ethnic group – religion, caste and *jati*. Which type of group is associated with the highest level of party system ethnification? In the 41 state-specific surveys that we have across the

---

[8]The question in the 2004 survey– "While voting some people give more importance to the work done by the state government while others give more importance to the work done by the central government. While voting in this (national) election, what mattered to you the most?"

[9]Elections to the 543 seats in the lower house of the national parliament in India are held every five years. The elections are held under the guidance of the Election Commission of India and according to single member, simple plurality rules. In the two elections under investigation in this paper, the 1999 and 2004 elections, the boundaries of the 543 constituencies remain unchanged in accordance to a freeze in constituency boundaries that had been in effect since 1976.

[10]For more information on the NES and other surveys by CSDS, please go to http://lokniti.org.

[11]In comparison, the World Values Survey (WVS) of India conducted in 2006 had 2000 respondents and was administered in ten major Indian languages.

two elections, *jati* produces the highest PVP score 24 times (or 59 percent of the time), caste produces the highest score 14 times (34 percent) and religion produces the highest score only 3 times (7 percent).

It is important to bear in mind that the PVP measures should be affected by the underlying distribution of groups in a state. If each group, for example, voted for a different party, then differences in PVP across states would not be a function of voting behavior but rather of differences in the number and size of groups. Thus, in assessing the levels of PVP it is important to control for the underlying level of ethnic polarization. Ethnic polarization ("EP") is a measure of ethnic diversity that considers only the number and size of groups, and it takes its maximal value when there are two equal size groups. The measure grows smaller as the number of groups proliferate or as one group becomes predominant in size.[12]

When we regress PVP on EP, an indicator for election year, and indicator variables for group type (with caste as the omitted category), the coefficient on the religion indicator is -.03 (p-value .05) and the coefficient on the *jati* indicator is .05 (p-value .006). This implies that controlling for EP and the year, using religion to tap group identity yields PVP scores that are on average .03 lower than those resulting from using caste, and that are .08 lower than those obtained using *jati*. And using *jati* results in PVP scores that are on average .05 greater than those resulting from using caste. These are fairly large differences given that the mean of PVP across all groups is .16 with a standard deviation of .09.

In sum, the evidence suggests that *jati* is the definition of group identity that is associated with the highest level of party system ethnification across the Indian states, followed by caste. Although there is a small handful of states where voting based on religion is very important, religion is the definition of group that is associated with by far

---

[12]The formal definition of EP is from Reynol Querol (2002), who draws on Esteban and Ray (1994):

$$EP = 1 - \sum_{i=1}^{G} \left(\frac{1/2 - s_i}{1/2} \right)^2 s_i, \tag{4}$$

where s_i is the size of group i and there are G groups.

the lowest levels on average of party system ethnification. In what follows, we will focus primarily on the *jati* definition of groups, although we will also present some analysis using the larger umbrella caste definition.

6   Party system ethnification across the Indian states

What is the relationship between party system ethnification and between-group inequality? Figure **??** shows a scatter plot of PVP and BGI for the three definitions of group. As noted above, in examining the relationship between PVP and any right-hand side variable, it is important to control for the underlying level of Ethnic Polarization. The values of PVP plotted in the figure are therefore the residuals from a simple regression of PVP on EP. The figure shows a strong positive relationship between PVP and BGI for the *jati* and caste definitions of group. The slopes are very similar, with a slightly stronger relationship when *jati* is used (the slope is .78 for *jati* and .71 for caste). By contrast, using religion, there is absolutely no relationship between BGI and PVP. We also see in the figure that BGI using religion is quite small compared with BGI using the other definitions of groups: the mean of BGI is .16 using *jati*, .13 using caste, and only .03 using religion. The figure therefore underscores that we should not expect a relationship between PVP and BGI for any definition of groups. Instead, these data suggest that when a group definition is not correlated with economic differences across groups, then voting polarization is typically not strong for this definition of groups, although obviously there can be isolated situations where such a definition nonetheless leads to strong ethnification.

OLS models. Does the relationship depicted in Figure **??** survive when controlling in an OLS regression framework for other variables that could affect PVP? Table **??** presents a number of models where PVP is the dependent variable, BGI is the right-hand side variable of central interest, and a number of other controls are included as well. All of the variables are standardized to have a mean of zero and standard deviation of 1 in order to facilitate
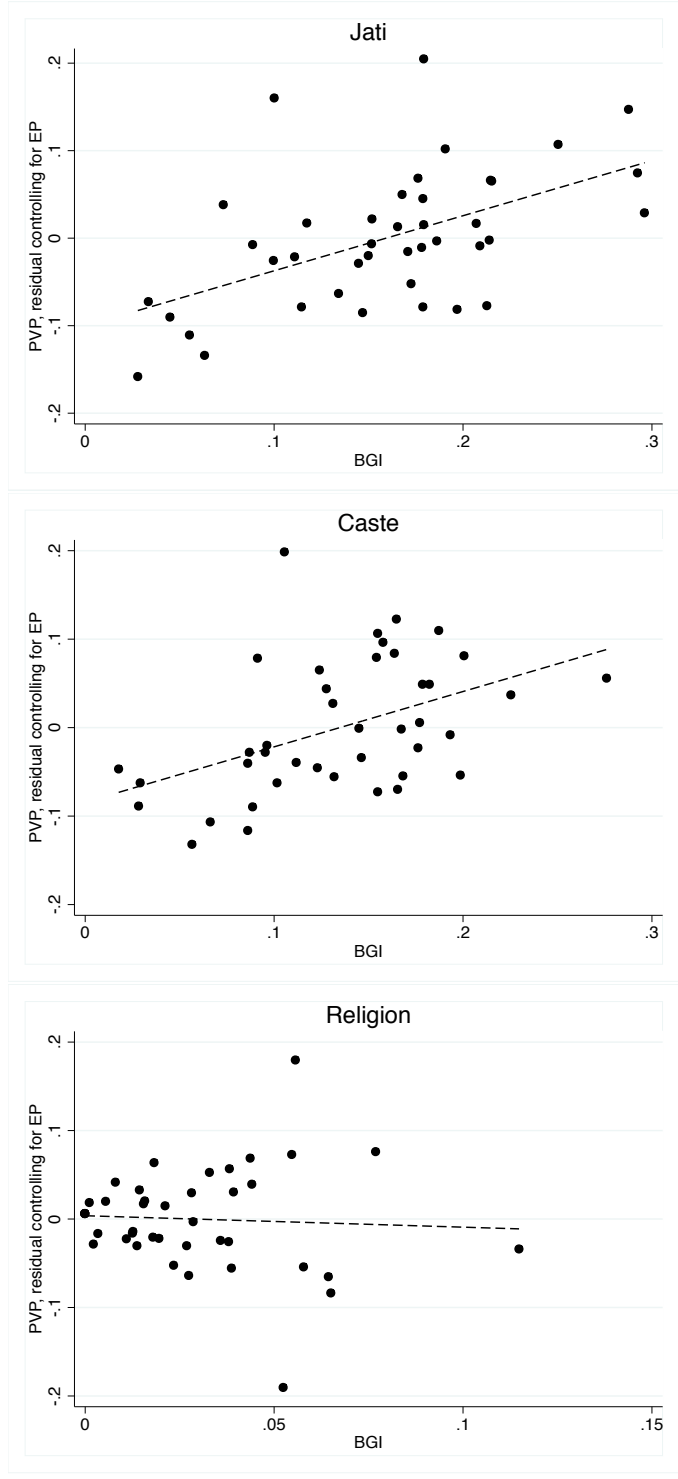
Figure 1: BGI and Party Ethnification using three definitions of group

comparisons of the coefficients.[13]

Each of the models includes three core control variables. As noted above, it is important to control for EP (so that we know variation in PVP is due to variation in voting patterns rather than variation in the number and size of groups). In addition, since BGI is one of three components of the Gini, to ensure any results we attribute to BGI are not in fact capturing other elements of inequality, we include in all models a variable, "Gini-BGI," which is the Gini minus its BGI component. Thus, we can interpret the coefficient for BGI as the effect of a change in BGI when all others aspects of inequality are held constant. Finally, we include a year fixed effect, an indicator variable for the 2004 election, to control for any systematic differences in ethnic voting that may have existed across the two elections.

Model 1 presents results where the right-hand side variables include only BGI, EP, Gini-BGI and the 2004 indicator variable. We see that the coefficient for BGI is the largest in magnitude and the most precisely estimated – a one standard deviation increase in BGI is associated with an increase in PVP of 0.630 standard deviations. This effect is even substantially larger than that of EP itself, the coefficient for which is also very precisely estimated. The coefficient on the indicator variable for the 2004 election is negative but estimated with considerable error, suggesting there is no significant difference in party ethnification across these two elections (when we control for EP and BGI). And there is negligible effect of inequality that is unrelated to BGI.

Are these results robust to the inclusion of other control variables? Particularly poor countries are often associated with patronage-based politics, thereby making the game of ethnic distributive politics especially salient. Model 2 includes a measure of economic development, a 5-year average of state GDP per capita (logged).[14] We find that BGI still has a large and precisely estimated coefficient, but the GDP variables and residual inequality

---

[13]A coefficient of 1 therefore indicates that an increase of one standard deviation in a right-hand side variable is associated with a 1 standard deviation increase in PVP.

[14]The state GDP values are available through the "Handbook of Statistics on Indian Economy" on the the Reserve Bank of India website www.rbi.org

Table 1: OLS regressions of PVP on BGI using asset 'income' data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BGI	0.630**	0.756**	0.653**	0.581**	0.617**	0.640**	0.465
	(0.253)	(0.311)	(0.269)	(0.264)	(0.254)	(0.256)	(0.278)
EP	0.433**	0.396**	0.483***	0.423**	0.445***	0.430**	0.087
	(0.160)	(0.169)	(0.169)	(0.161)	(0.161)	(0.161)	(0.244)
Gini-BGI	0.097	0.186	0.195	0.068	0.196		
	(0.200)	(0.237)	(0.227)	(0.205)	(0.229)		
WGI						0.126	-0.148
						(0.320)	(0.405)
Overlap						0.178	0.125
						(0.232)	(0.258)
GDP/cap(ln)		0.115					
		(0.162)					
Regional Parties			0.279				
			(0.318)				
Hindu Belt			0.216				
			(0.321)				
Northeast			-0.209				
			(0.462)				
Tax expenditure				0.010			
				(0.119)			
ELF					0.284		
					(0.320)		
2004	-0.373	-0.371	-0.328	-0.454	-0.386	-0.382	-0.556**
	(0.289)	(0.291)	(0.302)	(0.302)	(0.290)	(0.292)	(0.268) 218)
Constant	0.341	0.297	0.261	0.426	0.206	0.269	0.231
	(0.285)	(0.294)	(0.375)	(0.300)	(0.324)	(0.310)	(0.263)
Adj. R-squared	0.322	0.312	0.290	0.317	0.318	0.310	0.382
N	41	41	41	40	41	41	41
Group type	jati	jati	jati	jati	jati	jati	Caste

Note: Continuous variables are standardized to have a mean of 0 and a standard deviation of 1. Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.

do not.

Model 3 includes three regional indicator variables. The first is an indicator for states in the North East of the country.[15] These states are unique because of their distinct historical status under the British in the colonial period, the predominantly tribal composition of some of the states and because of secessionist conflicts in some others. While we see no compelling *ex ante* reason to expect that BGI by sub-caste group should work differently in these states, we should nonetheless control for this region to see if patterns of party ethnification are systematically different in these states, and if controlling for these differences affects the coefficient estimate for BGI. The second is an indicator for the Hindi speaking belt.[16] It is possible that the hindi-belt was more susceptible to the mobilization efforts by the hindu nationalist party, the BJP, and because of differing historical land-holding patterns and the greater numerical numbers of upper castes, experienced caste mobilization much later and more intensely than the states in the South (Jaffrelot 2005a). Third, we include an indicator variable, "Regional Party Strength," for states which have a strong regional party. We expect systems with regional parties to be associated with less ethnic parties because regional parties are believed to have mobilized voters across religious, caste and class groups by creating a sub-national unity amongst the states' residents (Singh 2010).[17] The results in Model 3 show no systematic differences in PVP by region. The coefficients for Hindi Belt and Regional Party Strength are positive, but they are estimated with considerable error. And the coefficient for Northeast states is negative (albeit very small) and also estimated with considerable error. And there are no statistically significant differences in the coefficient for any pair of these variables. The result for BGI is not affected by the inclusion of these variables.[18]

---

[15]The northeast includes the following: Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Tripura and Sikkim.

[16]Hindi-belt comprises of Bihar, Haryana, Himachal Pradesh, Madhya Pradesh, Rajasthan and Uttar Pradesh

[17]We include the states of Andhra Pradesh, Assam, Kerala, Punjab, Tamil Nadu and West Bengal in this category.

[18]We also included these regional variables one by one into the models and the results were similar to the ones presented here.

It is possible that state governments with relatively high bureaucratic professionalism and capacity to provide services may make voters less susceptible to ethnic appeals by parties. To measure bureaucratic capacity, we measured the states' expenditure on tax institutions as a proportion of overall expenditure. This variable is a proxy for the quality of tax and transfer institutions in the state – the higher the level of such expenditures, the more bureaucratic professionalism we should expect.[19] In Model 4 we find no effect of Tax Expenditure on PVP, and its inclusion has little effect on the main result for BGI.

If parties are trying to build the smallest winning coalitions of ethnic groups, they may be most successful when there are many small groups with which to bargain. This suggests that ethnification may be strongest when group fractionalization is highest. To test for this possibility, Model 5 includes a measure of ELF, or ethno-linguistic fractionalization. The coefficient for ELF is positive but not at all precisely estimated. The coefficients for BGI and EP continue to be large, positive and precisely estimated.

Model 6 examines the effect of within-group economic inequality on party system ethnification. We might expect that as groups become more economically heterogeneous, their members should vote less cohesively, driving down party system ethnification. We can test this using a measure of within-group inequality, or WGI, which is a weighted sum of the Gini coefficient for each group. As WGI increases, the economic heterogeneity of groups increases. The coefficient on WGI should therefore be negative if such heterogeneity is associated with decreased party system ethnification. When including BGI and WGI in the same model, it is also necessary to include Overlap (the residual component of the Gini related to stratification). Model 6 shows that the coefficients for all three variables are positive, but only BGI has a coefficient that is precisely estimated. Thus, our data suggest that any relationship between party system ethnification and inequality in India is working through inequality between groups and not any other component of group inequality.

---

[19]The data used to create these measures are available through the "Database on Indian Economy" on the Reserve Bank of India website. The state financial documents collect information on state expenditures and revenues available here: www.rbi.org.

Finally, we re-estimate model 6, but using caste to define group identity. In model 7, BGI has the expected positive sign, and the coefficient, though not statistically significant by standard metrics, is measured somewhat precisely (p=.104). The size of the BGI coefficient is also smaller than we find using *jati* (recall the coefficients are comparable across specifications because all variables are standardized). Thus, the association between BGI and PVP is strongest when we use *jati* to define groups; the relationship is also present but weaker using caste.

Measuring inequality by using information about individual assets is attractive because as noted above, these assets help distinguish the well-being of the many individuals who have little cash income. A drawback is that it can underestimate inequality by ignoring differences at higher levels of economic well-being. An alternative measure of economic well-being that captures some of this inequality among the more well-off is educational attainment. The NES surveys categorize respondents as non-literate, primary school, middle school, junior high school, high school, college (no degree), college (graduated), post-graduate degree and professional degree. We can use this 9 point scale as a measure of "income" to calculate educational inequality.[20] This measure provides an avenue for exploring the robustness of the findings in Table **??**. The results, presented in Table **??** in the Appendix, are, if anything, even stronger for the BGI variable when educational attainment is used to measure inequality.

Fixed effects. The regressions in Tables **??** and **??** demonstrate that the ethnification of party systems is strongest in the states where inequality between groups is largest. This association is robust to a range of model specifications, and it is robust when we measure inequality using either individual assets or educational attainment. But we might worry that BGI is correlated with some other right-hand side variable that we have not measured, decreasing our confidence in the results. And there is the possibility of reverse causation: if individuals vote by group, and if the winning parties favor their groups, then ethnic voting

---

[20]Respondents' educational scores are rescaled to their percentile rank (thus ranging from 0 to 100), which we use to measure educational inequality along with the three components of the Gini.

might cause inequality between groups rather than the contrary.

We can address such issues empirically by taking advantage of the fact that we have 15 states for which we have measures in both 1999 and 2004. This panel structure makes it possible to estimate fixed effects models (which also corresponds to estimating a difference-in-difference model given there are only two time periods). A fixed effects model obviously addresses concerns about possible biases from omitted regressors on the right-hand side. And it also addresses concern about reverse causation because any changes in the measure of BGI that occurs between the 1999 survey and the 2004 survey must precede the 2004 election itself, making it impossible that changes in 2004 voting caused the changes in inequality that we measure.

Figure **??** shows bivariate scatter plots where the y-axis plots the change in PVP from 1999 to 2004 and the x-axis plots the change in BGI from 1999 to 2004. For the plot on the left, BGI is measured using asset indicators and for the plot on the right, BGI is measured using education. In both plots, there is a strong positive relationship: those states with the highest increase in inequality between groups are the states with the highest increase in party system ethnification. The relationship is stronger using educational attainment, but this is due largely to the leverage of the outlier state, Assam ("AS"), which depresses the slope of the relationship in the plot using asset indicators.

The results from fixed effects regressions are presented in Table **??**. The first two columns use asset indicators to measure inequality. Column 1 includes the standard controls (Gini-BGI and EP), as well as the wealth of the state. The coefficient for BGI is positive and relatively precisely estimated (p=.08), even with the outlying state (AS) included. Model 2 drops the GDP variable, which has a coefficient estimated with considerable error in model 1. The coefficient for BGI is now estimated much more precisely (p=.03). Models 3 and 4 estimate the same models using educational inequality. In both models, the coefficient for BGI is positive and estimated very precisely (p< .005). It is important to bear in mind that these fixed effects regressions are based on only 15 states and one

Figure 2: Changes in PVP vs. changes in BGI in 15 states, 1999 to 2004
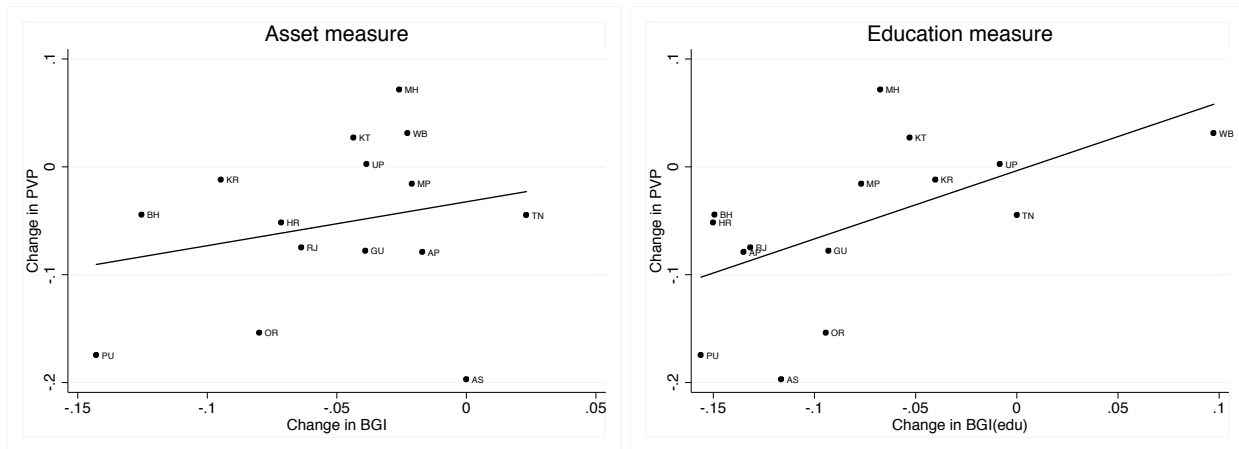
change in time period. Nonetheless, the impressive bivariate correlations in Figure **??** and the results from the empirical models in Table **??** provide evidence of a causal effect of inequality between *jati* groups on group voting behavior.

7   Group level analysis

We have presented evidence that the ethnification of parties is strongest when ethnicity and average income are strongly correlated, but this does not imply that party ethnification carries with it a strong class basis. High ethnification, for example, could be associated with members of poor *jatis* supporting different parties than members of rich *jatis*, resulting in a relatively strong class basis for party politics when ethnic voting is high. It could also be the case, however, that high BGI leads to high ethnification, but with no clear patterns of party support by *jati* that are based on income, with some parties drawing on support from certain rich and poor *jatis* and other parties drawing on support from other rich and poor *jatis*. This would lead to high ethnification, but to a weak class basis for party politics.

This section presents a group-level analysis of voting by *jati* to explore how the average income of groups is related to voting patterns. We draw on Huber (2012) to

24

Table 2: Fixed effects regressions of PVP on BGI

	(1)	(2)	(3)	(4)
BGI (assets)	0.966*	1.109**		
	(0.509)	(0.455)		
Gini-BGI (asset)	0.581	0.448		
	(0.752)	(0.711)		
BGI (edu.)			0.780***	0.699***
			(0.220)	(0.174)
Gini-BGI (edu.)			-0.124	-0.013
			(0.378)	(0.326)
EP	0.516*	0.512*	0.436*	0.466**
	(0.260)	(0.254)	(0.212)	(0.201)
GDP/cap(ln)	-0.093		0.076	
	(0.135)		(0.120)	
Constant	-0.483	-0.316	0.052	-0.107
	(0.373)	(0.278)	(0.297)	(0.157)
Adj. R-squared	-0.259	-0.204	0.238	0.276
N	30	30	30	30

Note: Models 1-2 measure inequality using asset indicators.
Models 3-4 measure inequality using educational attainment.
Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.

measure the "voting distance" between members of any two groups. The basic idea is to compare the distribution of support for the different parties by one *jati* group with the distribution of support by another *jati* group in the state. If the members of the first *jati* support parties in the same proportions as the members of the second *jati* (e.g., each *jati* gives 20 percent of its support to party 1 and 80 percent of its support to party 2), then the electoral distance is 0. The distance measure increases as the proportions that each *jati* gives to each party diverge, and it reaches its maximum when each *jati* has its own party (e.g. *jati* 1 gives 100 percent of its support to party 1 and *jati* 2 gives 100 percent of its support to party 2). Formally, let $\bar{r}_{ij}$ be the measure of electoral distance between groups i and j in a state. Then

$$\bar{r}_{ij} = \sqrt{\frac{1}{2}\sum_{k=1}^{N}(V_i^k - V_j^k)^2}.$$

where V_i^k and V_j^k are the proportion of members of group i and j respectively who support party k and there are N total parties in a state. The measure of $\bar{r}_{ij}$ is of course closely

related to $\tilde{r}_{ij}$, but it is based on group voting patterns rather than to the composition of support for parties.

Our goal is to understand if $\bar{r}_{ij}$ increases with the income differences between groups. We therefore regress $\bar{r}_{ij}$ on "Income Difference," which is the absolute difference in the mean income of groups i and j. Table **??** presents our results. In each model the dependent variable is the measure of $\bar{r}_{ij}$, and there are 2,233 pairs of groups across the states in the two elections.[21] The models also include state and year fixed effects (not reported). All continuous variables are standardized to have a mean of 0 and a standard deviation of 1.

Model 1 includes only Income Difference (and the state and year fixed effects). The coefficient is positive and very precisely estimated. We find that as income differences between groups increase, members of poor *jati* groups tend to vote for different parties than members of rich *jati* groups, suggesting that income differences between *jati* groups does lead to income-based electoral politics.

Does umbrella caste mitigate the effect of income distance? It may be the case that members of *jati*s belonging to the same umbrella caste group will vote in similar ways, regardless of their class interests. Model 2 examines this idea by including "Same Caste," an indicator variable taking the value 1 if the two *jati*s in a pair are from the same caste, as well as the interaction of Same Caste and Income Difference.

The results provide very uneven support for the idea that caste membership mitigates the effect of income difference on voting. The coefficient for Same Caste is negative and reasonably large in magnitude, but it is not at all precisely estimated. And the coefficient on the interaction term has the correct sign, but it too is very small in size and estimated with very large error. We find, however, that the coefficient for Income Difference remains large, positive and precisely estimated. The main conclusion to draw from

---

[21]We kept all pairs for the *jati*s belonging to Upper Castes, Peasant Castes, Upper backward Caste, Lower Backward Castes, Dalits and Tribals while removing *jati*s belonging to Christian, Muslims, Buddhists, Sikhs and Jain religions where the respondents did not also identify as a member of a larger caste group.

the model is that the effect of income differences on voting differences is essentially the same regardless of the caste affiliations of the *jati*s being compared.

Models 1 and 2 estimate the effect of the absolute difference in the average income of two groups on the voting behavior of the two groups. An alternative approach, which provides more information about which part of the income scale is producing the income differences, is to replace Income Difference with two variables, "Poor Income" (the average income of the poor group, which should have a negative coefficient if income differences lead to voting differences) and "Rich Income" (the average income of the rich group, which should have a positive coefficient). These variables allow us to understand if the results for income differences are driven disproportionately by changes at the high or low end of the income distribution of groups. Model 3 presents the results when only Rich Group Income and Poor Group Income are included (along with the fixed effects). We find that the coefficients for both variables are in the expected direction, that they are very precisely estimated, and that they are reasonably similar in absolute magnitude. Thus, there is no evidence from this model that the results we have found for income differences are driven disproportionately by the voting patterns of rich or poor *jati* groups.

Model 4 adds Same Caste and the interactions of Same Caste with the income variables. If shared caste diminishes the effect of income differences between *jati* on voting differences, then the interaction with Poor Income should be positive and the interaction with the Rich Income should be negative. When the caste variables are included in the model, we find that the coefficients for Poor Income and Rich Income remain precisely estimated with the expected sign. The coefficients for the interactions, however, are not precisely estimated, and the coefficient for the Rich Income*Same Caste has the wrong sign (suggesting increases in rich group income is associated with greater voting differences when *jati*s are of the same caste rather than different castes). When we calculate the standard errors for the interaction terms, however, we find that there is no statistically significant difference in the estimated marginal effect of Poor Income within caste and

Table 3: OLS regressions of Voting Distance on Income Distance between groups

	(1)	(2)	(3)	(4)	(5)
Income Difference	0.159***	0.153***			
	(0.019)	(0.021)			
Same Caste		-0.090		-0.109*	-0.113**
		(0.080)		(0.057)	(0.057)
Income Difference * Same Caste		-0.003			
		(0.030)			
Poor group income			-0.165***	-0.174***	-0.177***
			(0.022)	(0.024)	(0.035)
Rich group income			0.143***	0.129***	0.202***
			(0.022)	(0.023)	(0.040)
Rich inc.*Same Caste				0.045	0.043
				(0.069)	(0.069)
Poor inc.*Same Caste				0.072	0.077
				(0.064)	(0.064)
Small group size					-0.033
					(0.022)
Large group size					0.019
					(0.025)
Rich group Gini					0.090**
					(0.040)
Poor group Gini					-0.003
					(0.035)
State and year fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	-0.867	-0.861	-0.846	-0.826	-1.033
	(0.884)	(0.884)	(0.884)	(0.883)	(0.923)
Adj. R-squared	0.221	0.221	0.222	0.223	0.224
N	2233	2233	2233	2233	2233

Note: Continuous variables are standardized to have a mean of 0 and a standard deviation of 1. Standard errors are in parentheses: * p<.10, ** p<.05, *** p<.01

Poor Income across castes. The same is true for rich income.

Model 5 adds a number of group level controls to model 4. "Large Group" is the size of the largest group and "Small Group" is the size of the smallest group. If group size affects group heterogeneity or the appeal of groups in coalition building, these variables may affect voting patterns. Similarly, "Rich Gini" is the group Gini for the richest group and "Poor Gini" is the group Gini for the poorest group. Since a large Gini reflects greater income heterogeneity within a group, we might worry that the level of within-group inequality could mediate the effect of group income differences on voting outcomes.

There is little support for the hypothesis that group-level heterogeneity systematically affects voting patterns. Only one of the four group-level control variables – Rich group Gini – is measured precisely, and its positive coefficient suggests that all else equal, a rich heterogeneous group will have more distinctive voting patterns than a rich homogenous group. This effect, however, is not substantively large, and the inclusion of these group-level variables in the model, does not affect the results for income distance found in models 3 and 4. The coefficients for Poor Income and Rich Income remain large in absolute magnitude and are very precisely estimated (significant at the .01 level). And they are virtually identical in absolute size. Same Caste is more precisely estimated in model 5, negative and significant at the .05 level. But the coefficient of -.113 implies that a change in Poor Income from one standard deviation above the mean to one standard deviation below the mean would be associated with an increase in voting distance that is over three times the effect of a change from a within-caste *jati* comparison to a between-caste *jati* comparison.

As with the state-level regressions, for the group-level analysis we re-estimated the models in Table **??** using educational attainment rather than "asset income" distance between groups. The results, presented in Table **??** of the appendix, are remarkably similar to those in Table **??**. In each model, we find a robust and large effect of educational differences on voting differences.

In sum, we find robust evidence that as two *jati* groups in a state diverge in their average income or education, they tend to support different parties. This implies a strong class element in Indian ethnic electoral politics. In addition, we find that umbrella caste groups do little to mitigate the effect of sub-caste income differences on sub-caste voting patterns. This is a particularly important finding for Indian politics as it is often taken as conventional wisdom that these politicized higher caste categories are most critical to explaining variations in ethnic voting patterns in the country.

8   Cross-national analysis

Some might worry that the historic tie of *jati* groups to occupation makes *jati*-based politics in India rather unique. In particular, we might expect that the link between *jati*-based group identity and economic well-being to be particularly strong, with members of any given *jati* group unusually homogeneous with respect to economic well-being, and with inequality between *jati* groups unusually large. If this is true, India might be a difficult case from which to draw more general inferences about inter-group inequality and the salience of ethnicity in elections.

We doubt this concern is valid given the socio-political dynamic in India's post-independence period. In particular, affirmative action policies, economic reforms, growing urbanization and transformations in the agricultural economy have dramatically altered the relationship between *jati* and occupations to the extent that it is generally difficult to tell the *jati* of a person by their occupation (Karanath 1996). In addition, the average economic heterogeneity of *jatis* is quite high. When we calculate the Gini of each *jati* and take their average in each state, then the mean of these averages is .403. The mean of state-level Ginis is .440, only .037 higher. Thus, although there is variation across *jatis*, the average inequality within groups is very close to the average inequality within the states themselves.

We can also explore the issue of Indian exceptionalism directly by examining the relationship between BGI and PVP cross-nationally. Huber (2012) provides measures of PVP from 24 surveys in 13 majoritarian systems.[22] And Huber and Mayoral (2013), drawing on a wide range of individual level surveys, provide measures of the Gini decomposition for each of these countries. The cross-national data relies on the definition of groups in Fearon (2003). For India, we will use the data from the 1999 and 2004 surveys, but given that $jati$ vary considerably across states, it is impossible to use $jati$ at the national level.

---

[22]The surveys includes the Comparative Study of Electoral Systems, the World Values Survey and Afro-barometers. The majoritarian countries are Australia, Bangladesh, Botswana, Canada, France, India, Kenya, Madagascar, Malawi, Mali, Nigeria, United States and Zambia.
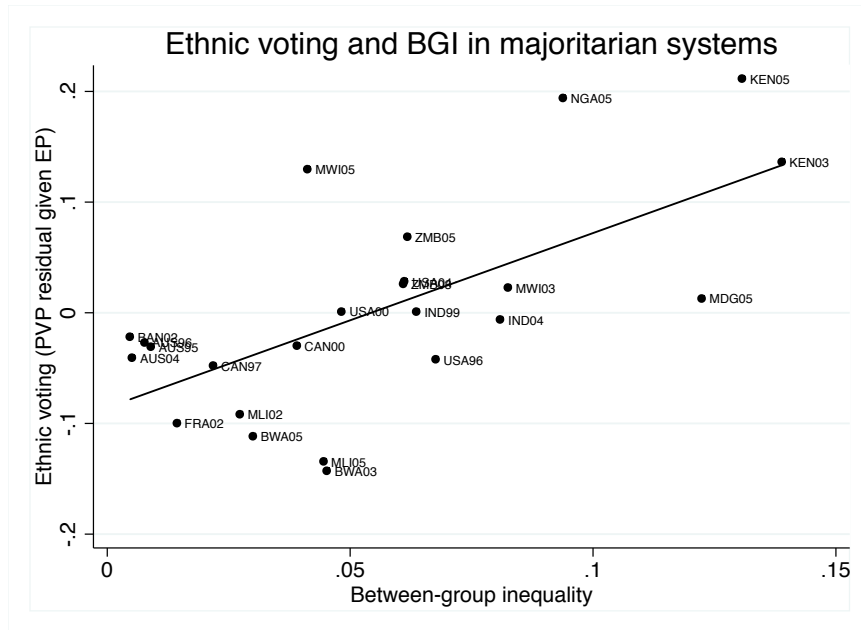
Figure 3: PVP and BGI in 13 majoritarian countries

But we show above that caste yields the second highest level of party ethnification in India and that it also has a positive relationship with BGI across the Indian states. We therefore calculate PVP and BGI using the Indian NES data for caste. The cross-national data analysis not only makes it possible to examine the empirical association between PVP and BGI outside India, but also to situate India in a broader comparative context.

Figure **??** shows the same plot using the cross-national data that we show for India in Figure **??**. That is, we regress PVP on EP and plot the residuals against BGI. Two things are noteworthy in the Figure. First, the relationship, if anything, is stronger in the cross-national data than in the cross-state data. The greater the level of inequality between groups, the greater the level of PVP. Second, India is not an outlier. Although there is slightly less ethnification of the parties than one would expect given the level of BGI, the India observations are not far from the bi-variate regression line.

Table **??** provides results from OLS models. Model 1 includes BGI, residual inequality (Gini-BGI), EP and other controls for economic development (the log of GDP per capita), federalism and the level of democracy (Polity2). All of the variables are stan-

dardized to have a mean of 0 and a standard deviation of 1 to facilitate comparisons of coefficients. We will focus our discussion on the inequality variables. In model 1, BGI has a positive coefficient that is very precisely estimated, and residual inequality (Gini-BGI) has a negative coefficient that is not at all precisely estimated. Model 2 introduces each of the components of the Gini, and only the coefficient for BGI is positive and precisely estimated. Model 3 drops the GDP and Federalism variables (whose coefficients are estimated with considerable error), and the coefficient for BGI is large, positive and precisely estimated. Thus, as in the analysis across the Indian states, in majoritarian countries, there is a strong, positive and robust association between inequality between groups and the ethnification of party systems.

For comparison, model 4 estimates model 3, but with PR rather than majoritarian systems. The results are quite different: BGI has no relationship with PVP but the coefficient for within-group inequality is negative and precisely estimated. Thus, the relationship between various components of inequality and ethnic voting are different under different electoral laws. While it is beyond the scope of this paper to carefully probe these differences, we suspect they may be due to the fact that when party entry is relatively easy (as under PR), multiple parties can compete for the same group, and they should be most successful in doing this when the groups are heterogeneous (as will be true when WGI is high). Different parties representing the same groups can then work together in post-election coalition politics. By contrast, in majoritarian systems where parties try to fashion governing coalitions by creating broad electoral coalitions, members of a group must vote together if they are to be influential in governance as a group. Their ability to do this will be greatest when the groups share an attribute – like being relatively rich or poor – that can be easily incorporated into the broader electoral coalition that parties are trying to build.

Table 4: Cross-national OLS regressions of PVP on inequality variables

	(1)	(2)	(3)	(4)
BGI	0.808**	0.693*	0.785**	-0.033
	(0.348)	(0.357)	(0.365)	(0.089)
Gini-BGI	-0.162			
	(0.117)			
WGI		-0.227	-0.172	-0.366**
		(0.160)	(0.149)	(0.176)
Ov		0.012	-0.332	0.232
		(0.361)	(0.279)	(0.146)
EP	0.605***	0.472*	0.731***	0.071
	(0.138)	(0.267)	(0.193)	(0.085)
GDP/capita(ln)	0.219	0.414		
	(0.321)	(0.470)		
Polity	-0.666***	-0.725***	-0.361*	-0.043
	(0.219)	(0.244)	(0.200)	(0.144)
Federalism	0.477	0.404		
	(0.492)	(0.527)		
Constant	0.451	0.425	0.696***	-0.221*
	(0.358)	(0.364)	(0.202)	(0.116)
Adj. R-squared	0.566	0.549	0.526	0.288
N	24	24	24	45
Elec. Law	SMD	SMD	SMD	PR

Note: Continuous variables are standardized to have a mean of 0 and a standard deviation of 1. Robust standard errors are in parentheses: * p<.10, ** p<.05, *** p<.01

9　Conclusion

Our empirical evidence finds that in India's majoritarian system, the degree to which there is a strong ethnic basis of support for political parties is strongly related to the level of inequality between groups. This is true when we measure inequality using asset indicators and using education levels, and it is true when we estimate fixed-effects models that can rule out omitted variables and (in the Indian case we study) reverse causation. We find very similar patterns in cross-national data from majoritarian systems. And the group-level analysis in India shows that the voting patterns of members from two different groups diverge when the mean incomes of these groups diverge.

We would emphasize two implications of this research. First, existing studies of ethnic politics typically ignore economic factors and focus instead on how patronage is used in places where identity politics are salient. Our findings emphasize that much can be gained by linking the study of economic factors and the study of ethnic politics. Parties and voters cannot turn the salience of particular group identities on or off like water at the tap. Voters have to have a reason for viewing particular identities as salient to their electoral behavior, and economic characteristics of groups can be central to determining whether identity politics are useful electorally. The strategies that parties use in appealing to ethnic groups in efforts to create diverse electoral coalitions are likely constrained by how economic well-being and group identity are related, as are the policies that winning parties can use to retain the confidence of electoral coalitions that put them into power.

Second, the findings suggest that when there is a strong element of ethnic politics in elections, we must be aware that this ethnic politics may be masking a strong element of class politics. If, for example, poor groups are not only voting cohesively, but are also voting with other poor groups, as we find here, and if the same is true for rich groups, then when electoral coalitions are related to group identity, they are also related to class. This is a quite different perspective on ethnic politics than is traditionally emphasized in the literature. At the same time, we would emphasize that our findings do not suggest

34

that ethnic politics are little more than class politics in disguise. We find that the economic heterogeneity of groups does not affect group voting patterns in India or other majoritarian systems, implying that when inequality between groups triggers the salience of ethnic identity, members of the same group tend to vote together independent of the degree to which they have common economic interests. Thus, inequality between groups might well contribute to group solidarity that transcends income differences within groups.

A central task for future research is to explore more carefully how this process works. A provactive finding here, for example, is that economic inequality works differently in the majoritarian systems we analyze than in PR systems. We suspect that this may be due to the different incentives that exist when governing coalitions must be created before elections occur (as they must under SMD) as opposed to when they can be formed after elections, during coalition building (as can occur under PR). Future research should therefore explore why economic inequality within groups might divide a group against itself more easily under PR than under majoritarian electoral rule.

Table A.1: OLS regressions of PVP on BGI using educational inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BGI (edu.)	0.590***	0.613***	0.528***	0.559***	0.544***	0.573***	0.419*
	(0.166)	(0.175)	(0.177)	(0.180)	(0.191)	(0.182)	(0.240)
EP	0.507***	0.489***	0.549***	0.488***	0.503***	0.513***	0.194
	(0.165)	(0.172)	(0.172)	(0.172)	(0.167)	(0.169)	(0.254)
Gini-BGI (edu.)	-0.190	-0.193	-0.252	-0.202	-0.196		-0.083
	(0.172)	(0.174)	(0.200)	(0.184)	(0.175)		(0.184)
2004	-0.190	-0.190	-0.144	-0.252	-0.205	-0.200	-0.407
	(0.302)	(0.305)	(0.310)	(0.329)	(0.306)	(0.309)	(0.312)
GDP/cap(ln)		0.060					
		(0.134)					
Regional Parties			0.134				
			(0.310)				
Hindu Belt			0.243				
			(0.320)				
Northeast			-0.350				
			(0.411)				
Tax expenditure				0.008			
				(0.120)			
ELF					0.137		
					(0.274)		
WGI (edu.)						-0.338	
						(0.436)	
Overlap						-0.148	
						(0.161)	
Constant	0.157	0.136	0.094	0.215	0.087	0.100	0.093
	(0.314)	(0.321)	(0.367)	(0.344)	(0.347)	(0.393)	(0.298))
Adj. R-squared	0.364	0.350	0.343	0.342	0.351	0.347	0.381
N	41	41	41	40	41	41	41
Group type	*jati*	*jati*	*jati*	*jati*	*jati*	*jati*	Caste

Note: Continuous variables are standardized to have a mean of 0 and a standard deviation of 1. Standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.

Table A.2: OLS regressions of Voting Distance on
Educational Distance between groups

	(1)	(2)	(3)	(4)	(5)
Education Difference	0.162***	0.150***			
	(0.020)	(0.021)			
Same Caste		-0.162**		-0.097*	-0.092
		(0.081)		(0.057)	(0.057)
Education Difference * Same Caste		0.026			
		(0.030)			
Poor group education			-0.205***	-0.204***	-0.211***
			(0.027)	(0.030)	(0.036)
Rich group education			0.172***	0.154***	0.102***
			(0.025)	(0.026)	(0.033)
Rich edu.*Same Caste				0.102	0.102
				(0.080)	(0.080)
Poor edu.*Same Caste				-0.001	0.005
				(0.076)	(0.076)
Small group size					-0.027
					(0.022)
Large group size					0.022
					(0.025)
Rich group Gini					-0.086**
					(0.039)
Poor group Gini					0.012
					(0.039)
State and year fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	-0.706	-0.696	-0.631	-0.607	-1.006
	(0.884)	(0.883)	(0.884)	(0.883)	(0.924)
Adj. R-squared	0.221	0.222	0.222	0.224	0.225
N	2233	2233	2233	2233	2233

Note: Continuous variables are standardized to have a mean of 0 and a standard deviation of 1.
Standard errors are in parentheses: * p<.10, ** p<.05, *** p<.01