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

John D. Huber*

Pavithra Suryanarayan†‡

May 4, 2013

Abstract

Why does group identity, such as ethnicity, become a salient element of electoral politics in some political systems but not others? We argue that inequality between groups plays an important role in answering this question: systems with the highest levels of group-based inequality should be the ones where identity is most salient to electoral competition. We test the argument using data from across the Indian states, finding that state-level party system ethnification is strongly correlated with economic inequality between groups in the states. We also find that when 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.

‡We thank Kanchan Chandra, Francesca Refsum Jensenius, Kimuli Kasara, Allison Patch, Joel Simmons, Neelanjan Sircar, Milan Vaishnav, seminar participants at Columbia University and the University of Maryland and participants at "the Ethnic Politics and Electoral Democracy" conference (Zurich 2012) for their comments on earlier drafts of this paper. We thank the Lokniti team at the Center for the Study of Developing Societies, New Delhi for making the Indian National Election Studies data available to us.

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? This paper addresses this question, arguing that the degree to which ethnicity and economic class are related is an important part of the answer.

Most existing research on ethnic politics emphasizes the intrinsic importance of ethnic categories. Ethnicity is held to provide a marker that elites can easily use to target voters and form electoral coalitions (e.g., Bates 1983, Horowitz 1985, Chandra 2004 and Posner 2005). Electoral politics is therefore akin to an "ethnic head count" where the contest is to form a "minimum winning coalition" of ethnic groups. Parties seek to represent winning coalitions of groups by strategically employing appeals to particular group identities, and where voters strategically invoke the particular 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 distribute government resources to the groups that support them. From this perspective, the specific attributes of group members are essentially irrelevant. 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 to this existing research, we argue that the relationship between group identity and economic well-being, which we will call "class", should have a central influence on the degree to which ethnicity is politicized in a given democratic system. If

ethnicity provides a cue about class, then ethnicity also provides a clear indication to parties about the potential policy and redistributive demands of group members. This can have at least two effects on the salience of ethnicity in electoral politics. First, when the class and ethnic cleavages overlap rather than cross-cut, it is it more attractive for group members to vote their ethnic identities because they do not face a trade-off, for example, between party appeals based on redistribution and party appeals based on group targeting. Second, the correlation of group and class helps parties to organize bargaining across groups, for example with parties representing poor groups allying with parties representing the middle class. This helps parties to credibly commit to specific policy objectives in a way that is impossible in purely distributive bargaining across ethnic groups. If voters can take group-specific commitments by parties more seriously in situations where group membership and economic class are correlated, parties have a greater incentive to make the group-based appeals in the first place. Ethnic identity, then, can play an important role in party electoral appeals, but the success of such appeals in coordinating the voting behavior of group members should be influenced by non-immutable group characteristics, such as cross-group differences in the material well-being of ethnic group members.

This paper develops this argument and tests it by comparing party system ethnification across the states in India. As we lay out in more detail below, India is a useful country in which to test our argument for a number of reasons, including the fact that recent changes in state party systems provide some reassurance that inequality between groups is likely driving ethnic voting patterns rather than the reverse. Using survey data from two national election studies 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 use the *jati* definition of groups to test our argument empirically.

We also use the survey to measure between-group inequality ("BGI") , a compo-

nent of the gini decomposition that measures differences in the average income of groups. When it 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. This correlation, however, does not imply a clear class dimension of ethnic politics: while BGI is correlated with party system ethnification, it may still be that *jati* voting patterns are not organized along economic lines. We therefore probe the relationship between group income differences and group voting patterns using a group-level analysis where we measure the voting differences between each pair of groups in a state, a strategy that also allows us to include state fixed effects. 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 income differences between *jati* groups is associated with class-based electoral politics.

The paper is organized as follows. The next section elaborates our argument about group-based economic inequality and the ethnification of party systems. We then present the measure of party system ethnification and between group inequality that will be used for the cross-state analysis, and we discuss the motivation for testing the argument on state politics in India. In the subsequent section we present the results of the three main empirical analyses. First we show that using *jati* groups – rather than religion or caste – yields the highest levels of party system ethnification in India. Second, we present empirical models of variation in ethnification across the states. Third, we present our group-level analysis. We conclude by discussing the implications of our results.

2 Party ethnification and inequality between groups

Most theoretical and empirical research on the ways in which parties develop an ethnic basis of support emphasize the intrinsic importance of ethnic categories in political competition rather than the link between ethnic group and income, or economic class. Horowitz (1985), for example, argues explicitly against the idea that ethnicity is a mask for class, or that ethnic affiliations simply provide a means for elites to satisfy class aspirations. Instead he maintains that ethnicity creates intrinsically compelling affiliations that are more important to individuals than class, which is why elites mobilize groups using ethnic appeals instead of class appeals. He goes so far as to argue that ethnic voting occurs because of its "expressive value," and given that ethnic identity is the over-riding factor shaping voting behavior, party elites must employ ethnic strategies to win elections. A key reason for the permeative role of ethnicity is that unlike class, ethnicity is rooted in descent. People are able to transcend their class in a lifetime or over a single generation in a way they are unable to do so with ethnicity. Consequently, while Horowitz acknowledges that ethnic societies can differ in the extent to which class and ethnicity coincide (i.e. ranked and unranked systems), he downplays the role of class in the politicization of ethnicity.

Horowtiz's seminal work has been enormously influential, but scholars have also recognized that the argument is incomplete. If ethnicity is an over-riding determinant of voting behavior, why do we see the ethnification of some party systems and not others? In addition, why are some ethnic categories salient in some party systems and not others?

Recent research has made strong inroads to answering these questions. Chandra (2004) for example, emphasizes not the "expressive value" of ethnicity (which she rejects), but rather its strategic value. She argues that in many contexts, voters believe that access to government benefits will depend on having co-ethnics in power, and party elites therefore commit to favoring particular ethnic groups by explicitly courting their support. Thus, electoral competition revolves around party elites mobilizing members of particular groups in efforts to win an "ethnic head count." Posner (2005) also presents a strategic argument,

arguing that an individual voter will choose to politicize the identity category that reaps the highest economic returns. In so doing, voters seek to adopt the identities that will allow them to be included in the smallest possible winning coalition (so that the spoils of victory are shared as narrowly as possible).

As with Horowitz, these authors assume that the attributes of group members themselves – such as whether they tend to be poor or otherwise different or excluded – are irrelevant to the emergence of ethnic-oriented electoral politics. Instead, group identity alone is sufficient to generate group-oriented voting behavior. Our central departure from this existing research is to argue for the value of introducing information about group attributes, focusing specifically on whether the level of economic inequality between groups plays a central role in explaining the importance of ethnic identity in electoral politics.

There are several reasons that inequality between ethnic groups should contribute to a heightened role of ethnicity. First, economic differences should lead to differences across groups in policy preferences. As Baldwin and Huber (2010, 644-45) emphasize, 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 redistribution across groups, and different 'class' identities by different groups." Their study finds that economic inequality between groups – not just the degree of group fractionalization itself – affects bargaining over public goods, resulting in lower levels of public good provision. Alesina et al (2012) also argue that inter-group inequality affects governance, and they find that economic development is lowest in countries where economic differences between groups are highest.

Second, inequality between groups should contribute to the ethnification of party systems because when group identity and income are related, it helps political parties to organize policy bargaining across groups. One limitation of existing arguments about the strategic efforts of groups to maximize their share of government resources is that such arguments do not explain how the generic instability of bargaining over distributive

benefits is resolved. These arguments assume that post-election bargaining is a purely distributive game across groups. But such bargaining is generically unpredictable – any coalition of groups in favor of some policy (or distribution of government revenues to groups) can be defeated by a coalition of other groups favoring a different policy. If group A and B are in a minimal winning coalition, for example, then an omitted group C has every incentive to cut a deal with A or B so it is not excluded. But then the excluded group has incentives to re-open the negotiations (or emphasize a different identity) so that they can be part of the winning coalition. Where this ends up is difficult to predict in any general fashion, making it difficult for parties to credibly commit to specific policies for specific groups.

When individual income and group identity are correlated, this creates a clear organizing principle for policy bargaining, one that takes into consideration the distinctive policy preferences of individuals from different groups. Parties representing poor groups, for example, may be able to commit to redistributive policy coalitions with parties representing other poor groups (or middle class groups). Or parties representing richer groups may credibly commit to coalitions with parties representing middle-class groups. Thus, when ethnic categories and class are correlated, it should be easier for parties to engage in group-based politics because they can more easily commit to pursuing specific policy outcomes that benefit group members.

Third, economic differences across groups should contribute to party system ethnification by reducing cross-pressures on voters. As Dunning and Harrison (2010) demonstrate in their study of Mali, cross-cutting cleavages can reduce incentives for ethnic voting. Income is an obvious potential source of cross-cutting cleavages: if poor individuals tend to be in different ethnic groups than rich ones, the group and class cleavages will be reinforcing, and individuals will not face a choice between supporting a party that represents their group and one that represents their income. If group identity is not correlated with income, then the group and class cleavages are cross-cutting, diminishing incentives for

ethnic voting.

In sum, we are likely to observe more party system ethnification in systems where inequality between ethnic groups is high. Such inequality implies that ethnic identity is correlated with income, which should create policy preferences that are correlated with group identity, with members of poorer groups, for example, preferring different policies than members of richer groups. This correlation of group and income also provides a natural way to order post-election bargaining processes across groups, and it reduces cross-pressures on voters, making it easier for parties to credibly commit to group-based policies. It is therefore important to examine empirically whether patterns of party system ethnification can be explained by the correlation of material well-being and group identity.

3 Measuring Party System Ethnification and Ethnic Inequality

Our objective is to quantify the degree to which the group basis of political support for parties varies across the Indian states. To this end, 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. The measure has two elements: (a) a measure of differences between the ethnic basis of support for all pairs of parties in a state, and (b) a method for aggregating these differences that takes into account the size of the parties.[1] In this section, we describe the measure, as well as our measure of inter-group inequality.

Measuring differences in the ethnic bases of support for all pairs of parties in a state. The first step to calculating PVP in a state is to calculate the "distance" between the ethnic profiles of supporters for each pair of parties in the state. This distance is determined by

[1]Heath (2005) and Chhibber, Jensenius and Suryanarayan (2012) create measures related to party ethnification in India, but the measures do not allow valid comparisons across states using different definitions of groups, or when the number of parties varies.

comparing the proportion of support that each party receives from each ethnic group. The distance is zero for a pair of parties when the group-basis of support for each party is the same. Suppose Party A, for example, receives 30 percent of its support from the Green group and 70 percent of its support from the Purple group. Suppose that the same was true for Party B. Then there would be zero distance in the ethnic basis of support for these two parties. By contrast, suppose that all of Party A's support came from the Green group and all of Party B's support came from the Purple group. This would yield the maximal distance between the ethnic basis of support for each party.

Formally, let $\tilde{r}_{ij}$ be the distance in the electoral bases of support for parties i and j, which is defined as

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

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.

For two hypothetical parties, A and B, Table 1 provides two examples of the calculation of $\tilde{r}_{AB}$.[2] In both examples, there are three groups, and we calculate the distance between the two parties using the composition of group support for each party. In the top example, Party A receives 40 percent of its support from Group 1, 35 percent of its support from Group 2, and 25 percent of its support from Group 3. Party B has a fairly similar profile, but a smaller proportion of its support comes from Group 1 and a larger proportion of its support comes from Group 3. Thus, $\tilde{r}_{AB} = .15$. In the second example at the bottom of the table, the support base for each party is skewed toward a particular group, with Party 1 receiving 80 percent of its support from Group 1 and no support from Group 3, and Party 2 receiving 80 percent of its support from Group 3 and none from Group 1. Thus, $\tilde{r}_{AB} = .8$, which is considerably larger than in the top example.

[2]This example is taken from the supplemental materials in Huber (2012).

The maximum $\tilde{r}_{ij}$ is 1, which occurs when all of party A's support comes from one group, and all of party B'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 A is the same as the proportion of support that comes from each group for party B. In general, as the distribution of groups supporting any two parties becomes more similar, $\tilde{r}_{AB}$ will decline.

Table 1: Example of the calculation of electoral distance ($\tilde{r}_{AB}$) between Party A and Party B

Example 1			
	P_k^A	P_k^B	$(P_k^A - P_k^B)^2$
Group 1	0.4	0.25	0.0225
Group 2	0.35	0.35	0
Group 3	0.25	0.4	0.0225
			$\sum_{k=1}^{3}(P_k^A - P_k^B)^2 = .0445$
			$\tilde{r}_{AB} = \sqrt{\frac{.0445}{2}} = .15$
Example 2			
	P_k^A	P_k^B	$(P_k^A - P_k^B)^2$
Group 1	0.8	0	0.64
Group 2	0.2	0.2	0
Group 3	0	0.8	0.64
			$\sum_{k=1}^{3}(P_k^A - P_k^B)^2 = 1.28$
			$\tilde{r}_{AB} = \sqrt{\frac{1.28}{2}} = .8$

Aggregating the distance measures in a state. The second step in calculating PVP is to aggregate these measures of distance in ethnic bases of support for parties, something that requires that we take into account party size. Suppose that there are three parties, each with its own unique ethnic basis of support (so that the ethnic distance between each pair of parties is 1). Should the measure of ethnification be independent of the size of the parties? If one party received nearly 100 percent of the votes, would this be

the same as if the parties were of equal size, or if two parties both received roughly half the vote (with the third party representing a small, negligible group)? As Huber (2012) describes, there are two theoretical frameworks for thinking about how to address this aggregation issue. One is the fractionalization perspective, which is based on the idea that more numerous ethnic parties should lead to more governance issues. Suppose that each party received support from one group, and it was a different group for each party. Then from a fractionalization perspective, we should aggregate the party distance measures in a way that yields a larger ethnic parties score as the number of parties proliferate (and, holding the number of parties constant, as the ethnic bases of support for the parties become more dissimilar).

"Polarization" is the major alternative to fractionalization. It emerged in large part from Horowitz's (1985) argument that many ethnic groups is less problematic for stable governance than two equal-sized groups. Many small groups, the argument goes, should not lead to substantial ethnification of politics because no group can impose its will. And one very large group should not lead to substantial problems because the dominant group can impose its will. 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 quite supportive.[3] If we apply the same logic to parties, we should expect the impact of ethnicity on politics to be greatest when there are two parties, each with different bases of ethnic support, and each roughly the same size.

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 obviously disproportionately reward the top two parties, so the role or impact of ethically-based support for parties becomes less clear in situations where this vote support is fragmented across parties. In addition, in majoritarian systems, we should also

[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.

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.

To implement the polarization perspective, Huber (2012) draws on Esteban and Ray (1994), who were the first to develop a measure of income polarization. We need to adapt their polarization measure to a situation where the distance between two parties varies from 0 to 1, as does our measure of $\tilde{r}_{ij}$. To this end, and drawing in part on Desmet et al. (2009), Huber (2012) applies the polarization perspective to party ethnification by defining Party Voting Polarization as

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

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. And holding the number and size of parties constant, the measure increases with increasing distance in the ethnic basis of support for parties (i.e., increasing $\tilde{r}_{ij}$). PVP equals its maximum of 1 when there are two parties, each receiving 50 percent of the vote, and each receiving their support from a single (different) ethnic group.

Next consider our central explanatory variable. To measure the degree to which group identity proxies for class, we use India's NES 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, there is a strong correlation of group and class because the average incomes of group members are strongly differentiated from each other. When BGI is small, there is very little correlation of class and group because the average incomes of groups are roughly the same. The formal definition is

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

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, 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 developing countries that employ various asset indicators to gauge economic well-being (see, for example, Filmer and Pritchett (2001), McKenzie (2005) and Huber, Orgazalek and Gore (2012)). Specifically, the Indian election surveys of 2004 and 1999 ask individuals if they own particular items. Both surveys contain the following 7 variables that are associated with economic well-being:

- 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 "income" using educational attainment rather than assets. This is an additional strategy we follow below, one which we argue also helps to address concerns about the direction of causation between PVP and BGI.

4 Ethnic politics in India

A central challenge in studies of ethnic politics in India is that individuals often have multiple identities, with political considerations playing a central role in determining which specific identities become politically salient in a given polity at a given point in time. Scholars have focused on three different identities that have become salient political cleavages: religion (e.g., Brass 1993, Wilkinson 2004, Jaffrelot 2005 and Tachil 2010), caste (e.g., Srinivas 1962, Rudolph 1965, Krishna 2003 and Chandra 2004) and subcaste, or *jati* (e.g., Chhibber 1999, Chandra 2004 and Dunning and Nilekani 2013).[5] Although there exists

[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

this menu of identities around which parties and individuals can mobilize, previous studies have tended to focus on one type of group identity or another, leaving unanswered the question of whether one type of identity is in fact associated with the greatest degree of party system ethnification. Examining which group type leads to the greatest party system ethnification will therefore answer an important question about Indian electoral politics.

Consequently we focus on the effects of religion, caste and sub-caste membership on party system ethnification. These ethnic categories in India are related to each other, with *jati* (hereditary groups based in large part on historical occupational categorization) nested within broad caste umbrellas (broad social status categories), and caste umbrellas nested within religion (even non-hindu citizens in many instances claim to have a *jati* or a caste).

The hindu-muslim cleavage in Indian politics has been well documented by scholars and has generated much theoretical debate. Brass (1993) argues that the vote for the hindu-nationalist Bharatiya Janat Party (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 therefore examine the extent to which party systems are ethnified by religion across the states by examining the voting behavior of any religious group with more than 20 respondents in the state in the NES surveys.

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." Caste is a "hereditary, endogamous, usually localized group, having a traditional association with an occupation and a particular position in the hierarchy of castes. Relations between castes are governed, among other things, by the concepts of pollution and purity, and generally, maximum commensality occurs within the caste" (Srinivas 1962, 3). Given the

as politically salient any more.

central role of occupation in caste categorization, some scholars have claimed that the link between caste and class "is nothing more than a sociological axiom" (Jayaram 1996, 82). If castes are simply hereditary occupational groups, using Indian caste groups to argue that inequality between ethnicity leads to greater party system ethnification might make the argument specific to the Indian case and India an especially convenient case to test this theory.

This concern is not valid. In the post-independence period, economic reforms, growing urbanization and transformations in the agricultural economy have dramatically altered the relationship between caste and occupations to the extent that it is no longer possible to tell the caste of a person by their occupation (Karanath 1996). While castes have been a principle means by which groups have mobilized politically to make claims on the state, decades of economic changes and reforms have also considerably weakened the role of castes in the economy. The most enduring facet of castes however may be in the social realm. Caste membership continues to play a key role in guiding religious and marriage customs, with endogamy being the norm across the country. In addition, while groups have used caste to mobilize their co-ethnics, they have also attempted through the adoption of ritualistic norms of upper castes to improve their social status leading to a persistence of the caste-system in social life. The caste system's link with occupational groups has considerably weakened in today's context and castes more closely resemble ascriptive ethnic kinship categories.

What groups do we mean when we refer to "Caste"? Caste categories could refer to the *varna* system, where 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). In this paper, however, we examine caste categories in two other ways. Using the question in the NES 2004 and 1999 on "What is your caste group?", we organize respondents into seven categories based on recent research in Indian politics that identifies these as politically salient

ethnic categories (Heath 2005, Chhibber, Jensenius and Suryanarayan 2012 and Dunning and Nilekani 2013). Respondents who provide a caste regardless of their religious affiliation are organized into the following categories: upper castes, peasant castes, upper backward castes, lower backward castes, schedules caste, scheduled tribes and muslims.

Next we use another question from the NES which asks "What is your Caste/Jati-biradari/Tribe name?" to categorize respondents into their sub-caste or *jati*. These sub-castes vary across states and regions in India and are a closer reflection of local, endogamous, descent-groups. This question therefore allows us to probe the effects of state-level sub-caste group membership on party system ethnification.

A focus on variation across the Indian states is not only intrinsically important, it also has significant advantages for understanding the more general relationship between group-based inequality and ethnic politics. First, exploring this argument within a single country avoids some of the difficulties associated with cross-national studies of ethnicity. In particular, across countries, the institutional context (like electoral rules and decentralization) can vary considerably, influencing the ethnification of electoral politics (Huber 2012); the substantive meaning and importance of ethnic categories can vary considerably; and economic differences across groups can take on quite different meanings in different societies. Across the Indian states, the electoral rules and the menu and meaning of ethnic identities are largely the same (even though the categories themselves might vary), and while the states differ in their economic development, individuals across the states find themselves in essentially the same political and economic framework.

Second, by examining the Indian elections in 1999 and 2004 we can capitalize on recent changes in the country's national party system and allay concerns about the causal direction of the relationship between party system ethnification and inter group inequality.[6] When parties with an ethnic base take control of government, we might expect them to adopt policies that influence inter-group inequality by favoring the groups that support

[6]We also argue below that we can address the endogeneity problem in part by looking at educational inequality.

them.[7] While observable changes in income distributions are very difficult to produce in short time periods, patterns of ethnic voting should have the greatest opportunity to influence inequality levels if the same groups stay in power long enough to substantially influence the average incomes of specific groups. We argue that this was not the case in India for two reasons.

On the one hand, there were important changes to the national party system in the 1990s. The "Congress System" (Kothari 1967), so named for a resilient agglomeration of interest groups and factions within the Congress Party, managed to dominate the national political landscape for the first four decades of the post-independence era. In fact, the Congress party won every election but one between 1947-1989.[8] The late 1980s and early 1990s, however, witnessed rampant intra-party factionalism as well as caste, religious and economic mobilizations that placed severe demands on the Congress party and its subsequent electoral fortunes. This led to the rise of regional and caste based parties, and led national parties such as the Congress and the BJP to modify their organizational structures, candidate nomination processes and electoral strategies to suit a state's political climate. Thus, to the extent that national election outcomes might affect group-based inequality differentially across states, it is important to recognize that the national election outcomes shortly before the elections we study represented a marked change from a long term status quo. It is therefore unlikely that these changes could have had a sharp impact on between-group inequality levels across the states in the short time between the upheaval in the party system and elections we study. And it is even more unlikely that the policy process at the national level would have produced outcomes that shape inequality between groups in a way that that mirrors the voting patterns of group members at the state level.

On the other hand, the early 1990s brought a change that for the first time introduced significant possibilities for states to influence inequality. In particular, although

[7] Kasara (2008) demonstrates that in Africa, in fact, the opposite occurs with ethnic leaders taxing their co-ethnics more heavily the non co-ethnics.

[8] The Congress party experienced a temporary setback in the general election of 1977 following the Emergency but came back to power in 1980.

the Indian states have always had constitutionally mandated policy subjects over which they have control, their role in economic policymaking began to increase substantially following major national level policy changes in 1991. That year, India began a new era of significant economic liberalization, which brought with it a more formal devolution of economic power to the states, and one which substantially increased the power of state-level political actors (Chhibber and Nooruddin 2004). Thus, if state-level electoral outcomes affected between-group inequality in 1999 and 2004, it must have done so extremely rapidly because there was not substantial autonomy to do so until after 1991.

In sum, there was a great deal of change in India in the time period immediately prior to the elections we study. These include the decline of the Congress party, the rise of regional and caste-based parties, the devolution of fiscal and economic planning in the post-1991 period and the ever increasing levels of discretion given to state administrative officials in managing and implementing federal projects. Together, these changes make it highly unlikely that national or state government election outcomes or policies could be producing state-levels of inter-group inequality that reflect state-level voting patterns by ethnicity.

5 Party ethnification in India using different definitions of group identity

If voters in India have multiple identities that can become relevant to electoral behavior, is there a specific identity that is associated with the highest level of party system ethnfication? We answer this question using the Part Voting Polarization (PVP) measure to explore patterns of party system ethnification across the Indian states. The National Election Studies of 1999 and 2004, each conducted in the aftermath of national parliamentary elections,

provide the core data for our analysis.[9] The NES surveys are distinctive for their large samples of voters across the Indian states. The final NES surveys are conducted face-to-face using a structured questionnaire and in the local dialect.[10] Consequently, the NES surveys are some of the most ambitious surveys undertaken on any single national polity. NES 2004, for instance, was conducted in 22 Indian languages and had 27,189 respondents.

We exclude two states from the analysis. Jammu and Kasmir has high levels of hindu-muslim tensions, and is the focus of an international conflict between Pakistan and India. Manipur, in the North East, has had a complex separatist movement dominating politics there since the late 1970s, leading to a strong presence by the Indian army. Given the highly unique situations in these two states, we feel it makes sense to present results that exclude them. This decision, however, does not affect the substantive results, which are essentially identical regardless of whether the two states are excluded.

We measure PVP in the states using three definitions of ethnic group – religion, caste and *jati* – that are calculated using respondents' self-identification into these groups. Which definition of group type yields the highest level of PVP? In the 41 state-specific surveys that we have across the 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 of the time) and religion produces the highest score only 3 times (7 percent of the time).

It is important to recognize 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 only 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 that implements Horowitz's (1985) idea that many

[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.

ethnic groups is less problematic for stable governance than two equal-sized groups (see Reynol Querol 2002 and Montalvo and Reyno Querol 2005). It 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.[11]

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, 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, whereas religion is the definition of group that is associated with the lowest levels of party system ethnification. This is particularly striking given that religion has on average the smallest number of groups and *jati* the largest number.[12] Given our measure incorporates the polarization perspective, this means that voters within and across *jati* are coordinating their votes to a very large extent.

In what follows, we will use the PVP measure using *jati* to determine if cross-states differences in party ethnification are associated with cross-state differences in between-group inequality. Before proceeding, however, it is worth noting that some might worry

[11]The formal definition of EP is from Reynol Querol 2002:

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

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

[12]For religion, the average number of groups is 2.7 and the effective number of groups is 1.5. For caste, these numbers are 6.7 and 4.2, and for *jati* they are 11.4 and 7.9

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.

Our discussion of India's caste structure in the previous section above pointed out that the link between *jati* and occupation in recent years has considerably weakened and that sub-castes now function more as endogamous kinship groups, which diminishes concerns in this regard. In addition, if the historic tie of *jati* groups to occupations results in economic homogeneity of group members, then we should expect that inequality within *jati* groups should be much lower than inequality in the state as a whole. We can test this by calculating the gini coefficient for each group and comparing these group-based ginis to the state-level ginis. For each state-year, Figure 1 plots the group ginis and the state ginis. The figure shows that it is simply not true that state-level inequality is significantly greater than group-level inequality. Instead we find that the inequality within the *jati* groups is often quite high and that in every state there are a large number of groups that have greater inequality than in the state itself. To quantify the difference between the ginis of the groups and the state ginis, note that the mean of the state ginis is .440, which is only .037 higher than the mean of the group ginis in a state. Thus, the average inequality within groups is very close to the average inequality within the states themselves. This finding is consistent with the works of scholars who have described the decline in the occupational homogeneity of caste groups since independence.

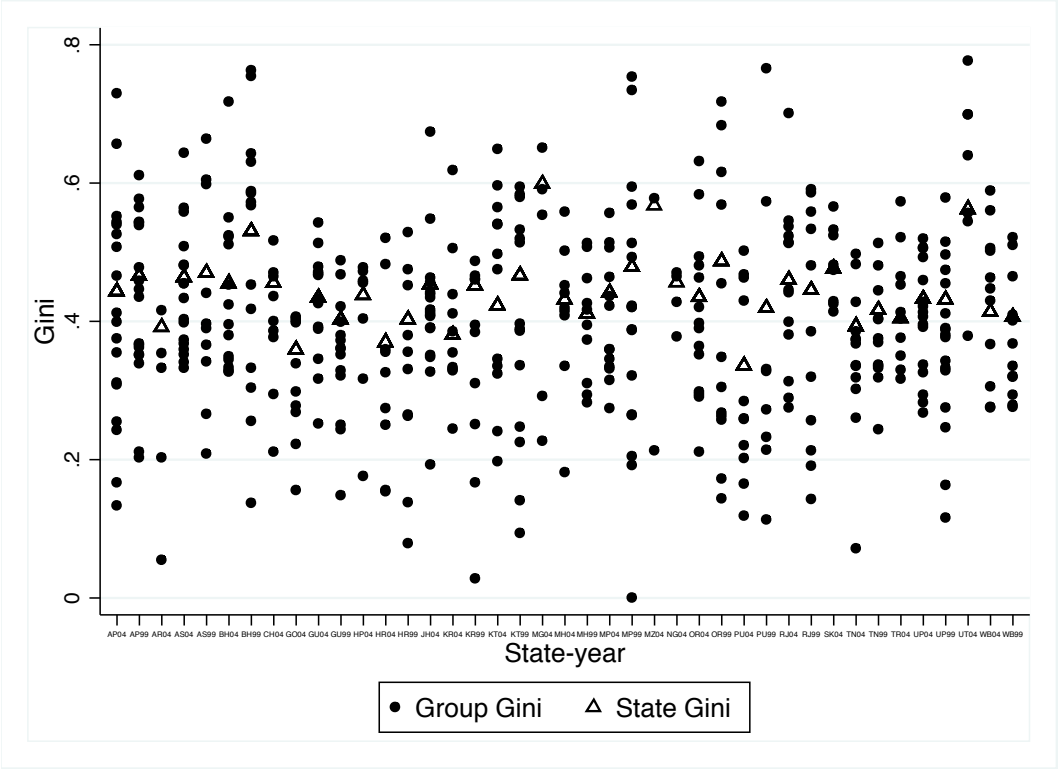

Figure 1: The Gini Coefficients of states and Jati groups

6 Variation in party system ethnification across the Indian states

We have argued that party system ethnification should be highest is systems where there is the strongest correlation of group and class, or where between-group inequality is highest. Is this true across the Indian states? Figure 2 shows a scatter plot of PVP and BGI. 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. In states where an individual's *jati* identity distinguishes him or her from others economically, support for political parties typically has a clear ethnic basis. In states where *jati* is most weakly correlated with economic well-being, the support bases of parties typically do not have a clear ethnic orientation.

Does the relationship depicted in Figure 2 survive when controlling in an OLS regression framework for other variables that could affect PVP? Table 2 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 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

[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.

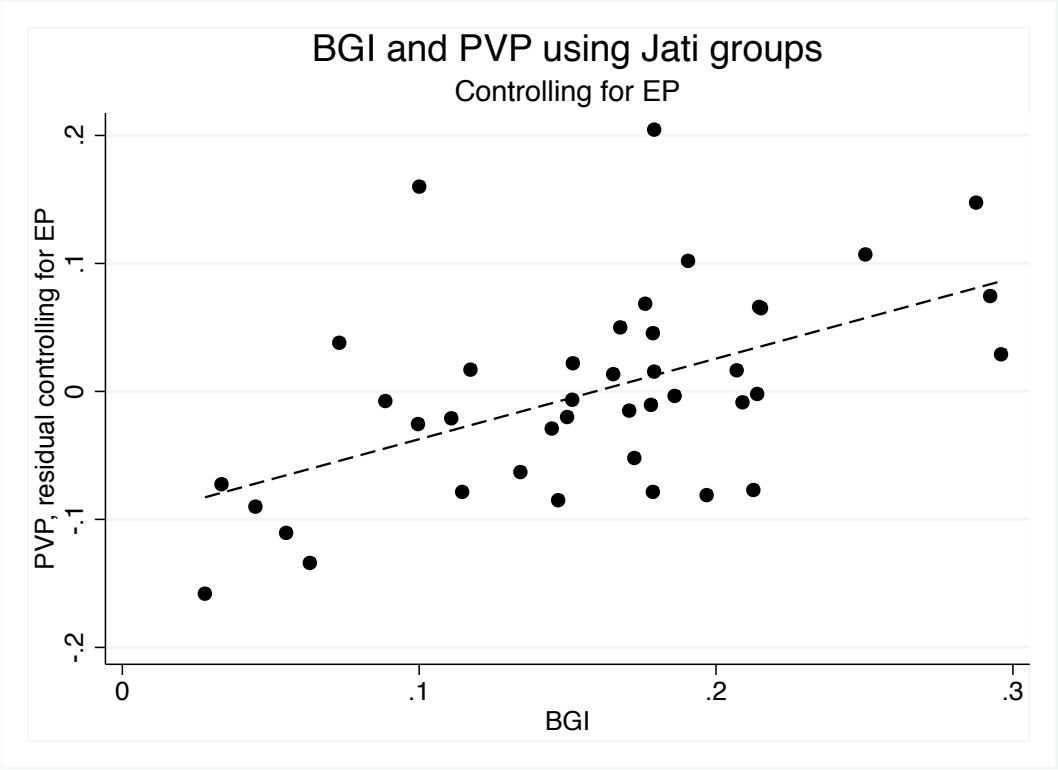

Figure 2: BGI and Party Ethnification

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. Also, 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 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 *jati* 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 dif-

[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

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

Table 2: OLS regressions of PVP on BGI using asset 'income' data

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

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.

ferences 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 2005). 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 coefficients for any pair of these variables. The result for BGI is not affected by the inclusion of these variables.[18]

It is possible that state governments with relatively high bureaucratic professionalism and capacity to provide goods and services to residents independent of their ethnic background 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

[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.

[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

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 will 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 it is not at all precisely estimated. The coefficients for BGI and for EP continue to be large, positive and precisely estimated.

Model 5 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. WGI is one of the three components of the gini decomposition, and it is a weighted sum of the gini coefficient for each group. Thus, as WGI increases, the economic heterogeneity of groups increases. The coefficient on WGI should therefore have a negative coefficient if such heterogeneity is associated with decreased party system ethnification. To test the effect of BGI when WGI is also included, we also need to include Overlap (a residual component of the gini that is related to stratification, see Yitzhaki and Lerman (1991)). 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. It is substantively important that the coefficient for BGI is precisely estimated and the coefficient for WGI is not – this suggests that economic heterogeneity does not diminish any relationship that exists between group identity and voting behavior.

Measuring inequality in educational levels. Using the "income" data based on individual assets, there exists a strong positive association between BGI and party system ethnification, an association that is robust to the inclusion of a wide range of controls. The asset-based income variable is attractive, not only because it conforms to how scholars

have typically measured inequality in poor countries, but also because India historically has been a highly agricultural country where economic well-being is strongly linked to land ownership. But education has become an increasingly important pathway to material well-being, making it useful to further probe the robustness of these results by estimating inequality using educational attainment. We use a variable that taps respondents' educational qualifications on a 9 point scale.[20] Respondents' educational scores were then rescaled to their percentile rank (thus ranging from 0 to 100), which was then used to measure educational inequality along with the three components of the gini.

Measuring educational inequality provides a useful way to test the robustness of the regressions presented above as educational attainment measures a different form of inequality. Using educational inequality also provides additional leverage in understanding the direction of causation between PVP and BGI. Educational attainment will change more slowly in a state than changes in asset variables, like owning cell phones or pumping sets, making it is more difficult to argue that ethnic voting outcomes influence contemporaneous educational inequality. Thus, if we find strong results for BGI using education, we can improve our confidence that it is BGI that is influencing voting patterns rather than voting patterns influencing BGI.

Table 3 presents the results for the models in Tables 2 but using education to measure inequality. The results for the BGI variable are, if anything, even stronger than the results using asset indicators to measure inequality. In each model, the coefficient on BGI is large, positive, and extremely precisely estimated. We find no precisely estimated coefficient for any other variables except for EP.

[20]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.

Table 3: OLS regressions of PVP on BGI using educational inequality

	(1)	(2)	(3)	(4)	(5)	(6)
BGI (edu.)	0.590***	0.613***	0.528***	0.559***	0.544***	0.573***
	(0.166)	(0.175)	(0.177)	(0.180)	(0.191)	(0.182)
GDP/cap(ln)		0.060				
		(0.134)				
Regional Parties			0.134			
			(0.310)			
Hindi 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)
EP	0.507***	0.489***	0.549***	0.488***	0.503***	0.513***
	(0.165)	(0.172)	(0.172)	(0.172)	(0.167)	(0.169)
Gini-BGI (edu.)	-0.190	-0.193	-0.252	-0.202	-0.196	
	(0.172)	(0.174)	(0.200)	(0.184)	(0.175)	
2004	-0.190	-0.190	-0.144	-0.252	-0.205	-0.200
	(0.302)	(0.305)	(0.310)	(0.329)	(0.306)	(0.309)
Constant	0.157	0.136	0.094	0.215	0.087	0.100
	(0.314)	(0.321)	(0.367)	(0.344)	(0.347)	(0.393)
Adj. R-squared	0.364	0.350	0.343	0.342	0.351	0.347
N	41	41	41	40	41	41

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.

7 Group level analysis with state fixed effects

We have found a strong positive relationship between the ethnification of parties and between-group inequality: when inequality between groups is largest, the ethnic basis of support for parties is strongest. This suggests there may be a strong class component of ethnic politics in India, one that has been previously ignored.

But though the ethnification of parties is strongest when ethnicity and average income are strongly correlated, this does not automatically imply that party ethnification carries with it a strong class basis. High ethnification, for example, could be associated with members of poor *jati* supporting different parties than members of rich *jati*, 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 *jati*s and other parties drawing on support from other rich and poor *jati*s. 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. The basic idea is to measure "voting distance" between every pair of groups in each state. If this distance is small, the patterns of support for parties are similar for the two groups. As the distance grows, the parties supported by members of one group diverge from those supported by members of the other group. Our goal is to estimate whether this electoral distance is related to income differences between the two groups. If income differences between two groups are associated with divergent voting patterns by members of the two groups, then group-based and class-based voting are strongly related, providing evidence that inequality between *jati* groups is leading to a strong class element of ethnic politics.

This approach also allows us to address a limitation with the state-level analysis. Although the cross-state regressions are run with a wide variety of controls, such analyses always run the risk that some omitted state-level variables are affecting the results. The

group-level analyses allay these concerns by including state fixed effects.

We draw on Huber (2012) to 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. 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 $\bar{r}_{ij}^m$ be the measure of electoral distance between groups i and j in state m. Then

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

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 state m.

Our goal is to understand if $\bar{r}_{ij}^m$ increases with the income differences between groups. We therefore regress $\bar{r}_{ij}^m$ on "Income Difference," which is the absolute difference in the mean income of groups i and j. Table 4 presents our results. In each model the dependent variable is the measure of $\bar{r}_{ij}^m$, 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

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

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 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. If members of *jatis* from the same caste have non-economic incentives to vote together, then we should expect Same Caste to have a negative coefficient – that is, voting differences between members of two *jati* groups from the same caste should be smaller than voting differences between *jati* groups from different castes. And if being from the same caste mitigates the effect of income differences on voting differences, the interaction term should have a negative coefficient.

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 the model, then, 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 Poor Income across castes. The same is true for rich income. We do find, however, that Same Caste has a negative sign and is relatively precisely estimated (significant at the .10 level).

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

Table 4: 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

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. We also find that the interactions between the income variables and Own Caste are not precisely estimated and that the Rich Income interaction with Own Caste has the wrong sign. Again, however, there is no statistically significant difference in the marginal effects of Poor Income for within or across caste comparisons, or for the marginal effects of Rich Income when making these comparisons.

As with the state-level regressions, for the group-level analysis we re-estimated the models in Table 4 using educational attainment rather than "asset income" distance between groups. The results, presented in Table 5, are remarkably similar to those in Table 4. 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 Conclusion

Our empirical analysis shows that in India, the highest level of party system ethnification occurs when individual identity is linked to one's *jati* group rather than to one's umbrella caste or religion, and that there is substantial variation across the Indian states in ethnification levels using *jati*. These cross-state differences in ethnification cannot be explained by variables such as economic development or ELF that one might expect to be important. There is, however, a large and robust relationship between party system ethnification and between-group inequality. The states where parties have the strongest ethnic bases of support are also the states where inequality between groups is highest. Moreover, the group-level analysis shows that the voting patterns of members from two different groups diverge when the mean income of these groups diverge. Thus, in states where party systems have a strong ethnic basis of support, this support has a strong relationship to class differences.

Although our data do not allow us to demonstrate conclusively that the direction of causation runs from economic inequality between groups to party system ethnification, we have argued that the particulars of the Indian case we examine provide reason to doubt that voting outcomes could be shaping between-group inequality in the time period we study. We have also shown that our results are robust when we measure educational inequality, which should respond more slowly than "asset inequality" to changes enacted by

Table 5: 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

governing coalitions, providing further assurance about the likelihood of reverse causation. And the state level analysis with state fixed effects provides assurance about a related endogeneity concern, which is that our state-level findings may be due to omitted state-level variables that are correlated with between-group inequality. But though we believe that in the specific context we study there is little worry about reverse causality, we also believe that over the long-run, we should expect a reciprocal relationship between ethnic politics and inequality. Understanding these dynamics is an important topic for future research.

The results also have an important implication for how we study identity politics. Given the empirical patterns we have found, it seems clear that 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 non-immutable attributes of group members – like income or class – can be central to determining whether identity politics are useful electorally. It is therefore important that scholars attempt to disentangle the effect of group identity from that of economic well-being when studying identity politics. We would underline, however, that if group-based inequality influences the salience of ethnic identity in politics, this does not mean that ethnic identity is epiphenomenal. We find that the economic heterogeneity of groups does not affect group voting patterns, 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.

References

Alesina, Alberto, Stelios Michalopoulos & Elias Papaioannou. N.d. "Ethnic Inequality." 2012.

Baldwin, Kate & John D. Huber. 2010. "Economic versus Cultural Differences: Forms of Ethnic Diveristy and Public Goods Provision." *American Political Science Review* 104 (4):644–662.

Bates, Robert. 1983. State versus Ethnic Claims: African Policy Dilemmas. In *Modernization, Ethnic Competition and the Rationality of Politics.*, ed. D Rothschild & V.A Olunsorola. Westview Press chapter Modernization, Ethnic Competition and the Rationality of Politics.

Brass, Paul R. 2003. *The Production of Hindu-Muslim Violence in Contemporary India*. New Delhi: Oxford University Press.

Cederman, L. E., Nils B. Weidmann & Kristian Skrede Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Politcal Science Review* 105(3):605–683.

Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed : Patronage and Ethnic Headcounts in India*. Cambridge studies in comparative politics Cambridge, UK ; New York: Cambridge University Press.

Chhibber, Pradeep. 1999. *Democracy without associations : transformation of the party system and social cleavages in India*. Ann Arbor: University of Michigan Press.

Chhibber, Pradeep, Francesca Refsum-Jensenius & Pavithra Suryanarayan. Forthcoming 2013. "Party Organization and Party Proliferation in India." *Party Politics* .

Chhibber, Pradeep & Irfan Nooruddin. 2004. "Do Party Systems Count: The Number of Parties and Government Performance." *Comparative Political Studies* 37 (2):152–187.

Desmet, Klaus, Ignacio Ortuno & Shlomo Weber. 2009. "Linguistic Diversity and Redistribution." *Journal of the European Economic Association* 7(6):1291–1318.

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

Dunning, Thad & Lauren Harrison. 2010. "Cross-cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104 (1):21–39.

Esteban, Joan & Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica* 62:819–52.

Esteban, Joan, Laura Mayoral & Debraj Ray. 2012. "Ethnicity and Conflict: An empirical study." *American Economic Review* 102(4):1310–42.

Fearon, James. 2003. "Ethnicity and Cultural Diversity by Country." *Journal of Economic Growth* 8:195–222.

Filmer, Deon & Lant H. Pritchett. 2001. "Estimating Wealth Effects Without Expenditure Data-Or Tears: An Application to Educational Enrollments in States of India." *Demography* 38 (1):115–132.

Heath, Oliver. 2005. "Party Systems, Political Cleavages and Electoral Volatility in India A State-Wise Analysis, 1998–1999." *Electoral Studies* 24(2):177–199.

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

Huber, John D. 2012. "Measuring Ethnic Voting: Do Proportional Electoral Laws Politicize Ethnicity?" *American Journal of Political Science* 56 (4):986–1001.

Huber, John D., Thomas K. Orgazalek & Radhika Gore. N.d. "Democracy, targeted redistribution and ethnic inequality." Working paper: Columbia University.

Jaffrelot, Christophe. 2005*a*. *India's Silent Revolution: The Rise of Lower Castes*. Hurst & Co, London.

Jaffrelot, Christophe. 2005*b*. *The Sangh Parivar: A Reader*. Oxford University Press.

Jayaram, N. 1996. Caste and Hinduism: Changing Protean Relationship. In *Caste: Its Twentieth Century Avatar*. Viking, India.

Karnath, G.K. 1996. Caste in Contemporary Rural India. In *Caste: Its Twentieth Century Avatar*, ed. M.N. Srinivas. Viking, India.

Kasara, Kimuli. 2007. "Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa." *American Political Science Review* 101(1):159–172.

Krishna, Anirudh. 2003. "What is Happening to Caste? A View from Some North Indian Villages." *The Journal of Asian Studies* 62 (4):1171–1193.

McKenzie, David J. 2005. "Measuring Inequality with Asset Indicators." *Journal of Population Economics* 18 (2):229–260.

Montalvo, Jose G. & Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict and Civil War." *American Economic Review* 95 (3):796–816.

Posner, Daniel N. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi." *American Politcal Science Review* 98 (4):529–545.

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

Reynal-Querol, Marta. 2002. "Ethnic Polarization, Potential Conflict and Civil War." *Journal of Conflict Resolution* 46(1):29–54.

Rudolph, Lloyd. 1965. "The Modernity of Tradition: The Democratic Incarnation of Caste in India." *The American Political Science Review* 59 (4):975–989.

Rudolph, Susanne H. & Lloyd Rudolph. March 22, 1993. "Modern Hate." *The New Republic* pp. 24–29.

Singh, Prerna. 2010. "We-ness and Welfare: A Longitudinal Analysis of Social Development in Kerela, India." *World Development* 38(2):282–293.

Srinivas, M N. 1962. *Caste in Modern India and other Essays*. Asia Publishing House, Bombay.

Thachil, Tachil. N.d. "Can Social Services Win Votes? A Study of Hindu Nationalism and the Poor in India." APSA 2010 Annual Meeting Paper.

Wilkinson, Steven. 2004. *Votes and Violence*. Cambridge University Press.

Yitzhaki, Shlomo & Robert I. Lerman. 1991. "Income Stratification and Income Inequality." *Journal of Population Economics* 37(3):313–28.