Measuring Ethnic Voting: Do Proportional Electoral Laws Politicize Ethnicity?

John D. Huber Columbia University

I develop four related measures of the "ethnicization" of electoral behavior. Each measure increases as ethnic identity becomes more central to vote choice, but the measures differ along two theoretical dimensions. The first dimension contrasts a group-based perspective (which focuses on cohesion in the voting patterns of group members) with a party-based perspective (which focuses on the composition of groups supporting political parties). The second dimension contrasts a fractionalization perspective (which assumes that more groups or parties cause more problems) with a polarization perspective (which assumes that problems are greatest when there are two equal-sized groups or parties). Using survey data to implement the measures in 43 countries, the article shows that proportional electoral laws are associated with lower levels of ethnicization—the opposite of what is widely assumed. I argue that the lower levels of ethnicization in PR systems should be unsurprising.

Scholars and constitutional engineers engage in heated debates about the role political institutions play in mediating the effects of ethnic diversity on governance, with particular attention being paid to electoral laws. The debate about electoral laws starts with the widely shared premise that proportional representation (PR) politicizes ethnicity. Since parties are easy to form under PR, political elites can make appeals based on ethnicity, and voters can choose parties that represent their groups, even if their groups are small. Thus, ethnic groups should be highly salient to electoral behavior in PR systems (e.g., Reilly and Reynolds 1999; Rokkan 1970; Sisk and Reynolds 1998; Wilkinson 2004). The main disagreement concerns whether such politicization of ethnicity is a good thing. Some scholars argue for PR: by allowing each group—even small ones—to have a party that represents the group's interests, it is possible to avoid problems that occur when individuals become alienated because their group is not represented (e.g., Cohen 1997; Lijphart 1977, 1999). Other scholars disagree, arguing that the politicization of ethnicity occurring under PR must be avoided. The goal instead should be to diffuse the salience of ethnicity by forcing parties to seek electoral coalitions that span different groups, for example by adopting electoral rules that force vote pooling (e.g., Horowitz 1985, 1991).

Despite these prominent debates, there is actually little evidence about which types of electoral laws—or any other macro factors that vary across countries—are associated with the politicization of ethnicity during elections. The problem is that although there is considerable research on ethnicization of electoral politics within particular countries, there currently exists no measures designed to compare such ethnicization across countries with different party systems or different levels of ethnic diversity. The central goal of this article is to develop and defend such measures and to use them to explore the relationship between electoral laws and the ethnicization of electoral politics.

I define four related measures of ethnicization in electoral politics. Each is based on the idea that ethnicization increases when group identity becomes more salient to vote choice. The measures are therefore based exclusively on information about individuals' group identity

John D. Huber is Professor of Political Science, Columbia University, New York, NY 10027 (jdh39@columbia.edu).
An earlier version of the article was presented at "Democracy and Representation: Conference honoring the contributions of G. Bingham Powell and his 40 years of teaching and research in the Department of Political Science at the University of Rochester" 2010, Rochester, New York. I am grateful for research assistance from Tom Ogorzalek, for research support from the National Science Foundation, and for helpful comments from Kate Baldwin, Johanna Birnir, Daniel Bochsler, Dawn Brancati, Bernard Groffman, Simon Hug, and seminar participants at Columbia, Essex, LSE, University of Geneva, ETH Zurich, and the Hebrew University in Jerusalem. I am also grateful to the Russell Sage Foundation, where I was a visiting research scholar while writing an early draft of this article. All data and replication materials can be found at http://dvn.iq.harvard.edu/dvn/dv/jhuber.

American Journal of Political Science, Vol. 56, No. 4, October 2012, Pp. 986–1001

 DOI: 10.1111/j.1540-5907.2012.00601.x

and vote choice. As the correlation between group and vote increases, the ethnicization measures increase. But this simple idea can be implemented in a number of different ways, and the four measures proposed here are distinguished from each other using two dimensions that are central to understanding the theoretical pathways by which ethnicization can create governance problems.

The first dimension draws a distinction between group-based and party-based perspectives on ethnic politics. A group-based perspective focuses on the level of voting cohesion by ethnic groups at election time. The perspective reflects the theoretical premise that as group members' propensity to vote together increases, the salience of group identity to politics increases, as does the cohesion of group members' interests. A party-based measure is based on the premise that conflict between groups plays out through parties. It therefore examines the degree to which political parties represent specific groups. If all parties are supported by broad cross-sections of society, ethnicization is low, whereas if each party tends to be supported by one or more specific groups that do not support other parties, ethnicization is high.

The second dimension draws a distinction between fractionalization and polarization perspectives on group conflict. A fractionalization perspective is based on the theoretical premise that the problems with ethnic diversity increase as the number of groups (or parties, from the party-based perspective) increases. A polarization perspective is based on the premise that group-based (or party-based) governance problems are greatest when there are two equal-sized groups. The four measures are derived from different combinations of assumptions along these two dimensions: there are group-based measures invoking fractionalization and polarization properties and party-based measures invoking fractionalization and polarization properties.

The article uses survey data from over 40 countries to implement the four measures and describe their relationships to each other. I then use the measures to examine correlations between electoral laws and the politicization of ethnicity. The central empirical finding is that contrary to commonly held beliefs, ethnicization is lower in proportional representation systems than in majoritarian ones. Perhaps this should not be too surprising. In majoritarian systems, small geographically dispersed groups can have strong incentives to vote cohesively for larger catch-all parties in efforts to be pivotal in elections. And in majoritarian systems with geographically concentrated groups, ethnically oriented parties can have strong chances to defeat larger catch-all parties. So there is ample space for ethnicization in majoritarian systems. In PR systems, by contrast, because it is easy for parties to form, it is easy for multiple parties to target members of the same group, often on issues unrelated to group identity. This divides the group against itself. Thus, ethnic identity will often be more relevant to voting behavior in majoritarian than in PR systems.

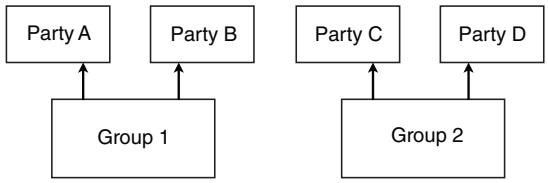
The article is organized as follows. The next section describes the four ethnicization measures. I then describe the data used for the analysis and present information on ethnic voting levels in 43 countries using 67 surveys. The empirical analysis describing the negative relationship between electoral laws and ethnicization follows. The article then explores in more detail the reasons why it should be unsurprising that the politicization of ethnicity is stronger in majoritarian than in PR systems.

Measures of Ethnicization in Elections

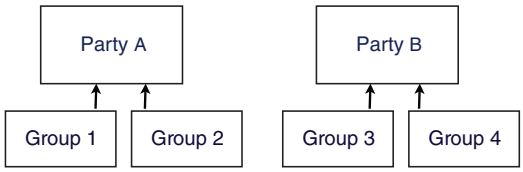
This section describes four country-level measures of the "ethnicization" of groups in elections. The measures are based on the voting behavior of group members, and the basic idea shared by each measure is simple: as the voting behavior of individuals becomes more closely correlated with group identity, the measures of ethnicization increase. If each group in a country has its own party, for example, group and vote are perfectly correlated and ethnicization is high. By contrast, if each group divides its support across all parties in the same proportions, ethnicization is low. The challenge is to implement such measures across countries that have large differences in the number and size of groups, as well as the number and size of parties. To this end, there are two important choices that one must make in developing any ethnicization measure.

Group-centered versus party-centered measures. The first choice concerns whether the measure should be group- or party-centered. A group-centered measure is based on the assumption that ethnicization increases when voting behavior by group members becomes more cohesive, making it easier to predict an individual's vote by knowing the individual's group. In the United States, for example, blacks vote overwhelmingly for the Democrats, making it easy to predict the vote of a black by simply knowing his or her race. If this link between race and vote diminished, with blacks splitting their vote between Democrats and Republicans, a group-centered measure (all else equal) would decline because knowing an individual was black would provide less information about vote choice. In Belgium, to take another example, each of the two language groups, the Flemish and the French,

FIGURE 1 Examples of Group- and Party-Based Ethnicization



Group-based ethnicization < Party-based ethnicization



Group-based ethnicization > Party-based ethnicization

support parties that receive virtually no votes from outside their language group. Thus, if one knows an individual's language, one knows which of several parties the individual might support, as well as which parties the individual would doubtfully ever support. But a group-centered conception of ethnicization assumes that the divisions among the French (or Flemish) lessen their potency in politics—if the French could coordinate on the same party, their cohesion would make them a more effective advocate for their group's interests.

A party-centered approach focuses on the composition of support for parties rather than on the cohesion of group behavior. Ethnicization from this perspective will increase when knowing an individual's vote choice provides clear information about group membership. If a party's support from various groups reflects the distribution of groups in society more generally, the party should not be viewed as ethnically politicized. But if support for a party comes disproportionately from one or more specific groups—which do not reflect the distribution of groups more generally—the politicization of ethnicity within that party should increase.

The group-centered and party-centered approaches are related, but they are conceptually distinct and can diverge. Suppose, for example, that there were two groups and each group split its vote across multiple parties, with parties receiving support from only one group, as depicted in the top panel of Figure 1. In this case, where parties receive votes from only one group but groups support more than one party, knowing an individual's vote choice reveals the individual's group, but knowing an individual's group provides only noisy information about the individual's vote choice. Thus, the group-centered perspective should indicate relatively lower ethnicization than the party-centered perspective. This is essentially what we find in Belgium, where the two language groups, the Flemish and French, spread their vote across a number of parties, but where all the parties receive support almost exclusively from members of only one language group. Knowing an individual's language group gives somewhat noisy information about vote choice in Belgium, but knowing an individual's vote choice gives accurate information about group membership.

Contrast this with the example depicted in the bottom panel of the figure, where there are four groups and two parties. Each group supports only one party, so one knows how a person votes if one knows the person's group. But each party receives support from more than one group, so knowledge of an individual's vote choice provides only noisy information about the individual's group. Thus, the group-centered perspective should indicate relatively higher politicization than the party-centered perspective. This is the type of situation found in South Africa, where a number of groups vote almost exclusively for the ANC. In this case, knowing group identity reveals very much about vote choice, but since a number of groups support the ANC, knowing that an individual supported the ANC does not reveal so much information about group identity.

Which approach is more appropriate obviously depends on one's theory about the pathways by which politicized ethnicity affects politics. Horowitz (1985), for example, worries about ethnic outbidding, where multiple ethnic parties try to outbid each other for ethnic support. If this were a principal concern, one might endorse a party- rather than group-centered approach. Similarly, if parties are the principal conduit by which group-based concerns are reflected in politics, then as the ethnic base of support for parties becomes more differentiated across parties, we should expect the ethnicization of politics to increase, again pointing to a party-centered approach. But if we want to know how salient groups are to individual behavior, then it seems quite reasonable to assume that when members of a group spread their support across multiple parties, this simply reflects that the group is divided. In such a case, one should assume that the salience of ethnicity to vote choice is lower—and that the group is a less homogenous force in politics—than in the case where members of the group support the same party. This would justify a group-centered approach. The reality is that existing theory and empirical research do not provide decisive justification for one approach over the other, and the strategy here will therefore be to develop and compare group- and party-centered measures.

Fractionalization- versus polarization-based measures. The second choice one must make in developing a measure of ethnicization concerns the appropriate assumption about how the number of groups influences governance outcomes. There are two theoretical frameworks that dominate the literature. The first, *fractionalization*, is based on the assumption that group-based problems increase with the likelihood that a given individual will encounter individuals from other groups. If members of groups have different interests, then group-based conflict will increase as group fragmentation increases. This perspective justifies the well-known index of ethno-linguistic fractionalization, or *ELF*, which measures the probability that two randomly chosen individuals will not belong to the same group. It is written as

$$ELF = 1 - \sum_{i=1}^{G} s_i^2, \tag{1}$$

where s_i is the proportion of individuals who belong to group i and there are G groups. If one believes that group-based problems will increase with the probability that individuals in society tend to be from different groups, this is an appropriate measure. For any number of groups, the measure increases as groups become more equal in size. And as groups proliferate, the measure increases, approaching 1 as the number of equal-sized groups grows large. High ELF has been negatively associated with a variety of governance outcomes, including long-run economic growth, the quality of government, redistribution, and investment (e.g., Alesina et al. 2003; Alesina and Glaser 2004; Easterly and Levine 1997; La Porta et al. 1999). And although considerable research has cast doubt on the relationship between ELF and civil conflict, recent research finds that higher ELF is associated with more civil conflict in cases where that conflict is over private (as opposed to public) goods (Esteban, Mayoral, and Ray 2012).

The second approach, *polarization*, builds on an observation in Horwitz's (1985) study that increasing ethnic diversity does not necessarily lead to more civil conflict. At the extremes of diversity—both high homogeneity and high heterogeneity—group-based conflict should be low. If there is one dominant group, for example, then conflict should be relatively rare, even if there are many smaller groups, because the dominant group can simply control outcomes. Similarly, if there are no dominant groups, but many smaller ones, then no group will be in a position to unilaterally impose its will, leading to minimal governance problems. But in a society where there is a large minority and a small majority, the potential for conflict should be high because the smaller group is reasonably powerful and is afraid of permanent exclusion from political power. Reynal-Querol (2002), whose main interest lies in the study of civil conflict, developed a measure of ethnic polarization ("*EP*") in an effort to capture this logic discussed by Horowitz (see also Montalvo and Reynal-Querol 2005a). The EP measure is based on the assumption that governance problems are maximized when there are two groups of equal size. It is written as

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

Like ELF, holding the number of groups constant, EP typically increases as groups become more equal in size. The measure takes a maximum value of 1 when there are two equal-size groups and decreases as the number of groups proliferates. Thus, if we start with a baseline of two equal-sized groups, EP will decline and ELF will increase as one of these groups is fractionalized into other groups. EP has been used primarily in empirical studies of civil conflict (e.g., Montalvo and Reynal-Querol 2005a), though Montalvo and Reynal-Querol (2005b) find it is also negatively related to economic development. Alesina et al. (2003), however, find that ELF has a stronger relationship with economic outcomes than does EP.

The electoral distance between groups or between parties. Developing a measure of ethnicization of elections that is comparable across countries therefore requires that one choose between a group- and party-centered approach, on one hand, and between a polarization and fractionalization perspective, on the other. But neither ELF nor EP explicitly incorporates information about the depth of differences between groups, a significant shortcoming if group-based differences are at the heart of governance problems. To develop a measure of ethnicization in elections, then, regardless of whether one adopts a fractionalization or polarization perspective, one needs to characterize the *electoral distance* between groups (from a group-based perspective) or between parties (from a party-based perspective). Incorporating information about electoral distance into fractionalization- and polarization-based measures can be done by drawing on existing measures that include generic information about the distance between groups.

From the fractionalization perspective, the measure for summarizing group-based differences was first proposed in Greenberg (1956). Let s_i be the size of group i, and let r_{ij} be the distance between group i and j on some dimension of interest. Then if there are G total groups, define r-Fractionalization (rF) as

$$rF = \sum_{i=1}^{G} \sum_{j=1}^{G} s_i s_j r_{ij}. \tag{3}$$

With $r_{ij} \in [0, 1]$, the measure has a minimum of 0 and will approach 1 as the number of equal-sized groups grows very large, with a large distance between them. Previous research has used linguistic differences between groups (Fearon 2003; Desmet, Ortuño, and Weber 2009) and economic differences between groups (Baldwin and Huber 2010) to measure r_{ij}. To implement *rF* in a group-based measure of electoral ethnicization, one needs a measure of the "electoral distance" between each group. And a party-based approach requires an analogous measure capturing the distance between the ethnic support bases of parties.

An analogous approach incorporates group-distance information into the polarization perspective. Esteban and Ray (1994) motivate their measure of polarization by proposing that individuals feel identification toward their own group proportional to the size of their group, and they feel antagonistic toward other groups proportional to their *economic* distance from them. The total polarization in society ("P") is a weighted sum of each group's identification with itself multiplied by its antagonism toward each other group (Esteban and Ray 1994, 830–31). Their measure is defined as

$$P = k\left(\sum_{i=1}^{m}\sum_{j=1}^{n} p_i^{1+\alpha} p_j \mid \bar{y}_j - \bar{y}_i \mid\right) \tag{4}$$

where k is some constant greater than 0, α is a parameter greater than 0 that can take on any value up to 1.6, and $\bar{y}_j$ is the mean income of group j. Desmet, Ortuño, and Weber (2009) defend a simple operationalization of P for any distance metric between two groups, r_{ij}. If $k = 4$ and $\alpha = 1$, then r-Polarization ("*rP*") is

$$rP = 4\sum_{i=1}^{G}\sum_{j=1}^{G} s_i s_j^2 r_{ij}. \tag{5}$$

When $r_{ij} \in [0, 1]$, $rP \in [0, 1]$. It takes its maximal value 1 when $r_{ij} = 1$ and there are two groups of equal size.

The key to developing group-based or party-based measures from the fractionalization and polarization perspectives is therefore to measure the "electoral distance" between any two groups or parties. First consider the group-based approach. The strategy employed here is to characterize differences in the voting patterns of any two groups. At one extreme, the distance should be zero if the voting patterns of the two groups are identical. This would occur when the proportions of the first group that supports each party are identical to the proportions of the second group that supports each party. In the United States, for example, if 90% of blacks supported the Democrats (and the rest supported Republicans) and 90% of Hispanics also supported the Democrats (with the rest supporting Republicans), the voting patterns of the two groups would be identical, and the electoral distance between these two groups should be zero. In this case, knowing an individual's group would convey zero information about the individual's vote. But as the proportion of blacks supporting the Democrats diverges from the proportion of Hispanics supporting the Democrats, the measure of electoral distance should increase. At the extreme, if 100% of blacks supported the Democrats and 100% of Hispanics supported Republicans, knowing an individual's group would provide perfect information about vote choice. This would represent maximal electoral distance between the two groups, or $r_{BlackHispanic} = 1$.

The specific metric used here for comparing the electoral distance between any two groups borrows from Gallagher's (1991) index of electoral disproportionality.[1] This disproportionality index measures the degree to which the proportion of seats each party receives in a system reflects the proportion of votes each party receives. It takes the value 0 if for all parties, the vote share exactly equals the seat share. The index moves toward 1 as the disjunction between seats and votes increases. In a two-party system, for example, if one party received no votes and all the seats, and the other party received all the votes and no seats, the disproportionality index equals 1. Similarly, to measure electoral distance, $\bar{r}_{ij}$, between groups, i and j, we can measure the degree to which the proportion of votes that group i gives to each party reflects the proportion of votes that group j gives to each party. Let

$$\bar{r}_{ij} = \sqrt{\frac{1}{2}\sum_{k=1}^{P}\left(V_i^k - V_j^k\right)^2}, \tag{6}$$

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 P total parties.

The measure of electoral distance between two groups, A and B, is therefore based on comparing the support by Group A for each party with the support by Group B for each party. At one extreme, if each group gave the same proportion of support to each party (i.e., $(V_A^k - V_B^k)^2 = 0$ for all parties), $\bar{r}_{AB} = 0$. At the other extreme, if each group gives all its support to one party that receives no support from the other group, $\bar{r}_{AB} = 1$. In general, as the distribution of support across the parties diverges, $\bar{r}_{AB}$ increases.

Table 1 provides two examples. In Example 1, 40% of Group A members support Party 1 and only 25%

[1] See Taagepera and Grofman (2003) for a useful analysis of why it makes sense to embrace this particular index.

TABLE 1 The Electoral Distance ($\bar{r}_{AB}$) between Group A and Group B

Example 1			
	$\mathbf{V_A^k}$	$\mathbf{V_B^k}$	$(\mathbf{V_A^k} - \mathbf{V_B^k})^2$
Party 1	0.4	0.25	0.0225
Party 2	0.35	0.35	0
Party 3	0.25	0.4	0.0225
			$\sum_{k=1}^{3}(V_A^k - V_B^k)^2 = .045$
			$\bar{r}_{AB} = \sqrt{\frac{.045}{2}} = .15$
Example 2			
	$\mathbf{V_A^k}$	$\mathbf{V_B^k}$	$(\mathbf{V_A^k} - \mathbf{V_B^k})^2$
Party 1	0.8	0	0.64
Party 2	0.2	0.2	0
Party 3	0	0.8	0.64
			$\sum_{k=1}^{3}(V_A^k - V_B^k)^2 = 1.28$
			$\bar{r}_{AB} = \sqrt{\frac{1.28}{2}} = .8$

support Party 3 (with the remaining 35% going to Party 2), whereas Group B gives 40% of its support to Party 3 and only 25% to Party 1. Thus, each group skews its support toward a different party, but the vote distributions are fairly similar to each other, and thus the electoral distance between these groups should be relatively small, and in this case $\bar{r}_{AB} = .15$. In Example 2, the vote distributions are much more distinctive from each other, with Group A giving 80% of its vote to Party 1 and none to Party 3, and Group B giving 80% of its vote to Party 3 and none to Party 1. Here each group's voting distributions are much more distinct from each other than was true in Example 1, and the electoral distance between the two groups is much higher, with $\bar{r}_{AB} = .8$.

Next, consider party-based measures. In the group perspective, $\bar{r}_{ij}$ characterizes the distance between the vote profiles of groups i and j. Analogously, from the party perspective, define $\tilde{r}_{ij}$ as the distance between the ethnic profiles of two parties, i and j. If the proportion of support that party i receives from each group is the same as the proportion of support that party j receives from each group, then the distance between the two parties will be zero. If all of i's supporters are from one group and all of j's supporters are from a different group, then the distance between these two parties will be 1. Thus, let

$$\tilde{r}_{ij} = \sqrt{\frac{1}{2}\sum_{g=1}^{G}\left(P_g^i - P_g^j\right)^2}, \tag{7}$$

where P_g^i and P_g^j are the proportion of supporters of parties i and j who come from group g. Table A in the supplemental materials provides examples of how $\tilde{r}_{ij}$ is calculated.

The measures. We can use $\bar{r}_{ij}$ to derive a measure of Group Voting Fractionalization, GVF, and Group Voting Polarization, GVP. Define

$$GVF = \sum_{i=1}^{G}\sum_{j=1}^{G} s_i s_j \bar{r}_{ij}, \tag{8}$$

where there are G total groups and s_k is the size of group k. Holding the number and size of groups constant, Group Voting Fractionalization will increase when the electoral distance between any two groups increases. It will approach 1 when $\bar{r}_{ij} = 1$ for all pairs of groups, groups are of equal size, and the number of groups becomes very large.

Similarly, define

$$GVP = 4\sum_{i=1}^{G}\sum_{j=1}^{G} s_i s_j^2 \bar{r}_{ij}. \tag{9}$$

Like GVF, holding the number and size of groups constant, Group Voting Polarization will increase when the electoral distance between any two groups increases. It will approach 1 when there are two equal-sized groups and $\bar{r}_{ij} = 1$ for these two groups.

Similarly, we can use $\tilde{r}_{ij}$ to derive a measure of Party Voting Fractionalization, *PVF*, and Party Voting Polarization, *PVP*. Define

$$PVF = \sum_{i=1}^{P}\sum_{j=1}^{P} p_i p_j \tilde{r}_{ij}, \tag{10}$$

where P is the total number of parties and p_k is the proportion of voters who support party k. Holding the number and size of parties constant, Party Voting Fractionalization will increase when the electoral distance between any two parties increases. It will approach 1 when $\tilde{r}_{ij} = 1$ for all pairs of parties, parties are of equal size, and the number of parties becomes very large.

Similarly, define

$$PVP = 4\sum_{i=1}^{P}\sum_{j=1}^{P} s_i s_j^2 \tilde{r}_{ij}. \tag{11}$$

Data for Estimating the Four Ethnicization Measures in 43 Countries

The measures of ethnicization—GVF, GVP, PVF, and PVP—provide four ways of estimating ethnicization,

based on whether one's theory requires a group- or party-based perspective, on one hand, and a fractionalization or polarization perspective, on the other. This section utilizes three existing cross-national surveys taken from 1995 to 2006—the World Values Survey (WVS, Wave 4), the Afrobarometer (Rounds 2 and 3), and the Comparative Study of Electoral Systems (CSES)—to create estimates of these four measures. These surveys make it possible to categorize respondents according to their group, as well as their vote intention. By including surveys from all three studies, it is possible to create a data set that includes a wide range of democratic systems that vary in their ethnic diversity, political and economic development, and political institutions.

To identify the relevant ethnic categories in the countries for which surveys exist, the main approach in this study follows the ethnic categories identified in Fearon (2003).[2] Fearon argues that any list of groups in a country should be based on "the idea that members and nonmembers recognize the distinction [on which group identity is based] and anticipate that significant actions are or could be conditioned on it" (198). He lists seven features that a "prototypical" ethnic group should have, based in large part on whether groups can be understood as "descent groups" (into which individuals are born) and whether groups are locally viewed as socially or politically consequential. Groups are, therefore, based on a range of characteristics, including religion, ethnicity, language, tribe, and race. While it is possible to debate Fearon's list of groups in particular countries, Fearon (2003) is an attractive source of group names because of his effort to apply reasonable criteria consistently across a wide range of countries.

In implementing Fearon's list of groups, I impose a rule for excluding surveys that do not adequately reflect the Fearon groups. The rule is to include surveys in the data set only if the percentage of the population (per Fearon's data) that cannot be assigned to any of Fearon's groups is less than 10%. For example, if the "purple" group is one of Fearon's groups and this group cannot be identified in a survey, then the survey is discarded if the purple group has more than 10% of the population in Fearon's data. If there are multiple Fearon groups that cannot be identified, then the survey is excluded if these groups together represent more than 10% of the population. This rule ensures that surveys are included in the study only if the group data from them can be used to reasonably approximate the groups identified by Fearon.

[2]Below, I also allow the definition of groups to emerge endogenously from survey categories.

Since the focus here is on voting, surveys are included only from countries that are at least nominally democratic (Polity2$>$0). This low bar for inclusion makes it possible to explore whether the quality of democratic institutions affects ethnic voting. This Polity2 inclusion rule, along with the "10%" selection rule for keeping surveys based on Fearon groups, results in the use of 68 surveys from 43 countries. The countries and surveys are listed in Table B of the supporting materials. Importantly, one can be confident that the surveys adequately represent the size of groups in the Fearon data because the surveys generate ethnic diversity scores that mirror quite closely the ethnic diversity in the Fearon data—as Baldwin and Huber (2010) show, the surveys produce an ELF for which the correlation with Fearon's ELF is .95.

Each survey has some form of "vote" variable that can be used to measure the distribution of support for each party. The CSES is a post-election survey, so it contains a question asking individuals which party they supported in the last election. I use the vote in the lower house election if it exists and use the presidential election vote otherwise. The Afrobarometer (Round 3) and the WVS surveys, which are not conducted post-election, ask which party the respondent would support if there were an election tomorrow. The Afrobarometer (Round 2) asks the respondent not about vote intent, but rather whether the respondent feels close to a particular political party. I use these various survey instruments to measure the proportion of respondents in each group who support each party.

The different wording of the "vote" questions—and of the timing of the surveys vis-à-vis elections—could create biases. Voters may not actually pull the lever for the party they say they feel closest to, for example, or they may be more inclined to say they support (or oppose) the incumbent between elections than right after the election. Despite these possible biases, this study uses the range of different surveys in order to bring as much data as possible to bear on the empirical relationships between macro political factors and ethnic voting. I also include controls for survey types and check the robustness of the results to different subsets of the data.

Comparing the Four Ethnicization Measures

The four measures of ethnicization should be closely related to each other empirically because each measure relies on the same information about voting behavior and group identity. The four measures should all increase,

for example, as individual voting behavior becomes more strongly correlated with group identity. But though the measures are based on the same information, they are also clearly distinct. So how, more generally, are the four measures related to each other?

Figure 2 plots the two fractionalization-based measures against each other (top panel) and the two polarization-based measures against each other (bottom panel). The two fractionalization-based measures are obviously correlated, although for most surveys, PVF is larger than GVF. But the measures can diverge substantially, as is particularly true for Belgium and Macedonia. Party-based fractionalization scores are large relative to group-based fractionalization scores when voting is highly correlated with groups (otherwise both scores would be low), and there are a large number of parties relative to the number of groups. We have already discussed this pattern in Belgium, where individuals vote only for parties in their language group, but where there are numerous such parties. A similar pattern exists in Macedonia. There are only two major groups, the Macedonians and Albanians, keeping the fractionalization-based score relatively small. Each group supports parties that receive no significant support from the other group. But each group also spreads its vote across several parties. Thus, if one knows the party an individual supports, one can be virtually certain of the individual's group. But if one knows an individual's group, there will be more uncertainty about party choice because members of each group support a range of parties.

Next, consider the polarization-based measures depicted in the bottom panel of Figure 2. Again, the group- and party-based measures are highly correlated, though they can diverge substantially at higher levels of ethnicization. But unlike in the fractionalization-based measures, countries like Macedonia and Belgium now have higher group-based measures than party-based measures. This, of course, is simply because group-based polarization measures increase as the number of "groups" goes to two, and thus countries like Belgium and Macedonia, with two groups and many parties, have higher scores in the group-based measures using the polarization framework.

Since the group-based measures are more sensitive to the number of groups and the party-based measures are more sensitive to the number of parties, a useful way to think about the relationships between all four measures is to consider how each measure is affected by the ratio of parties to groups. In the group-based measures, GVF will decrease as $\frac{\#parties}{\#groups}$ increases (because all else equal, group-based fractionalization measures increase with the number of groups). By contrast, GVP will increase as $\frac{\#parties}{\#groups}$ increases (because group-based polarization measures decrease as the number of groups increase, all else constant). In the party-based measures, PVF should increase as $\frac{\#parties}{\#groups}$ increases (because a party-based fractionalization measure should increase with the number of parties), whereas PVP should decrease with $\frac{\#parties}{\#groups}$.

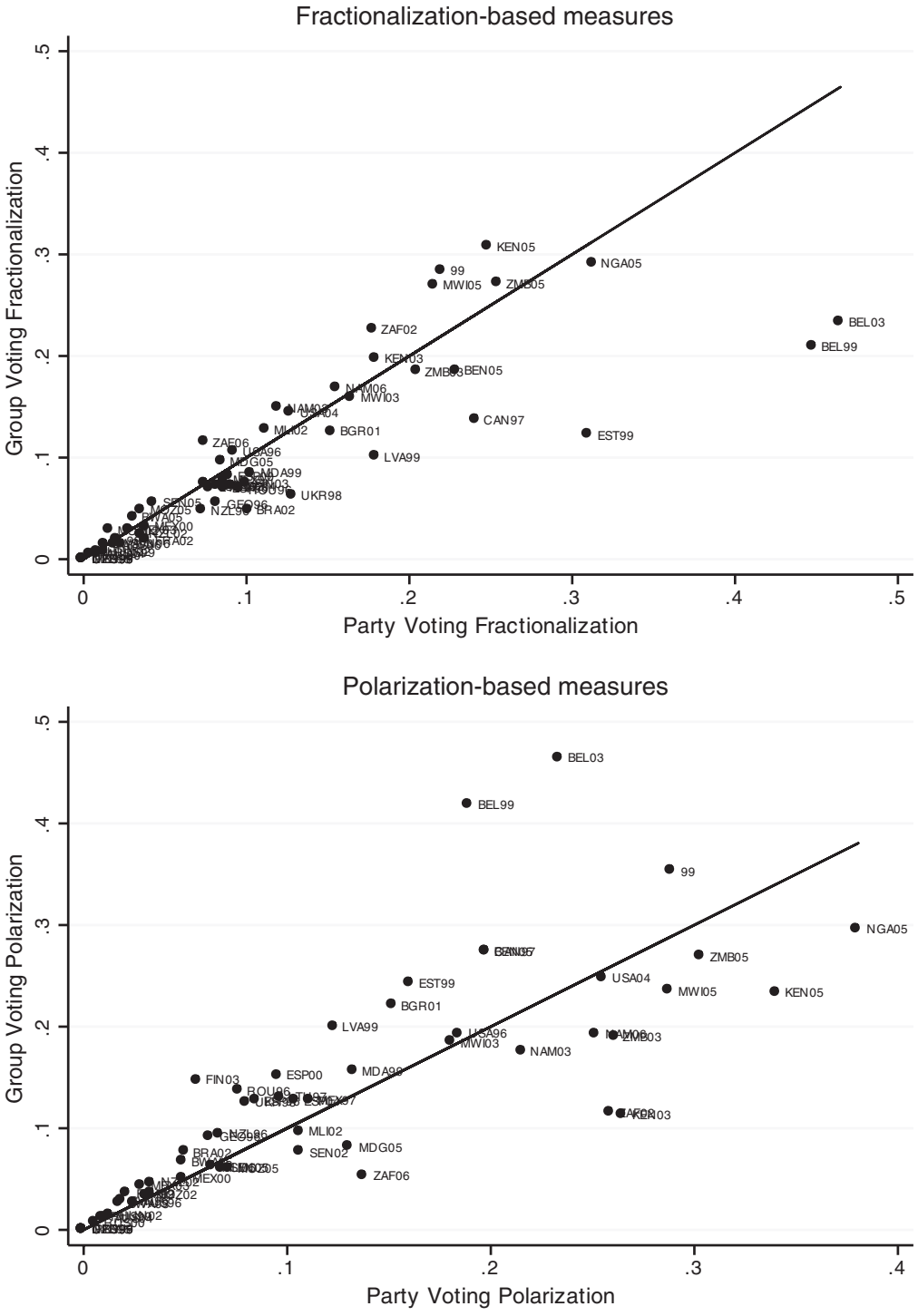
The empirical measures, then, respond in the ways one would expect to the number of groups or parties. What is crucial to observe, however, is that GVF and PVP each move in the same direction as $\frac{\#parties}{\#groups}$ increases, as do GVP and PVF. Thus, GVF and PVP should be strongly correlated, as should GVP and PVF. Table 2 shows the bivariate correlations of the two general diversity measures (ELF and EP),[3] the four ethnicization measures, and $\frac{\#Parties}{\#Groups}$. Several points are worth underscoring. First, adding information about the electoral distance between groups or parties clearly leads to different measures of ethnicization then generic ELF or EP measures. This is particularly true for EP, where the strongest correlation with any of the four ethnicization variables is only .41. Second, the four ethnicization measures proposed here are strongly correlated with each other. The weakest pairwise correlation is .82, and the strongest correlations are the two pairs described above, both above .9. Finally, although the four measures move in systematic ways with $\frac{\#Parties}{\#Groups}$, the measures are hardly mere proxies for this ratio. Indeed, the correlation between $\frac{\#Parties}{\#Groups}$ and any of the four measures has a maximum of .36 (in absolute value). Thus, the voting patterns of group members obviously play a central role in accounting for differences in the values of the ethnicization variables.

The Electoral Law and Ethnicization

This section employs the four measures of ethnicization to explore empirically the relationship between electoral laws and ethnic voting. As noted in the introduction, it is commonly assumed that proportional electoral laws politicize ethnicity by making it easier for parties to form by appealing to ethnic groups, even small ones. To test this relationship, I consider three measures of the electoral law. The first is the log of average district magnitude (DM), where DM measures the mean size of all electoral districts in a country. Two data sets were consulted to create this commonly used measure: Johnson and Wallack (2007) and the World Bank's Database of Political Institutions (Beck et al. 2001). In cases where there were disagreements between these two sources, I conducted further research to determine the mean district magnitude. I take the log because beyond a certain threshold,

[3] These measures are from Fearon (2003).

FIGURE 2 The Four Measures of Ethnicization



the variable should have no further effect on the number of parties. If PR systems encourage the politicization of ethnic politics, DM should have a positive coefficient.

It is important to recognize that a country's average district magnitude can mask attributes of electoral laws that make it easy for parties to form by appealing to groups. Countries like Russia (during the time period of this study) have an average district magnitude of only 2, but half of the legislators are elected from one large PR district (with the rest being elected in

TABLE 2 Cross-Correlations

Variables	ELF	EP	GVF	GVP	PVF	PVP	$\frac{Parties}{Groups}$
ELF	1.00						
EP	0.45	1.00					
GVF	0.67	0.22	1.00				
GVP	0.38	0.41	0.83	1.00			
PVF	0.47	0.35	0.87	0.93	1.00		
PVP	0.65	0.25	0.96	0.83	0.82	1.00	
$\frac{Parties}{Groups}$	−0.59	−0.16	−0.31	−0.04	0.01	−0.36	1.00

single-member districts). Since one large PR district eases party formation, a more useful measure of the electoral law than DM taps the total proportion of all legislators who are elected by PR. Johnson and Wallack (2007) provide such a measure, which I label PropMMD. Finally, it is useful to consider a simple indicator that takes the value 1 if the country uses any form of PR. Such a measure implies that the mixed systems like Russia are now simply coded as PR systems. There are 45 observations in the data with PR systems and 23 (from 14 countries) with majoritarian electoral systems. The countries coded as majoritarian (with number of surveys in parentheses) are Australia (3), Bangladesh (1), Botswana (2), Canada (2), France (1), India (1), Kenya (2), Madagascar (1), Malawi (2), Mali (2), Nigeria (1), United States (3), and Zambia (2).

Results using the Fearon-defined groups. Table 3 presents estimates of the relationship between the ethnicization measures and the three measures of electoral law. Each model regresses one of the ethnicization scores on a measure of electoral law, along with the appropriate group-based variable (ELF or EP, depending on the dependent variable), survey indicator variables (to account for biases that may be inherent to the different wording used for the "vote" questions across the surveys, with the CSES survey as the omitted category), and a number of controls. As noted above, the measures respond to the ratio of parties to groups, so this variable is included as a control. In addition, since it is widely held that the politicization of ethnicity should be particularly likely in early stages of democratic development (e.g., Birnir 2007; Lijphart 2002, 38), the Polity2 measure of political development is included. Similarly, since low levels of economic development could also increase the salience of ethnicity in politics (as groups struggle for resources), the models include a measure of national wealth (the log of GDP/capita using purchasing power parity from the World Development Indicators of the World Bank). If political or economic development diminishes the salience of ethnicity in politics, the Polity and GDP variable should have negative coefficients. Scholars have widely argued that decentralization can influence the politicization of ethnicity, though they have not agreed on how this works (Brancati 2009). I therefore include a federalism indicator variable from Treisman (2002), supplemented by my own research for missing cases. Finally, if individuals from the same group tend to live together, and to be exposed mostly to members of their own group, then it should be more likely that they will form group-specific viewpoints and interests, and thus that they will vote together with their own ethnic group. The models therefore include a measure of the Geographic Isolation of groups, which measures "the extent to which minority members are exposed only to one another" (Massey and Denton 1988, 288). Geographic Isolation, which can range (in theory) from 0 to 1, is calculated using the region variable included in each survey. Details are provided in the online supplemental materials.

I estimate the models using OLS with standard errors clustered by country. Looking across all 12 models in the table, for each measure of ethnicization and each measure of electoral law, the coefficients on the electoral law variables are always negative. When either GVF (Models 1a–1c) or PVF (Models 3a–3b) is the dependent variable, the negative coefficients are measured precisely, and when PVP is the dependent variable, the electoral law coefficients are also measured precisely, except when the log of DM is used to measure electoral laws (Models 4a–4c). Only when GVP is the dependent variable (Models 2a–2c) are the electoral law coefficients measured usually with considerable error. These regressions, then, show no evidence whatsoever that there is a positive correlation between proportional electoral laws and ethnicization. On the contrary, particularly for those measures that increase with the number of groups, PR seems associated with less ethnicization.

Considering the controls, Polity2 has the expected negative sign but is estimated with considerable error. GDP has a positive sign and is at times significant, but this does not seem robust. Federalism has a consistent negative sign and is usually relatively precisely estimated, suggesting that federalism is associated with less politicized ethnicity. Geographic isolation and $\frac{\#groups}{\#parties}$ have a sign that depends on the dependent variable, with both variables taking a positive sign for the fractionalization-based variables and a negative sign for the polarization-based variables.

To further explore the robustness of the results for proportional electoral laws, I estimated the models in Table 3 on subsets of the data. I did not estimate models for GVP, which yields very large standard errors on the electoral law coefficients in Models 2a–2c, and I

TABLE 3 Electoral Laws and Ethnicization

	(1a) GVF	(1b) GVF	(1c) GVF	(2a) GVP	(2b) GVP	(2c) GVP	(3a) PVF	(3b) PVF	(3c) PVF	(4a) PVP	(4b) PVP	(4c) PVP
Avg. DM (ln)	−0.236*			−0.190			−0.273**			−0.210		
	(0.121)			(0.148)			(0.109)			(0.181)		
Prop. MMD		−0.296**			−0.201			−0.235**			−0.326*	
		(0.110)			(0.155)			(0.110)			(0.179)	
PR			−0.686**			−0.665*			−0.637**			−0.770*
			(0.259)			(0.334)			(0.259)			(0.394)
ELF	1.044***	1.084***	1.061***				1.300***	1.312***	1.303***			
	(0.173)	(0.168)	(0.173)				(0.209)	(0.213)	(0.212)			
EP				0.541***	0.575***	0.590***				0.241	0.309*	0.300*
				(0.174)	(0.186)	(0.179)				(0.160)	(0.178)	(0.167)
$\frac{\#parties}{\#groups}$	0.043	0.053	0.043	0.024	0.028	0.032	0.194*	0.190	0.187	−0.068	−0.047	−0.058
	(0.075)	(0.080)	(0.079)	(0.112)	(0.112)	(0.115)	(0.106)	(0.113)	(0.114)	(0.096)	(0.088)	(0.091)
GDP/capita(ln)	0.221	0.179	0.228	0.166	0.124	0.174	0.354**	0.300*	0.347**	0.047	−0.005	0.058
	(0.173)	(0.152)	(0.162)	(0.267)	(0.254)	(0.269)	(0.170)	(0.158)	(0.166)	(0.330)	(0.323)	(0.333)
Polity2	−0.040	−0.003	−0.051	−0.037	−0.002	−0.042	−0.097	−0.058	−0.101	−0.035	0.014	−0.042
	(0.097)	(0.081)	(0.084)	(0.141)	(0.130)	(0.127)	(0.105)	(0.091)	(0.092)	(0.152)	(0.134)	(0.134)
Federalism	−0.287	−0.420	−0.407	−0.141	−0.223	−0.260	−0.436	−0.528*	−0.538*	0.184	0.047	0.045
	(0.264)	(0.259)	(0.262)	(0.392)	(0.398)	(0.415)	(0.286)	(0.279)	(0.282)	(0.416)	(0.451)	(0.449)
Geo. Isol.	0.293*	0.302**	0.294*	0.242	0.260	0.248	0.457**	0.478***	0.466***	−0.047	−0.036	−0.042
	(0.157)	(0.140)	(0.147)	(0.191)	(0.189)	(0.195)	(0.173)	(0.172)	(0.172)	(0.197)	(0.191)	(0.202)
WVS	0.366*	0.319*	0.264	0.549*	0.510	0.461	0.583**	0.524*	0.475*	0.441	0.404	0.342
	(0.185)	(0.181)	(0.183)	(0.313)	(0.312)	(0.310)	(0.266)	(0.260)	(0.265)	(0.264)	(0.267)	(0.255)
Afrobarometer 2	−0.110	−0.219	−0.425	0.635	0.609	0.394	−0.471	−0.555*	−0.761**	0.980	0.964	0.704
	(0.395)	(0.408)	(0.433)	(0.656)	(0.672)	(0.675)	(0.325)	(0.316)	(0.344)	(0.705)	(0.763)	(0.733)
Afrobarometer 3	0.254	0.186	−0.040	1.069	1.078	0.844	−0.217	−0.253	−0.478	1.261*	1.285*	1.001
	(0.389)	(0.392)	(0.388)	(0.636)	(0.658)	(0.638)	(0.360)	(0.367)	(0.386)	(0.662)	(0.708)	(0.645)
Constant	−1.853*	−1.721	−1.176	−1.623	−1.517	−1.074	−2.816**	−2.583**	−2.107*	−0.474	−0.419	0.149
	(1.100)	(1.045)	(0.989)	(1.855)	(1.846)	(1.787)	(1.171)	(1.170)	(1.090)	(2.197)	(2.254)	(2.154)
Adj. R^2	0.474	0.508	0.516	0.114	0.119	0.163	0.456	0.443	0.468	0.188	0.238	0.250
N	67	67	67	67	67	67	67	67	67	67	67	67

Note: OLS models with clustered standard errors (by country) in parentheses. The dependent variable is listed under column number.
To facilitate comparisons, all continuous variables are standardized to have a mean of 0 and a standard deviation of 1. *p < .10, **p < .05, ***p < .01.

focused on one measure of electoral law, the proportion of legislators from multimember districts. I considered three different subsets of the data. First, there are a number of very homogenous countries in the data and if such countries happen to be correlated with the electoral law, one might worry that these cases are driving the results. I therefore estimated the models without the seven surveys for which ELF is less than .2. Second, in Africa there are a number of countries with dominant parties receiving a huge percentage of the vote. In such cases, ethnicization will obviously be low, which is a concern if these dominant-party countries are correlated with the electoral law. I therefore further eliminated the 10 surveys in which there exists a party receiving more than 70% of the vote. Finally, the Afrobarometer 2 survey is the one survey that asks respondents not whom they would vote for (or have voted for), but rather whether they feel close to a particular party. Since this "feel close" variable may tap different aspects of electoral behavior than a vote variable, I further eliminated the remaining observations from Afrobarometer 2. The full results are provided in Table C of the supplementary materials. They show that the results from Table 3 are rather robust. In all nine models, the coefficient on the electoral law is negative. When GVF or PVP is the dependent variable, the coefficient on the electoral law is precisely estimated in all three models, and when PVF is the dependent variable, the coefficient on the electoral law is precisely estimated in one of the three models.

Results using "endogenous" groups. Fearon's (2003) definition of groups is attractive because it represents one of the few efforts to name politically relevant groups across much of the world. But as noted, any effort to name such groups always involves subjective judgments that can be criticized, making it useful to explore whether the results here are robust to alternative definitions of groups. To this end, one is constrained by the categories that exist in the surveys. But within this constraint, it is possible to allow the definition of politically relevant groups to "emerge endogenously" from the data. Each of the surveys typically has one or more "ethnic" group variables—some have an "ethnicity" variable, some a "race" variable, and some a "language" variable. At times, the validity of the group names under these variables is highly suspect for the purposes here. In one Australian survey, for example, the ethnicity groups include Australian, British Isles, and Southern Europe, which are hardly the key salient groups in that country. Similar examples of surveys with suspect groups include a Canadian survey (groups include English, French, Irish, Scottish, German/Austrian, and Italian) and a U.S. survey (American Indian, English, Irish, Scottish, French, German, Polish, Italian, White, Black, Hispanic, and American). Using the Fearon-defined groups is, therefore, more attractive than using the survey-defined groups in many instances.

But the fact that there often exist multiple variables tapping group identity in a given survey makes it possible to look empirically for group identities that are most "ethnicized." To this end, in each survey, in addition to the Fearon-based measures, I have calculated up to three additional ethnicization measures using the survey-defined ethnicity, language, and race categories. For a given survey, I have then set the measure of ethnicization to the maximum score obtained from the different possible group definitions. In Canada's 1997 CSES survey and the 1999 WVS survey, for example, one can calculate the four ethnicization variables using the Fearon groups, the survey ethnicity variables, or the survey language variables. For the CSES survey, GVF is .19 using the ethnicity variable, .14 using the Fearon groups, and .14 using the language variable. Thus, for the 1997 CSES survey I set GVF equal to .19. But in the WVS survey, the value of GVF for Fearon groups—.15—is higher than the value of GVF using other group definitions. Thus, for this survey I set GVF equal to .15.

Are the previous results regarding proportional representation robust using the endogenously defined groups? The full results are found in Table D of the online supplementary materials.[4] I estimated the models for the three dependent variables analyzed in Table C (available online) using the full data set, as well as the three subsets of data used in Table C. Thus, there are four models for each of the three measures of ethnicization. The results are actually stronger than was the case using just the Fearon groups. The coefficient is negative and rather precisely estimated in 11 of 12 models, with a p-value of less than .01 in 6 of the 12 models. Only when the most restrictive sample is used and PVF is the dependent variable is the negative coefficient on the electoral law variable estimated with considerable error.

Why Should Proportional Representation Be Associated with Lower Levels of Ethnic Voting?

It is widely believed that by making party formation low cost, PR should make ethnicity more salient because it allows parties to form by making ethnic appeals to even

[4] Allowing the definition of groups to be determined endogenously expands the number of countries and surveys because it is no longer necessary to eliminate those countries where Fearon's groups are inadequately represented.

small groups. But in none of the models estimated do we find a positive correlation between PR and ethnicization. And for the two ethnicization variables that increase with the number of groups, GVF and PVP, there is a robust and precisely estimated negative relationship between PR and ethnicization.

I cannot establish that PR causes lower ethnicization. It does, however, seem possible to at least diminish concerns about reverse causation. That is, we probably need not worry that these results are due to PR being chosen in cases where ethnicization is low. First, it is often argued that PR *is* used in precisely the type of situations that make ethnic relations most difficult, and there is a widespread belief that PR is most appropriate in ethnically divided societies. Rokkan (1970), for example, argues that PR was adopted in Europe's most ethnically and religiously divided countries precisely because by allowing the representation of minority groups, it fostered territorial consolidation. More recent research also argues that PR is most likely to be adopted in divided societies (e.g., Boix 1999; Lijphart 1992). To the extent that PR is chosen in places where ethnicity presents particularly difficult problems, this should bias *against* finding a negative relationship between ethnicization and PR. This makes the empirical results all the more striking. Second, despite the existing arguments about ethnic diversity and PR, there is essentially no relationship in the data used here between ethnic divisions and PR. The bivariate correlation between ELF and the proportion of legislators from multimember districts, for example, is a minuscule –.11 (i.e., more heterogeneous societies are more majoritarian), while the correlation between EP and Prop. MMD is only .14. I calculated two additional variables related to the structure of groups: (a) the size of the largest group, and (b) the number of small groups (with a size less than 10% of the sample). Again, the correlation between the electoral law and these measures is very small (and has the "wrong" sign).[5] The fact that the choice of electoral law is unrelated to the nature of underlying ethnic divisions is consistent with recent research emphasizing the wide variety of strategic considerations that drive the choice of electoral law (e.g., Andrews and Jackman 2005; Benoit 2004).

Given that I find the opposite relationship between electoral law and ethnic voting than that typically assumed, and that I cannot establish empirically a causal relationship, it is especially important to ask whether in fact there could be a reasonable explanation for this negative relationship between PR and ethnicization. I believe that there is. I argue that part of the answer lies in the incentives for politicized ethnicity in majoritarian systems, regardless of whether groups are geographically concentrated or geographically diverse. Another part of the answer is that easy party formation in PR systems makes it possible to divide group members against each other in electoral politics.

TABLE 4 Voting by Ethnic Group in Bangladesh and Nigeria

Bangladesh 1999 WVS			
Party	**Muslims**	**Hindus**	**Total**
Awami League	45	89	49
BNP	30	8	33
Jatiya	14	3	13
Jamat	6	1	6

Nigeria 2003 Afrobarometer				
Party	**Housa**	**Yoruba**	**Ibo**	**Total**
PDP	39	63	39	51
APP	60	9	16	35
AD	0	23	3	6
APGA	0	1	40	7

Note: Cells give the percent of the vote by the group for the party.

In majoritarian systems, where party entry is difficult in a given district, there are two different patterns that result in high ethnic voting. The first pattern is exemplified by Bangladesh and the United States, where there is a large majority group and a relatively small (but nontrivial) minority group that is geographically dispersed. In Bangladesh, a country that has struggled to establish free and fair elections, the Muslims (93% of respondents in the WVS 1999) are the dominant group and the Hindus (7%) are the small minority group. The top of Table 4 shows the support for the four main parties by these two groups. There is no ethnic party in the sense that the Awami League, a center-left party that supports the rights of minorities, receives a plurality of support from both groups. But the Hindus overwhelmingly support this party, whereas the Muslims show substantial levels of support for the other parties, particularly the Bangladesh Nationalist Party, a center-right party that is not friendly to Hindu rights. Given the division of Muslims between the center-left and center-right, the cohesive support of the Hindus for the Awami League can be pivotal in making it the plurality party. Indeed, in the most recent election, the Awami Party won a clear majority with Hindu

[5]The correlation between the log of district magnitude and the size of the largest group is .15, and the correlation between the log of district magnitude and the number of small groups is −.22. Virtually identical results for all correlations are obtained using proportion of MMD as the measure of electoral law.

support. The story in the United States is similar in that the white majority leans toward the Republicans, but heavily supports both parties. This puts the minority blacks in a potentially pivotal situation. The blacks overwhelmingly support the Democrats, and it is virtually impossible to imagine the Democrats winning national elections without this black support. More generally, in majoritarian systems with geographically dispersed groups, there can be strong incentives for a minority group to vote together for a mainstream party in an effort to become pivotal in determining election outcomes.

The second pattern of strong ethnic voting in majoritarian systems is exemplified by some of the African majoritarian countries with high ethnic diversity. These countries typically follow a pattern like that of Nigeria, given in the bottom of Table 4 (for the three main groups and the largest parties). The PDP receives support from each of the ethnic groups, but a plurality from each group supports a different party: 60% of Hausa support the APP, 63% of Yoruba support the PDP, and 40% of Ibo support the AGPA. Several factors combine to create strong incentives for ethnic voting during majoritarian elections in Nigeria. First, the groups are geographically concentrated, making ethnic appeals possible. Second, no group has a majority, making it more difficult for any group to exercise the strategy of the Hindus in Bangladesh, whereby they vote cohesively for a catch-all party in an effort to be pivotal to that party's success. Third, with majoritarianism, there is restricted entry of parties to challenge ethnic parties in any given district, making it easier for parties to achieve success by winning support from a dominant group in the district.

The dynamic is different in PR systems. The data show that individuals often demur when presented with the opportunity to vote for ethnic parties. Take the example of Catalans in Spain. There are two parties that are "ethnic" in the sense that one can predict reasonably well a person's ethnicity based on knowledge that he or she supports the party. Data from the 2004 CSES survey reveal that for the Republican Left of Catalonia (ERC, which advocates independence), 81% of supporters are Catalan, and for the Convergence and Union party (CiU, which does not advocate independence), 65% of supporters are Catalan. But only 50% of all Catalan voters support one of these two ethnic parties, as many support the PSOE or other parties. And those individuals who identify as Catalan are a relative small proportion of all Spanish—only 7%. So the Catalans are a small group, about half of them vote for nonethnic parties, and those who support ethnic parties split their vote between two such parties that have quite different positions on issues central to Catalans. Another group almost as large as the Catalans is the Galicians, who are even less inclined to vote cohesively. Only about 5% report supporting the Galician party (the Galician Nationalist Bloc, or BNG), with the rest splitting their vote between the PSOE and the PP. Thus, although PR makes it possible for ethnically oriented parties to form, it also makes it easy for nonethnic parties to compete for ethnic votes and for multiple ethnic parties to compete for the same ethnic group. This diminishes the intensity of ethnicity in vote choice.

TABLE 5 The Relationship between PR and the Ratio of Parties to Groups

	(1)	(2)	(3)	(4)
Prop. MMD	1.033*	1.048**	0.785*	0.934*
	(0.558)	(0.459)	(0.453)	(0.493)
ELF		−4.673***	−3.120*	−3.588***
		(0.718)	(1.591)	(1.093)
GDP/capita (ln)			0.393	0.447
			(0.335)	(0.335)
Polity2			−0.138	−0.126
			(0.122)	(0.131)
Federalism			−0.601	−0.600
			(0.531)	(0.533)
Geo. Isol.			0.210	
			(0.331)	
Constant	1.653***	4.017***	1.164	0.872
	(0.321)	(0.534)	(2.797)	(2.543)
Adj. R^2	0.050	0.390	0.362	0.388
N	69	69	67	69

Note: OLS models with standard errors clustered by country. The dependent variable is $\frac{\#parties}{\#groups}$. *p < .10, **p < .05, ***p < .01.

A different approach to illustrating how PR diminishes the cohesiveness of voting by groups is to examine the correlation between the electoral law and the number of parties per group. Table 5 shows several OLS regressions. The dependent variable is $\frac{\#parties}{\#groups}$. Model 1 presents the simple bivariate regression on Prop. MMD. The coefficient is positive and measured relatively precisely (p = .07). It is important, of course, to control for how fractionalized the society is, and Model 2 adds ELF as a control variable. The model fit improves dramatically and Prop. MMD remains large and is now very precisely estimated. Model 3 adds standard controls, and Model 4 includes the controls except Geographic Isolation (which is measured with considerable error). The coefficient on Prop. MMD remains positive, significant, and large. Using the coefficient from Model 4, the analysis suggests in a country with five groups, there would be roughly 4.5 more parties in a pure PR system than in a single-member district plurality system.

The argument here is not that the politicization of ethnicity will always be small in PR systems—the cases of Belgium and Macedonia make clear this is not the case. But perhaps the politicization of ethnicity occurs less, on average, in PR systems precisely because it is so easy to make electoral appeals of any sort. If one party attempts to exploit ethnic identity with strong ethnically based appeals under a permissive PR system, then this very electoral permissiveness also allows other parties to make appeals that attract voters on issues other than ethnicity. Given that members of an ethnic group typically have heterogeneous preferences, in PR systems they may often conclude that supporting ethnically oriented parties is not the most effective way of advancing their own interests because they should have attractive options among nonethnic parties. One implication of this argument would be that if the United States adopted PR, we should expect the salience of race in electoral politics to recede because multiple parties could form to appeal to the black vote, thereby dividing it. There may be, for example, a party that is socially conservative but supportive of affirmative action and a party that is socially liberal but supportive of affirmative action, with both parties receiving support from blacks.

Conclusion

Proportional representation is often advocated as the most attractive electoral law in divided societies, and recent cross-national empirical studies by Cohen (1997), Saideman et al. (2002), and Schneider and Wiesehomeier (2008) find that PR is indeed associated with lower levels of civil conflict. These empirical studies, like so much other research on ethnic divisions, invoke Lijphart's logic to explain why PR is attractive: by allowing parties to represent groups in parliament, groups will adopt peaceful as opposed to violent means for advancing their cause. The key assumption in this research, then, is that groups are in some sense exogenously given, that they have clear group-specific interests, and that the best electoral institutions are ones that allow direct representation of groups by parties.

The central empirical finding of this article suggests that the apparent success of PR electoral laws in managing ethnic conflict likely follows a quite different logic. The Lijphart-based argument leads us to expect a stronger relationship between voting behavior and group identity in PR-type systems, but the empirical analysis here finds the opposite: there is a weaker relationship between vote choice and group identity in PR systems than in plurality ones, regardless of whether one takes a group- or party-based perspective. The main reason seems to be that while ethnic identity can be an important element of vote choice, voters have other interests or identities that are equally or more important. By allowing relatively easy party formation, PR allows parties to form that appeal on bases other than ethnic identity, with the result being that voters from the same group often divide their support across a number of parties, often nonethnic ones. Proportional representation, then, may facilitate good governance not by giving each group its own party, but by diminishing the salience of ethnicity in elections. Ironically, this implies that if one accepts the Horowitz argument that the goal should be to depoliticize ethnicity in elections, one should adopt the electoral institutions advocated by Lijphart.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(June): 155–94.

Alesina, Alberto, and Edward Glaeser. 2004. *Fighting Poverty in the US and Europe: A World of Difference.* New York: Oxford University Press.

Alesina, A., and E. La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3): 762–800.

Andrews, Josephine, and Robert W. Jackman. 2005. "Strategic Fools: Electoral Rule Choice under Extreme Uncertainty." *Electoral Studies* 24(1): 65–84.

Baldwin, Kate, and John D. Huber. 2010. "Cultural vs. Economic Differences: Forms of Ethnic Diversity and Public Goods Provision." *American Political Science Review* 104(4): 644–62.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1): 165–76.

Benoit, Kenneth. 2004. "Models of Electoral System Change." *Electoral Studies* 23(3): 363–89.

Birnir, Jóhanna Kristín. 2007. *Ethnicity and Electoral Politics.* New York: Cambridge University Press.

Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93(3): 609–24.

Brancati, Dawn. 2008. "The Origins and Strength of Regional Parties." *British Journal of Political Science* 38(1): 135–59.

Brancati, Dawn. 2009. *Peace by Design: Managing Ethnic Conflict through Decentralization.* New York: Oxford University Press.

Cohen, Frank S. 1997. "Proportional versus Majoritarian Ethnic Conflict Management in Democracies." *Comparative Political Studies* 30: 607–30.

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

Easterly, W., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4): 1203–50.

Esteban, Joan-Maria, Laura Mayoral, and Debraj Ray. 2012. "Ethnicity and Conflict: An Empirical Study." *American Economic Review* 102(4): 1310–42.

Esteban, Joan-Maria, and Debraj Ray. 1994. "On the Measurement of Polarization." *Econometrica* 62: 819–51.

Fearon, James. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8(2): 195–222.

Gallagher, Michael. 1991. "Proportionality, Disproportionality and Electoral Systems." *Electoral Studies* 10: 33–51.

Greenberg, Joseph H. 1956. "The Measurement of Linguistic Diversity." *Language* 32: 109–15.

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

Horowitz, Donald. 1991. *A Democratic South Africa? Constitutional Engineering in a Divided Society.* Berkeley: University of California Press.

Johnson, Joel W., and Jessica S. Wallack. "New Electoral Systems Dataset: Electoral Systems and the Personal Vote." University of California, San Diego. Typescript.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization* 15(1): 22–79.

Lijphart, Arend. 1977. *Democracy in Plural Societies.* New Haven, CT: Yale University Press.

Lijphart, Arend. 1992. "Democratization and Constitutional Choices in Czecho-Slovakia, Hungary, and Poland, 1989–1991." *Journal of Theoretical Politics* 4(2): 207–23.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in 35 Countries.* New Haven, CT: Yale University Press.

Lijphart, Arend. 2002. "The Wave of Power-Sharing Democracy." In *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy*, ed. Andrew Reynolds. Oxford: Oxford University Press, 37–54.

Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67(2): 281–315.

Montalvo, Jose G., and Marta Reynal-Querol. 2005a. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review* 95: 796–815.

Montalvo, Jose G., and Marta Reynal-Querol. 2005b. "Ethnic Diversity and Economic Development." *Journal of Development Economics* 76: 293–323.

Reilly, Ben, and Andrew Reynolds. 1999. *Electoral Systems and Conflict in Divided Societies.* Washington, DC: National Academy Press.

Reynal-Querol, M. 2002. "Ethnicity, Political Systems, and Civil Wars." *Journal of Conflict Resolution* 46(1): 29–54.

Rokkan, Stein. 1970. *Citizens, Elections, Parties: Approaches to the Comparative Study of the Process of Development*. Oslo: Universitetsforlaget.

Saideman, Stephen M., David J. Lanoue, Michael Campenni, and Samuel Stanton. 2002. "Democratization, Political Institutions, and Ethnic Conflict: A Pooled Time-Series Analysis, 1985-1998." *Comparative Political Studies* 35(1): 103–29.

Schneider, Gerald, and Nina Wiesehomeier. 2008. "Rules That Matter: Political Institutions and the Diversity–Conflict Nexus." *Journal of Peace Research* 45: 183–203.

Sisk, Timothy D., and Andrew Reynolds. 1998. *Electoral Systems and Conflict Management in Africa.* Washington, DC: U.S. Institute of Peace Press.

Taagepera, Rein, and Bernard Grofman. 2003. "Mapping the Indices of Seat-Votes Disproportionality and Inter-Election Volatility." *Party Politics* 9(6): 659–77.

Treisman, Daniel. 2002. "Defining and Measuring Decentralization: A Global Perspective." University of California, Los Angeles. Typescript.

Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India.* Cambridge: Cambridge University Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Formula for Geographic Isolation, from Baldwin and Huber (2010). Note about inclusion of parties or groups. Data sources.

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

Table B: Countries and surveys in study.

Table C: OLS results using Fearon-defined groups and subsets of data.

Table D: OLS results using endogenously defined ethnic groups.