

# Measuring ethnic voting: The political context and the politicization of ethnicity\*

John D. Huber<sup>†</sup>

January 13, 2011

## Abstract

I develop two related measures of ethnic voting in a country. One measure taps the amount of information that is revealed about an individual's vote choice by knowing only the individual's ethnic group. The other measure taps the amount of information that is revealed about an individual's ethnic group by knowing only the individual's vote choice. Each measure also takes into consideration information about the size of groups. The paper uses existing survey data to estimate both measures and then examines the relationship between characteristics of political systems and the politicization of ethnicity. A central empirical finding is that the politicization of ethnicity is lowest in systems using proportional electoral laws – the opposite of what is typically assumed. The results also show that political decentralization is associated with lower levels of ethnic voting behavior, economic development is associated with higher levels of ethnic voting behavior, and the level of democratic development is uncorrelated with ethnic voting levels.

---

\*An earlier version of the paper 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” October, 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, Joanna Birnir, Dawn Brancati, Simon Hug and seminar participants at Columbia University. I am also grateful to the Russell Sage Foundation, where I was a visiting research scholar while writing an early draft of this paper.

<sup>†</sup>Professor, Department of Political Science, Columbia University

# 1 Introduction

Ethnic diversity has been associated with a variety of governance issues in developed and developing countries, including lower levels of public goods provision, higher corruption, political instability, violence and slower economic growth. Ethnic diversity does not, however, condemn a country to economic and political problems. Many countries seem to govern quite well even though their population is composed of multiple groups.

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 and decentralization. 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. Thus, ethnic groups should be highly salient to electoral behavior in PR systems.<sup>1</sup> The main disagreement concerns whether such politicization of ethnicity is a good thing. Some scholars argue for PR: by allowing each group 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., Lijphart 1977, Lijphart 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 and 1991).

There is less agreement about whether decentralization affects the politicization of ethnic groups in national politics (Brancati 2009). Scholars like Lijphart (1977) argue that by allowing groups autonomy over their own social, economic and political affairs, decentralization diffuses the salience of groups in national level politics. Others scholars, such as Kymlicka (1998) and Snyder

---

<sup>1</sup>See, for example, Reilly and Reynolds 1999, Sisk and Reynolds 1998, Tselis 1990 and Wilkinson 2004.

(2000), argue that decentralization will reinforce ethnic identity, inviting conflict between groups at the national level. This is particularly true in “ethnofederal” systems, where the boundaries of subnational units coincide with the geographic distribution of groups.

Each of these arguments about institutional design in ethnically divided societies rests on assumptions about whether particular institutional forms affect the politicization of ethnicity. The validity of these assumptions, however, has never been established because there exists no measure of politicization that can be applied across a range of political systems. Consequently, social scientists do not actually know whether proportional representation politicizes ethnicity, as is invariably assumed, or whether decentralization politicizes ethnicity or not.

This paper has two goals. The first is to develop two election-related measures that can be used to describe the politicization of ethnicity in democratic systems. Both measures are based on the premise that the politicization of ethnicity is high when voting behavior and group identity are closely related. The first measure, “Ethnic Voting,” taps how easy it is to predict an individual’s vote choice by knowing only the individual’s group. If the distribution of support for each party is the same across groups, for example, then group identity conveys no information about vote choice and Ethnic Voting is very low. But if the distribution of support for the various parties is quite different for each group (with each group supporting parties that receive little support from outside the group), considerable information about vote choice is conveyed by group identity and Ethnic Voting will be high.

“Ethnic Parties,” the second measure, is closely related to Ethnic Voting. It provides information about how easy it is to predict an individual’s group by knowing only the party the individual supported. While the two measures are closely related, they diverge when a group supports multiple parties that receive little or no support from outside the group. In such cases, knowing which party an individual supports will provide considerable information about the individual’s group (because parties receive support primarily from single groups), but knowing an individual’s group will convey less information about the individual’s voting behavior (because members of the same

group spread their support across multiple parties).

Ethnic Voting and Ethnic Parties are country-level measures. Although it is possible to think about the politicization of specific groups, a country-level measure is necessary to study empirically the relationship between system-level factors – like the electoral law or decentralization – and the politicization of ethnicity. In addition, focusing on a country-level measure makes it possible to take account of group size in measuring the politicization of ethnicity. We should expect, for example, that if everyone from a particular ethnic group supported the same party (and no one outside this group supported the party), then the degree to which this politicization of ethnicity is a problem in society should depend on the size of the group. As a small group becomes larger, the overall politicization of ethnicity in electoral politics should be said to increase.

The second goal is to implement the measures empirically across a wide range of political systems in order to understand the relationship between ethnic voting behavior and macro political factors like electoral laws and decentralization. To this end, the paper draws on Afrobarometer surveys, the World Values Survey (WVS) and the Comparative Study of Electoral Systems surveys (CSES) to measure Ethnic Voting and Ethnic Parties in 43 countries. The data show that at any given level of ethnic diversity, there is a wide range of Ethnic Voting and Ethnic Party scores.

The empirical measures make it possible to assess claims in the literature about political institutions and the politicization of ethnicity. With respect to debates about the electoral law, the results show that the common premise – that PR politicizes ethnicity – is not supported by the data. On the contrary, the data reveal a robust negative relationship between PR and both measures of ethnic behavior. With respect to the two opposing views about decentralization, the results support those who contend that decentralization *depoliticizes* ethnicity.

The paper is organized as follows. The next section describes the two measures, Ethnic Voting and Ethnic Parties. 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 relationship between the political context and ethnic voting follows. The paper 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.

## **2 Measuring Ethnic Voting and Ethnic Parties**

This section describes the two ways in which this paper conceptualizes the politicization of ethnicity during elections, Ethnic Voting (which describes how much information about vote choice is conveyed by knowing only an individual's group) and Ethnic Parties (which describes how much information about an individual's group is conveyed by knowing only the individual's vote choice).

Both measures are based on information about voting behavior and group identity. First consider Ethnic Voting and consider the example at the top of Table 1, which examines voting patterns by group in two countries, Hungary and Finland.<sup>2</sup> In Hungary, the two main groups, Hungarians and Roma, distribute their vote in roughly the same way. The Roma are somewhat less supportive than the Hungarians of the Socialists (.36 vs. .47), and the Hungarians are somewhat less supportive of Fidesz than are the Roma (.41 vs .54), but the differences are relatively small. In Finland, the two main groups are the Finns and Swedes, and the Finns distribute their vote across almost all of the parties, whereas 83 percent of Swedes support the Swedish People's Party (SPP). Thus, for each party (that is non-trivial in size), the difference between the proportion of Finns that support the party and the proportion of Swedes that support the party is quite large. Ethnic Voting uses information like that in Table 1 to create a measure that will reflect the fact that group identity provides more information about vote choice in Finland than in Sweden (because of the highly predictable behavior of the Swedes in Finland).

Ethnic Voting includes two types of information about the relationship between voting behavior and groups. The first is how well one can predict voting behavior by knowing an individual's group. The measure should obviously have a higher score in Finland than Hungary because there

---

<sup>2</sup>The data in the table are from the Comparative Study of Electoral Systems survey.

Table 1: Proportion of vote from each ethnic group for parties in Finland and Hungary

Hungary 2002						
	Center Party	Justice and Life	Socialists	Free Dems	Workers	Fidesz
Hungarians (96%)	.03	.03	.47	.05	.02	.41
Roma (3%)	.04	0	.36	.04	.04	.54

Finland 2003							
	SDP	Center	Nat. Coalit.	Left Alliance	Swedish PP	Green	CDP
Finns (93%)	.28	.23	.16	.10	.03	.09	.08
Swedes (7%)	.06	.04	0	.04	.83	.02	.02

Calculating EV in Finland

(1) Total Vote	.2646	0.2167	.1488	.0958	.0860	.0851	.0758
(2) $(V_{Finns}^j - V^j)^2$	.00024	.00018	.00013	.00002	.00314	.00002	.00002
(3) $(V_{Swedes}^j - V^j)^2$	.04186	.03122	.02214	.00311	.55354	.00424	.00311

$$(4) \sum_{j=1}^p (V_{Finns}^j - V^j)^2 = .0037 \quad EV_{Finns} = \sqrt{\frac{1}{2} \cdot .0037} = .04$$

$$(5) \sum_{j=1}^p (V_{Swedes}^j - V^j)^2 = .6592 \quad EV_{Swedes} = \sqrt{\frac{1}{2} \cdot .6592} = .57$$

$$(6) EV' = (.04 * .93) + (.57 * .07) = .08 \quad \text{Weight} = \frac{1}{\sqrt{\frac{2-1}{2*2}}} = 2$$

$$(7) EV = 2 * .08 = .16$$

is a tighter link between group and vote in Finland. In general, as voting behavior by individuals from different groups becomes differentiated, Ethnic Voting increases. The second is group size. Suppose there were two groups in a country, and members of each group supported their own party, with the party for each group receiving no support from outside the group. In this case, independent of the size of the groups, one would know with certainty an individual's party by knowing the individual's group. Yet if ethnic voting behavior is of concern because it affects some aspect of governance, then we should expect the effect of this sort of "perfect ethnic voting" to depend on the size of the groups. If one of the two groups is 99 percent of the population, for instance, the potential for destabilizing behavior from perfect ethnic voting would seem small compared with a situation where the two groups were much more equal in size. Or, to take the example in the table, one would expect that the effect of the Swede's ethnic voting behavior on Finnish politics would be different if the Swede's constituted 40 percent of the voting population. Thus, holding the distribution of support for parties constant, the measure of ethnic voting proposed here increases as groups become more equal in size.

To make possible meaningful comparisons across countries, the measure has a theoretical minimum (0) and maximum (1) that is independent of the number of ethnic groups. The theoretical minimum is reached when the support for each party is the same across all groups. Hungary approaches this, and ethnic voting there should be 0 if the Roma and Hungarians had exactly the same proportions of respondents supporting each party. This implies that the measure should decrease if the distribution of voting support by a group becomes more similar to the distribution of voting support in the population as a whole. Ethnic voting in Finland would therefore decrease if there were marginal shifts by Swedes away from the SPP and toward any other party (because the proportion of Swedes who support any other party is less than the proportion of the total population that supports those parties), or if there were marginal shifts by the Finns toward the SPP.

The theoretical maximum of 1 should be reached when two conditions are satisfied: (a) each group votes for its own party (and no other group supports that party) and (b) each group is the

same size. Condition (a) ensures that simply knowing an individual's group will reveal the party that the individual supports. It also implies that the ethnic voting measure will take higher values if a particular group supports one party (that receives no support outside the group) than if that group splits its support between two or more parties (that receive no support from outside the group). Condition (b) captures the importance of group size as summarized in the previous paragraph.

It is possible to create a measure of ethnic voting satisfying these three properties by drawing on indices of proportionality that have been used to study electoral systems. These indices measure the relationships between seats and votes in a party system, and perfect proportionality is achieved when the percent of seats that each party receives is the same as the percent of votes that the party receives. A number of proportionality indices exist but the Gallagher index of disproportionality (Gallagher 1991) has become the most widely accepted in political science.<sup>3</sup> The Gallagher index of disproportionality is given by:

$$D = \sqrt{\frac{1}{2} \sum_{i=1}^n (V_i - S_i)^2}, \quad (1)$$

where  $V_i$  is the vote share for party  $i$  and  $S_i$  is the seat share for party  $i$ . The index 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,  $D$  would equal 1.

The disproportionality index can be used to describe how vote support for each party by a particular group compares to vote support for each party in the population as a whole. If the proportion of a group that supports each party is very similar to the proportion of the entire population that supports each group, then disproportionality would be low. The disproportionality scores for each group can then be aggregated to create a country-level measure of ethnic voting.

---

<sup>3</sup>See Taagepera and Grofman (2003) for a useful analysis of why it makes sense to embrace this index.

For the group-level component of the aggregate measure, let Ethnic Voting by group  $g$  be given as

$$EV_g = \sqrt{\frac{1}{2} \sum_{j=1}^P (V_g^j - V^j)^2}, \quad (2)$$

where  $V_g^j$  is the proportion of individuals in group  $g$  who support party  $j$ ,  $V^j$  is the proportion of individuals in society that support  $j$ , and  $P$  is the number of parties.  $EV_g$  therefore is a variant of  $D$ , but instead of calculating differences between votes and seats for all parties, it calculates, for a given group and for each party, the difference between in-group support for the party and support for that party in the entire population ( $V_g^j - V^j$ ).

To create a measure of ethnic voting in a country, one possibility would be to sum each  $EV_g$ , weighting by the size of each group. Let  $EV' = \sum_{g=1}^G (EV_g * s_g)$ , where  $G$  is the total number of groups in the country and  $s_g$  is the proportion of group  $g$  in the country's voting population.  $EV'$  is not, however, an adequate measure of ethnic voting because its theoretical maximum can never equal 1, and because it is sensitive to the number of groups. The hypothetical maximum should occur when each group is the same size and each group supports a different party than other groups. With two groups,  $EV'$  has a maximum of .5 (where there are two equal sized groups, and each group supports one party that is not supported by the other group). If there are three equal-sized groups, each supporting different parties,  $EV' = \frac{1}{3}$ .

One can address this issue by weighting  $EV'$  by a function of the number of groups. To solve for this function, note that if all groups are of equal size and each group supports a different party than other groups, then for each group,

$$\begin{aligned} EV_g &= \sqrt{\frac{1}{2} \left[ \left(1 - \frac{1}{G}\right)^2 + \sum_{j=1}^{G-1} \left(0 - \frac{1}{G}\right)^2 \right]} \\ &= \sqrt{\frac{G-1}{2G}}. \end{aligned}$$

Thus, in this case of equal-sized groups,  $(EV_g * s_g) = \left(\frac{1}{G} * \sqrt{\frac{G-1}{2G}}\right)$ , and  $EV' = \sum_{g=1}^G (EV_g * s_g) = \sqrt{\frac{G-1}{2G}}$ . Consequently, if we weight  $EV'$  by the reciprocal of  $\sqrt{\frac{G-1}{2G}}$ , the measure of ethnic voting will equal 1 when ethnic voting is at its maximum, independent of the number of groups:

$$EV = \frac{1}{\sqrt{\frac{G-1}{2G}}} \sum_{g=1}^G (EV_g * s_g). \quad (3)$$

By construction,  $EV$  takes the value 1 when groups are of equal size and each group supports its own party. Since  $EV$  is a function of  $(V_g^j - V^j)$ , it will equal 0 for any number and size of groups if the proportion of each group that supports each party is the same. Basing  $EV$  on a proportionality score also ensures that as vote proportions become more equal across groups,  $EV$  will decline. And since each  $EV_g$  is based on the comparison of vote proportions by the group with vote proportions *in the population*, as the size of groups becomes more equal,  $EV$  will increase.

The bottom of Table 1 uses Finland to walk through an example of how  $EV$  is computed using real data. The first step is to calculate the (squared) difference between the vote in the group for each party and the vote in the population for each party. Line (1) provides the proportion of the total vote received by each party, and line (2) depicts the squared difference between the proportion of Finns supporting a party and the population proportion supporting this party (e.g., the entry in line 2 under SDP is  $(.28 - .2646)^2 = .00024$ ). Line (3) depicts the same squared differences for the Swedes. To calculate  $EV_{Finns}$ , first sum the entries in line 2, which yields .0037 (see line (4)). Dividing this number by 2 and taking the square root yields  $EV_{Finns} = .04$  (also line (4)). Line (5) depicts the same calculation for the Swedes. Given the concentrated support by the Swedes for the SPP, we see that  $EV_{Finns}$  is much smaller than  $EV_{Swedes}$ . To calculate  $EV$ , we must calculate  $EV'$  (which is .08, as shown on line (6)), and the weight (which is 2 when the number of groups is 2, also show on line (6)).  $EV$  equals  $EV'$  times the weight, or .16.

Figure 1 provides an example of how  $EV$  changes with changes in the size of groups, and with

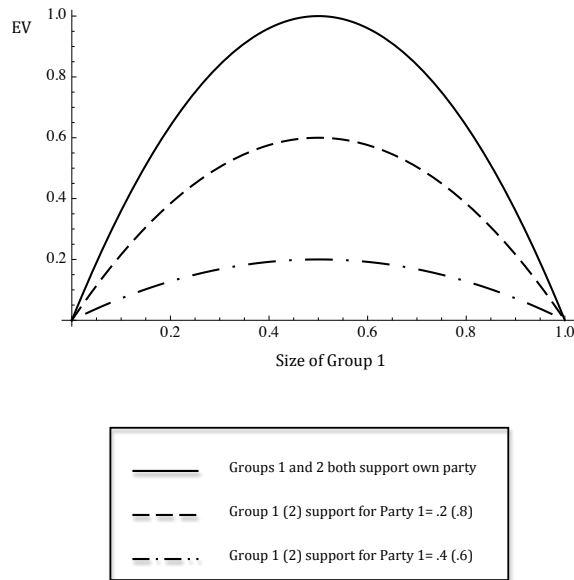


Figure 1: EV examples

changes in their voting patterns. There are two groups and two parties. The solid line depicts the case when each group supports a different party, the dashed line when 20 percent of Group 1 supports Party 1 (and thus 80 percent support Party 2), and 80 percent of Group 2 supports Party 1. The dash-dot line depicts the case where the voting patterns for the two groups are most similar to each other: support for Party 1 is 40 percent by Group 1 and 60 percent by Group 2. The figure illustrates that no matter the distribution of support for the parties by each group, EV goes to zero as one group becomes very large (and of course equals 0 when there is only one group), and EV is at its maximum of 1 when the groups are of equal size and each group supports its own party. And for any size of the groups, EV is increasing as the levels of support by each group for each party diverge.

*Ethnic parties.* Ethnic Voting decreases as it becomes more difficult to predict an individual's vote by knowing only the individual's group. Consequently, the measure will be lower when members of an ethnic group support multiple parties (none of which receive support from outside the

group) than when members of the group support one party (that receives no support from outside the group). One argument justifying a measure with this property is that when group members do not agree on which party should represent the group's interests, then 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. From this perspective, the impact of the Swedes on Finnish politics would diminish if Swedes split their vote between two (Swedish) parties.

There may be situations, however, when a group supports multiple parties, but these parties act cohesively on issues that are salient to the group. Members of the group, for instance, may differ on general policies regarding business, labor, or public goods, but they may have common goals on constitutional issues pertaining to regional autonomy, or on policy issues surrounding education. In this case, if a group divides its support across multiple group-linked parties, this may not diminish the impact of the group if the various parties have a common voice on group-specific issues. It is therefore worth exploring a measure of the politicization of ethnicity that does not decrease when members of the same group divide their vote across multiple group-linked parties.

To this end, Ethnic Parties is a measure tapping the information that is conveyed by an individual's party choice about the individual's group identity. The measure is constructed in the same way as Ethnic Voting. For each party,  $k$ , the measure calculates

$$EP_k = \sqrt{\frac{1}{2} \sum_{g=1}^G (V_k^g - s_g)^2}, \quad (4)$$

where  $V_k^g$  is the proportion of party  $k$ 's support that comes from group  $g$  (and, recall,  $s_g$  is the proportion of group  $g$  in society and  $G$  is the total number of groups). Thus, if the proportion of party  $k$ 's support from each group is the same as the proportion of each group in society,  $EP_k$  will equal 0. And  $EP_k$  increases as the proportion of voters from a particular group for party  $k$  diverges from the proportion of voters in society.

Following the same aggregation and weighting strategy used for  $EV$ , we can sum each  $EP_k$ , weighting by the percent of votes received by  $k$ , and dividing the sum by a factor of the number of parties (to ensure that the maximum score is invariant to the number of parties). Let

$$EP = \frac{1}{\sqrt{\frac{P-1}{2P}}} \sum_{k=1}^P (EP_k * p_k), \quad (5)$$

where  $P$  is the total number of parties and  $p_k$  is the proportion of the vote received by party  $k$ .

$EP$  and  $EV$  are very closely related. If each group supports only one party (that receives no support outside the group) and the groups are the same size, then  $EP = EV$ . Similarly, if the distribution of support across the parties is the same for each group, the two measures will equal 0.

But although the two measures are related, 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 of Figure 2. 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 Ethnic Party score would be greater than the Ethnic Voting score. Contrast this with the example depicted at the bottom 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 Ethnic Party score is less than the Ethnic Voting Score.

*Related measures.* I am aware of no existing measure of ethnic voting that can be used for cross-national research, but there do exist indirect cross-national measures and country-specific measures. Birnir (2007) uses electoral volatility (changes in support for parties over time) as a proxy for ethnic voting in her study of ethnic voting and political stability, but she does not

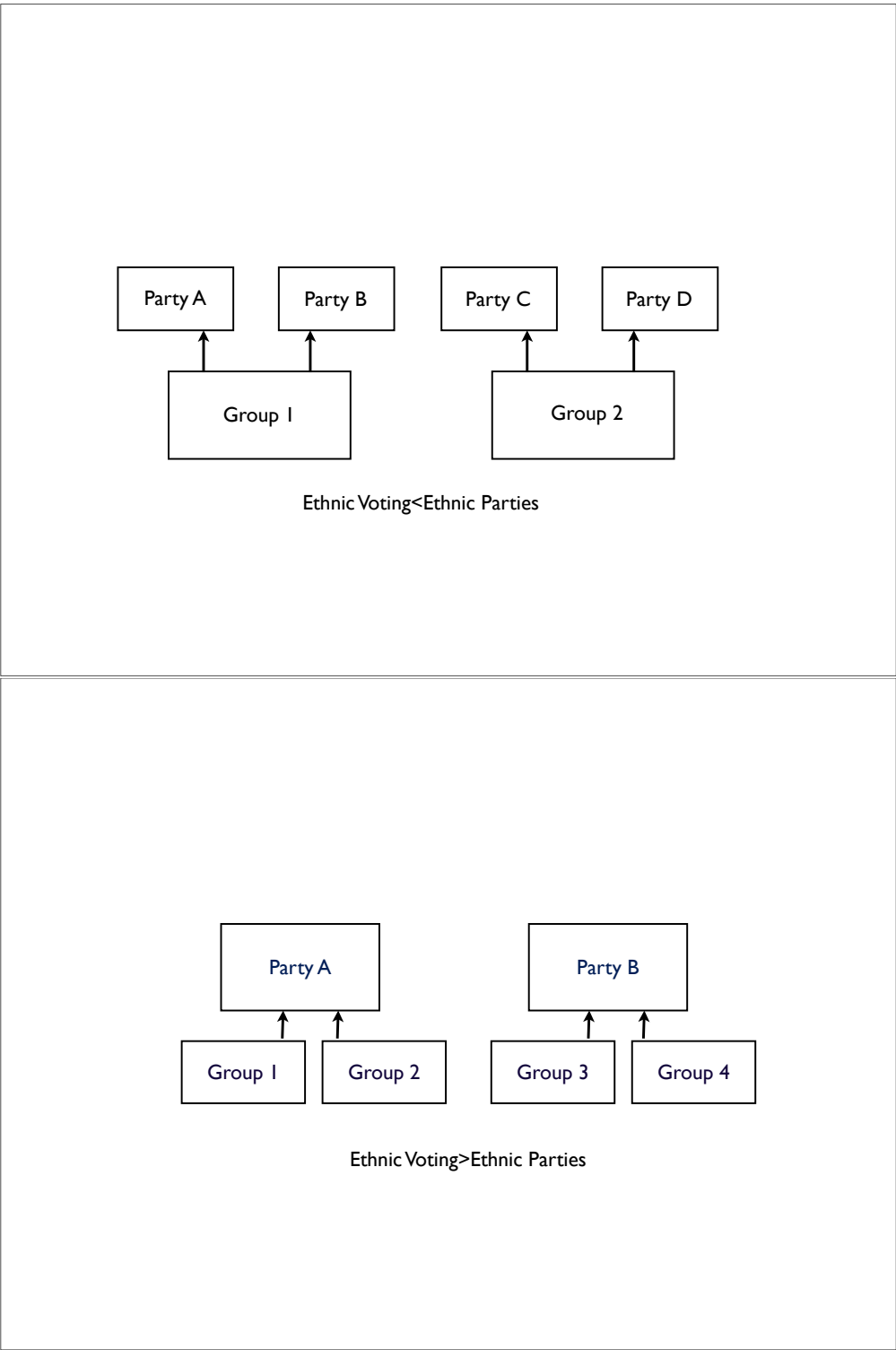


Figure 2: Ethnic Voting versus Ethnic Parties

establish that such volatility is related to ethnic-based behavior. Brancati (2008 and 2009) also offers a measure that could be relevant to cross-national studies of ethnic voting. She focuses on the level of support in a country for regional parties, which are often ethnically based. But Brancati is not specifically interested in ethnic voting, and she emphasizes that not all regional parties are ethnic (e.g., the Basque National Party in Spain) and that not all ethnic parties are regional (e.g., the BJP in India).

It is important to recognize, however, that there are other ways to conceptualize ethnic voting besides the focus here on patterns of support across groups for political parties. Chandra's (2004) analysis of India also focuses on political parties, but focuses on "ethnic parties" – parties that make an explicit appeal to specific ethnic groups and that advocate excluding other groups (see also Chandra 2005). The ethnic voting measure proposed here does not take account of the nature of appeals, and could thus take high values even without concerted efforts by party elites mobilize an ethnic vote. The implications of party systems where voters sort themselves by ethnic group even when elites do not make explicit ethnic appeals might be different than the implications of party systems where explicit ethnic appeals motivate group sorting.

Posner (2005, see also 2004a), in his study of which ethnic divisions become politically salient in Zambia, does not focus on the role of parties. Instead, he considers the extent to which voters support *candidates* from their own tribe (in single-party elections) or their own language group (in multiparty ones). His study reminds us that the politicization of ethnicity can be high even in one-party states if this politicization operates through candidate choice rather than party choice.

Although the measures used by Posner and Chandra would be difficult to calculate in large-n cross-national research, their studies raise conceptual issues that one should bear in mind when drawing conclusions about the measures proposed here. There may be cases, for example, where Ethnic Voting or Ethnic Parties understate ethnic voting behavior because such behavior operates through candidates rather than parties. But in competitive democracies, political parties are typically the main vehicle through which ethnic representation occurs, and the politicization of

ethnicity can be approximated using voting patterns – without the difficult and at times subjective process of classifying the extent to which parties make explicit ethnic appeals, and without coding parties themselves as “ethnic” or not. Much can therefore be gained by utilizing a measure of ethnic voting that is based on voting support for parties across groups, and that can be used across a range of democratic systems.

### **3 Ethnic Voting and Ethnic Parties in 43 countries**

Implementing the two measures requires information about individual ethnicity and party preference, which can be found in three existing cross-national surveys – the World Values Survey (“WVS”, wave 4, 1996 and 1999), the Afrobarometer (rounds 2 and 3, ) and the Comparative Study of Electoral Systems (“CSES”, all available waves). These surveys contain questions that make it possible to categorize respondents according to their ethnic group, as well as their vote intention. And including surveys from all three studies ensures that the data set includes democratic systems that vary in their ethnic diversity, political and economic development, and political institutions.

Considerable research now accepts that ethnicity is not primordial, but rather is “constructed” – it is endogenous to the social context, can change, and is often multidimensional (e.g., Horowitz 1985, Laitin 1998, Chandra 2004 and Posner 2004b). Group identities are very salient politically in some contexts, while these same group identities may not be salient in other contexts (Posner 2005). The challenge therefore is to identify individual group attachments in a way that makes sense in the particular context in which the individual finds him or herself. This inevitably entails subjective judgment, but without a willingness to engage in such subjectivity, it is impossible to conceive of a broadly cross-national empirical study of ethnic voting.

To identify the relevant ethnic categories in the countries for which surveys exist, this study follows the ethnic categories identified in Fearon (2003). Fearon takes seriously the constructivist

critique of primordialism, and the problems it creates for creating a clear list of groups in a country. But he also underlines that much research requires such a list, and argues that any list of groups in a country should be based on the “the idea that members and non-members recognize the distinction [on which group identity is based] and anticipate that significant actions are or could be conditioned on it” (p. 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” (to which individuals are born into), and whether groups are locally viewed as socially or politically consequential.<sup>4</sup> Groups are therefore based on 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 information about groups because of the care he takes in using reasonable criteria consistently across a wide range of countries. The measures proposed here do not, of course, depend on any specific definition of groups, and it would be possible to substitute another list in place of Fearon’s to calculate the two measures.

Fearon’s list of groups is of little use for our purposes if the surveys do not make it possible to identify the relevant groups. The rule followed here 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 percent. 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 percent 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 percent of the population. This rule assures that surveys are included in the study only if the group data from

---

<sup>4</sup>The seven prototypical features are: “1. Membership in the group is reckoned primarily by descent by both members and non-members. 2. Members are conscious of group membership and view it as normatively and psychologically important to them. 3. Members share some distinguishing cultural features, such as common language, religion, and customs. 4. These cultural features are held to be valuable by a large majority of members of the group. 5. The group has a homeland, or at least ‘remembers’ one. 6. The group has a shared and collectively represented history as a group. Further, this history is not wholly manufactured, but has some basis in fact. 7. The group is potentially “stand alone” in a conceptual sense that is, it is not a caste or caste-like group (e.g., European nobility or commoners)” (p. 201).

Table 2: Countries and surveys in study

Western Europe	Other Europe	Asia	Latin America	Africa	Other
Belgium (WVS)	Bulgaria (CSES)	Bangladesh (WVS)	Brazil (CSES, WVS)	Benin (Afrobar.)	Australia (CSES (2), WVS)
Finland (CSES)	Czech Republic (CSES)	India (WVS)	Colombia (WVS)	Botswana (Afrobar. (2))	Canada (CSES, WVS)
France (CSES)	Estonia (WVS)		Dominican Republic (WVS)	Kenya (Afrobar. (2))	New Zealand (CSES (2), WVS)
Germany (WVS)	Georgia (WVS)		Mexico (CSES (3), WVS)	Madagascar (Afrobar.)	USA (CSES (2), WVS)
Netherlands (WVS)	Hungary (CSES)		Uruguay (WVS)	Malawi (Afrobar. (2))	
Spain (CSES (3), WVS)	Latvia (WVS)		Venezuela (WVS)	Mali (Afrobar. (2))	
	Lithuania (CSES)			Mozambique (Afrobar. (2))	
	Macedonia (WVS)			Namibia (Afrobar. (2))	
	Moldova (WVS)			Nigeria (Afrobar.)	
	Romania (CSES)			Senegal (Afrobar. (2))	
	Russia (CSES (2))			South Africa (Afrobar. (2))	
	Slovenia (CSES)			Zambia (Afrobar. (2))	
	Ukraine (CSES)				

them can be used to reasonably approximate the groups identified by Fearon.

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 percent” selection rule for keeping surveys based on Fearon groups, results in the use of 67 surveys from 43 countries. The countries and surveys are depicted in Table 2.<sup>5</sup> Importantly, one can be confident that the surveys adequately represent the 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. The surveys therefore provide useful data for examining ethnic behavior across countries.

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 what party the respondent would

<sup>5</sup>Several surveys are also omitted from the analysis because the relevant right-hand side variables do not exist for these surveys in the regression analysis below.

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 that support each party, as was done in Table 1 for Sweden and Hungary.

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. But since it is possible to explore the incidence of systematic bias by controlling for survey type in the empirical analysis, 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.<sup>6</sup>

The measures of Ethnic Voting from these surveys ranges from .01 to .42 with a mean of .14 and a standard deviation of .11. The level of ethnic voting in any country will obviously be affected by the level of ethnic diversity, so to compare the levels of ethnic voting across countries, it is important to control for the overall level of ethnic fractionalization. The top panel in Figure 3 plots Ethnic Voting against ELF, along with the bivariate regression line derived from regressing Ethnic Voting on ELF. At just about any level of ethnic fractionalization, there is variation in Ethnic Voting. Among the low ELF countries, for example, Finland has high Ethnic Voting relative to other countries. Among countries with a middling level of ethnic fractionalization, Macedonia and Belgium have very high levels of Ethnic Voting and Mexico and Brazil have relatively low levels. It is therefore clear that Ethnic Voting cannot be explained exclusively by underlying levels of ethnic diversity.

Table 3 describes the countries that have the highest and lowest levels of Ethnic Voting, controlling for the level of ethnic diversity (measured by ELF). The list of countries is based on the

---

<sup>6</sup>The analysis eliminates respondents who supported parties receiving less than 2 percent of the vote, as such small parties do not have enough respondents to accurately characterize the vote distribution by group.

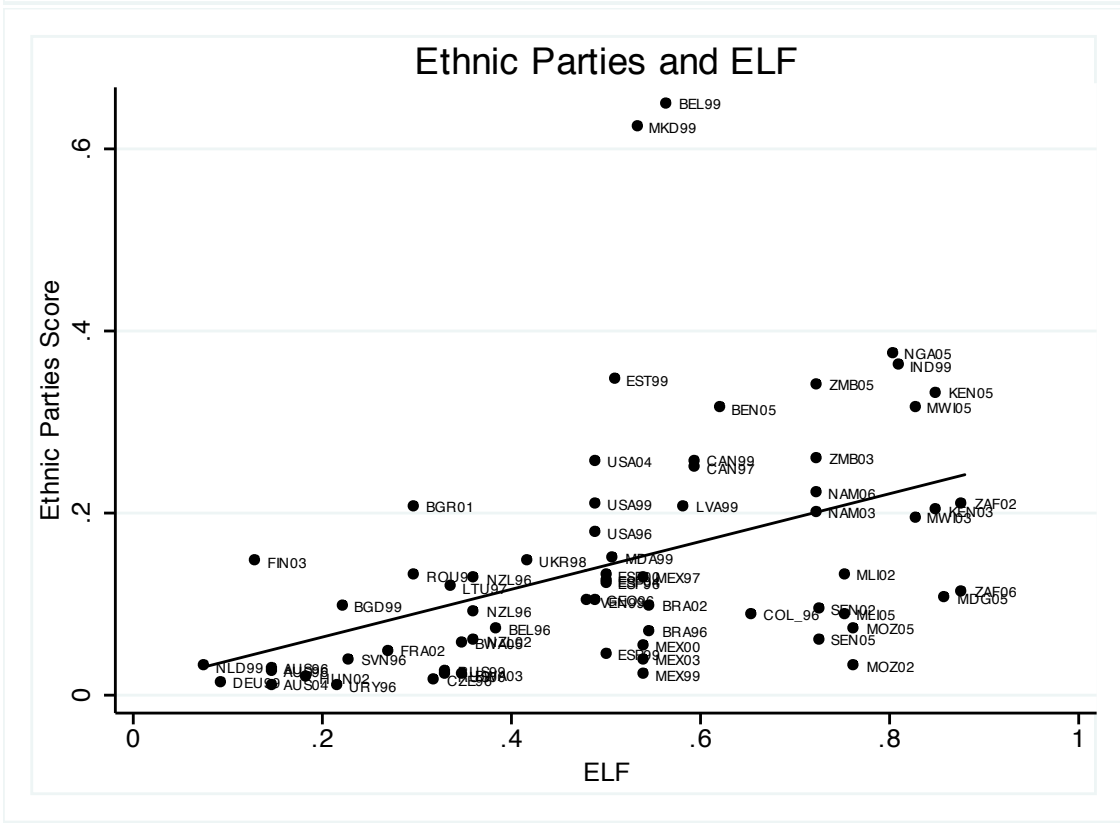
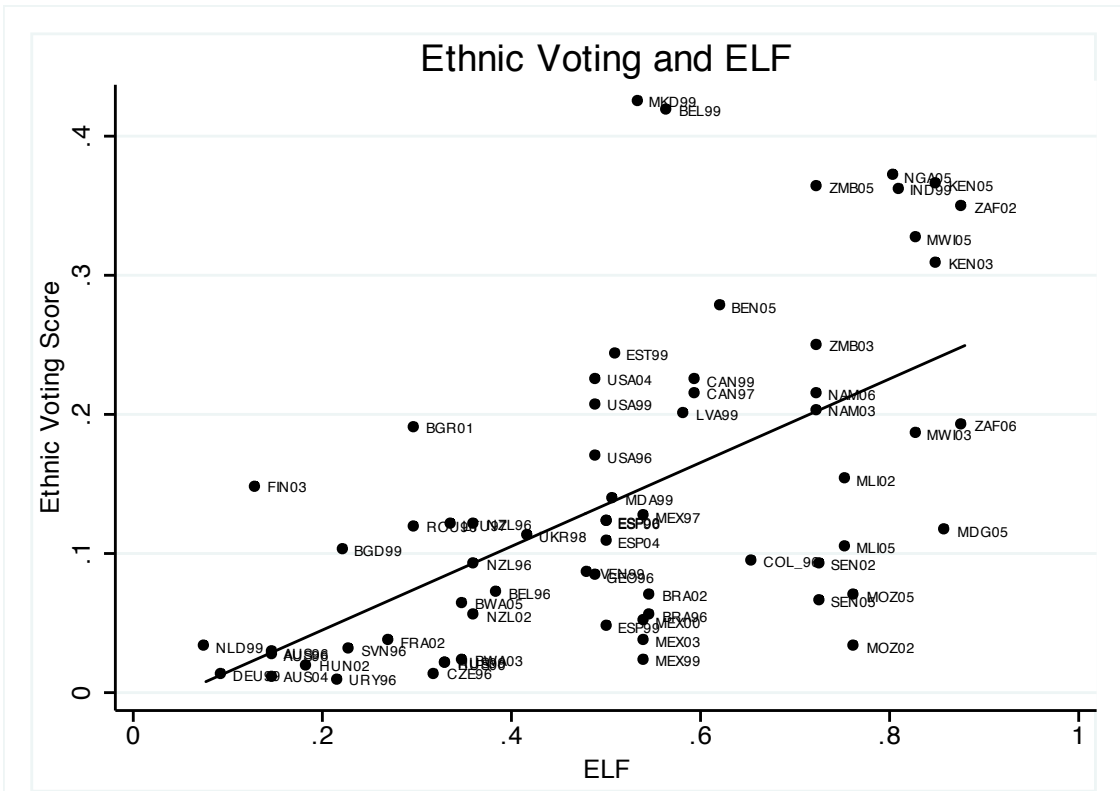


Figure 3: Ethnic Voting, Ethnic Parties and ELF

bivariate regression of  $EV$  on  $ELF$ . The countries in the top of the table are those with the lowest level of ethnic voting given  $ELF$  – that is, they are the countries with the largest negative residuals from the regression. In Mozambique, for example,  $ELF$  is very large, but  $EV$  is very small, making this the country with the lowest level of ethnic voting (given  $ELF$ ). There are two types of vote distributions that will lead to low  $EV$  in high  $ELF$  countries. In some countries, Ethnic Voting is low because a dominant party receives a very large proportion of support across groups. This is what happened in 5 of the 6 African surveys listed in the top of the table. Among the African countries, only Mali stands out as a country that has low ethnic voting without a dominant party.

Mali is an example of the second type of vote distribution that leads to low ethnic voting in ethnically diverse countries. In Mali, the largest party was named by only 36 percent of respondents, and three other parties were named by at least 10 percent of respondents. Not only is there no dominant party, there also is little correlation between group and the vote. Dunning and Harrison (2010) examine the reason for low ethnic voting in Mali, attributing it to an informal institution called *cousinage*, which creates cross-cutting cleavages. Mali is joined by Mexico and Brazil as countries that have much lower than expected ethnic voting, but no dominant party.

The bottom half of the table lists the surveys that have the highest levels of Ethnic Voting (controlling for  $ELF$ ). The list includes countries with a wide range of  $ELF$ , and a wide range of “largest parties.” Belgium and Macedonia have roughly the same  $ELF$  as Mexico, but their Ethnic Voting scores are roughly 10 times larger than Mexico. The list is geographically diverse, including four countries from Africa, two from eastern Europe, two from western Europe, and India.

The bottom panel of Figure 3 displays the relationship between Ethnic Parties and  $ELF$ . We again find that for any level of  $ELF$ , there is typically a wide range of Ethnic Parties scores. Ethnic Parties has a wider range than Ethnic Voting, from .01 to .65, with a mean of .15 and a standard deviation of .13. Macedonia and Belgium are the countries that increase the most (about .2) when moving from the Ethnic Voting measure to Ethnic Parties. When the surveys from these countries are excluded, the two measures have the same mean (.13) and the same standard deviation (.10).

Table 3: Largest outliers in regression of EV on ELF

<i>Surveys with largest negative residuals (lower EV than predicted)</i>					
Country (survey)	ELF	Largest party	EV	Predicted EV	Residual
Mozambique (Afrobarom. 2002 )	0.77	0.95	0.03	0.22	-0.18
Mozambique (Afrobarom. 2005 )	0.77	0.91	0.07	0.22	-0.15
Senegal (Afrobarom. 2005 )	0.73	0.76	0.06	0.20	-0.14
Madagascar (Afrobarom. 2005 )	0.86	0.78	0.12	0.24	-0.13
Mexico (WVS 1999 )	0.54	0.53	0.02	0.15	-0.13
Senegal (Afrobarom. 2002 )	0.73	0.68	0.09	0.20	-0.11
Mexico (CSES 2003 )	0.54	0.35	0.04	0.15	-0.11
Mali (Afrobarom. 2005 )	0.75	0.38	0.10	0.21	-0.11
Mexico (CSES 2000 )	0.54	0.48	0.05	0.15	-0.10
Brazil (WVS 1996 )	0.55	0.34	0.05	0.15	-0.10
<i>Surveys with largest positive residuals (higher EV than predicted)</i>					
Macedonia (WVS 1999 )	0.53	0.39	0.42	0.15	0.28
Belgium (WVS 1999 )	0.57	0.19	0.42	0.16	0.26
Zambia (Afrobarom. 2005 )	0.73	0.40	0.36	0.20	0.16
Nigeria (Afrobarom. 2005 )	0.80	0.50	0.37	0.23	0.14
India (WVS 1999 )	0.81	0.36	0.36	0.23	0.13
Finland (CSES 2003 )	0.13	0.27	0.15	0.02	0.12
Kenya (Afrobarom. 2005 )	0.85	0.53	0.36	0.24	0.12
Bulgaria (CSES 2001 )	0.30	0.56	0.19	0.07	0.11
Benin (Afrobarom. 2005 )	0.62	0.35	0.28	0.17	0.10
Estonia (WVS 1999 )	0.51	0.27	0.24	0.14	0.10

Table 4: Largest outliers in regression of EP on ELF

<i>Surveys with largest negative residuals (lower EP than predicted)</i>					
Country (survey)	ELF	Largest party	EP	Predicted EP	Residual
Mozambique (Afrobarom. 2002 )	0.77	0.95	0.03	0.21	-0.18
Senegal (Afrobarom. 2005 )	0.73	0.76	0.06	0.20	-0.14
Mozambique (Afrobarom. 2005 )	0.77	0.91	0.07	0.21	-0.14
Mexico (WVS 1999 )	0.54	0.53	0.02	0.15	-0.13
Madagascar (Afrobarom. 2005 )	0.86	0.78	0.11	0.24	-0.13
South Africa (Afrobarom. 2006 )	0.88	0.89	0.11	0.24	-0.13
Mali (Afrobarom. 2005 )	0.75	0.38	0.09	0.21	-0.12
Mexico (CSES 2003 )	0.54	0.35	0.04	0.15	-0.12
Senegal (Afrobarom. 2002 )	0.73	0.68	0.09	0.20	-0.11
Mexico (CSES 2000 )	0.54	0.48	0.05	0.15	-0.10
<i>Surveys with largest positive residuals (higher EP than predicted)</i>					
Belgium (WVS 1999 )	0.57	0.19	0.65	0.16	0.49
Macedonia (WVS 1999 )	0.53	0.39	0.62	0.15	0.47
Estonia (WVS 1999 )	0.51	0.27	0.34	0.15	0.20
Nigeria (Afrobarom. 2005 )	0.80	0.50	0.37	0.22	0.15
Benin (Afrobarom. 2005 )	0.62	0.35	0.31	0.17	0.14
Zambia (Afrobarom. 2005 )	0.73	0.40	0.34	0.20	0.14
India (WVS 1999 )	0.81	0.36	0.36	0.22	0.14
Bulgaria (CSES 2001 )	0.30	0.56	0.21	0.09	0.12
United States (CSES 2004 )	0.49	0.53	0.26	0.14	0.12
Finland (CSES 2003 )	0.13	0.27	0.14	0.05	0.10

South Africa is the country that moves most sharply in the other direction, with an Ethnic Parties score that is .14 points less than its Ethnic Voting score (using the 2002 Afrobarometer).

Table 4 shows the largest outliers for Ethnic Parties. For the most part, the countries are the same as those found in Table 3. South Africa is a large negative outlier for Ethnic Parties but not Ethnic Voting, and the US is a large positive outlier for Ethnic Parties but not Ethnic Voting. But there is an impressive correlation in the names of the countries on the lists in the two tables.

Figure 4 plots the relationship between Ethnic Voting and Ethnic Parties, along with the 45-

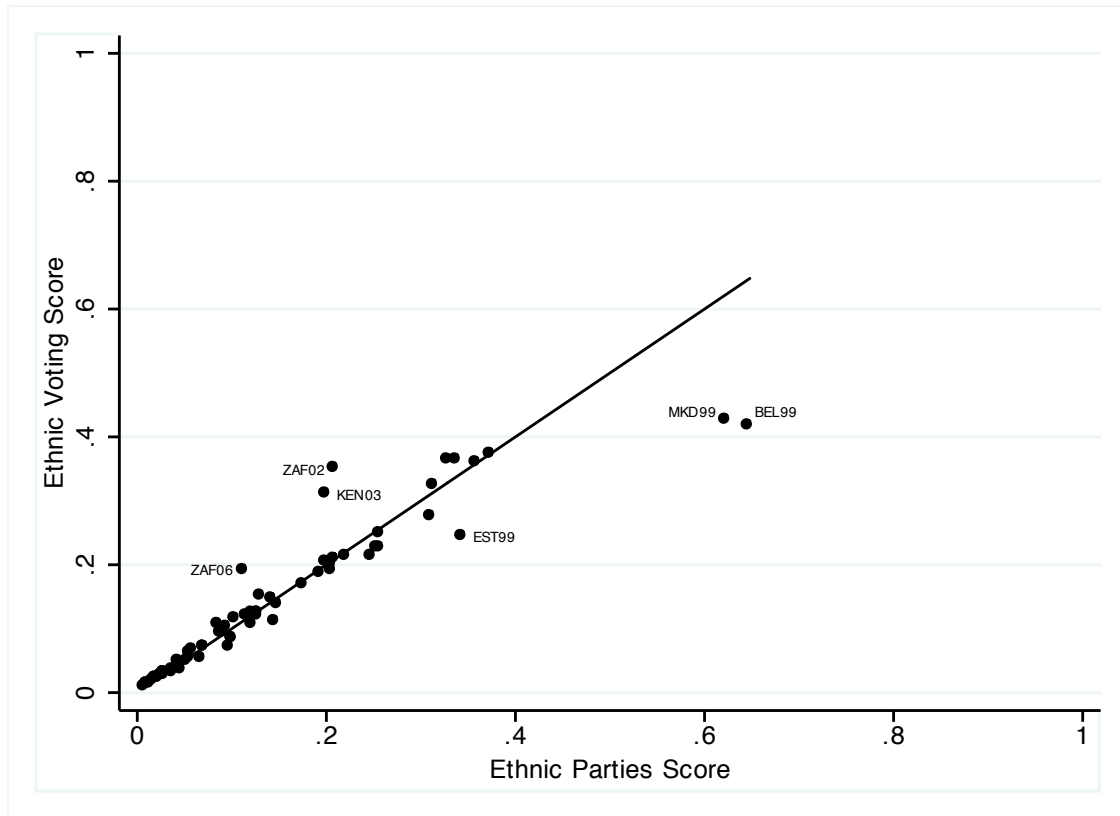


Figure 4: Ethnic Voting vs. Ethnic Parties

degree line. The two variables are obviously extremely closely related, with a correlation of .94. As noted previously, Belgium and Macedonia stand out as parties with much higher Ethnic Parties than Ethnic Voting scores. South Africa and Kenya have higher Ethnic Voting scores, but the difference between Ethnic Voting and Ethnic Parties is much lower than is the case for Belgium and Macedonia.

To understand why countries can have a much higher Ethnic Parties score than Ethnic Voting score, consider Table 5, which depicts results from the 1999 WVS survey in Macedonia. There are two major groups, the Macedonians and Albanians, and 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. The same dynamic unfolds in Belgium. Due to the substantial tensions between the two main groups – French and Flemish speakers – Belgian’s party system has a unique history. Starting in 1968, the main political parties began to divide along linguistic lines. In 1968, the Catholic (now Christian Democratic) Party split into two, a Flemish party (CVP) and a French one (PSC). In the years following, the other main parties followed suit, and now all significant parties are tied to one or the other of the linguistic communities. There are Flemish and French Liberals, Greens and Social Democrats, and there are Flemish-only parties that advocate separatism. The pairs of parties, such as the French and Flemish Social Democrats, cooperate and formally coordinate on many issues, but they maintain separate entities because of the salience of the regional and linguistic cleavage, and because of the geographic segregation of these groups from each other. Thus, as in Macedonia, one can be almost certain of which group a voter comes from by doing which party the voter supported, but one knows less about party choice given knowledge of an individual’s group because each group supports multiple parties.

The opposite situation exists in South Africa. There are a range of groups that throw almost all of their electoral support to the ANC. Thus, for individuals from many groups, group identity provides considerable information about group choice. But given that many groups support the ANC, knowing that an individual supported the ANC provides rather noisy information about the individual’s likely group. Thus, Ethnic Voting is higher than Ethnic Parties in South Africa.

Two points are worth underscoring from this examination of the two measures. First, although Ethnic Voting and Ethnic Parties are conceptually distinct, empirically, for almost all countries, the two measures are very similar. That is, the conceptual distinction between the two is greater than the empirical one. Second, using both measures, there clearly exists substantial cross national differences, even after controlling for ELF. While this should come as no surprise, these differences have not been previously quantified, leaving us without clear empirical knowledge of which countries have the highest level of ethnic voting. With this data in hand, it is possible to examine

Table 5: Voting by ethnic group in Macedonia

Party	Macedonia 1999		Total
	Macedonian (66.7%)	Albanian (29.7%)	
DA	3.2	0.0	2.2
DPA	0.0	40.3	12.0
LDP	7.30	0.0	4.9
LP	4.1	0.7	2.9
New Democratia	2.4	0.0	1.6
NDP	0.0	20.0	5.9
PDP	0.0	39.1	11.6
SDSM	54.3	0.0	36.2
Socialistic Party	3.8	0.0	2.5
VMRO - DPMNE	22.2	0.0	14.8
VMRO - VMRO	2.7	0.0	1.8

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

the relationship between macro political variables and the politicization of ethnicity.

## 4 Ethnic Voting, Ethnic Parties and the macro context.

In this section, I regress Ethnic Voting and Ethnic Parties on variables measuring the electoral law, decentralization, and other elements of the macro context.

As noted in the Introduction, it is widely assumed that PR politicizes ethnicity. To test this assumption, I consider three measures of the electoral law. The first is average district magnitude (“DM”), which 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 Data Base of Political Institutions (Beck, Clarke, Groff, Keefer, and Walsh 2001).<sup>7</sup> In cases where there were disagreements between these two sources, I conducted further research

<sup>7</sup>For Johnson and Wallack, see <http://dss.ucsd.edu/~jwjohnso/espv.htm>. For the World Bank Data, see <http://go.worldbank.org/2EAGGLRZ40>.

to determine the mean district magnitude. I take the log of average district magnitude because beyond a certain threshold, the variable has no further effect on the number of parties (Ordeshook and Shvetsova 1994, Cox 1997). If PR systems encourage the politicization of ethnic politics, Average District Magnitude should have a positive coefficient.

A country's average district magnitude can mask proportional attributes of electoral laws. Countries like Russia have an average district magnitude of 2, but half of the legislators are elected from one large PR district (with the rest being elected in single member districts). Since one large PR district can allow small parties that make ethnic appeals to form, it is useful to explore a measure of the electoral law that taps the total proportion of all legislators who are elected by PR. Johnson and Wallack (2007) provide such a measure, which I label Proportion MMD. 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 that took intermediate values using the measure average district magnitude or proportion of legislators from multi-member districts are now simply coded as PR systems. There are 44 observations in the data with PR systems and 23 (from 14 countries) with majoritarian electoral systems.<sup>8</sup>

The Introduction also described disagreements that exist in the literature regarding whether decentralization politicizes ethnicity. To explore the relationship between decentralization and ethnic behavior, I use a federalism indicator variable from Treisman (2002), supplemented by my own research for missing cases. If decentralization makes ethnicity more salient politically, the Federalism variable should have a positive coefficient.

The statistical models also include a number of other variables. Ethnic voting will obviously be relatively limited when ethnic diversity is low, so the models all include a control for ELF, as measured by Fearon (2003). In addition, it is widely held that the politicization of ethnicity should be particularly likely in early stages of democratic development. Lijphart (2002, 38) states, for

---

<sup>8</sup>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 (2), Tanzania (1), USA (3) and Zambia (2).

example, that “the problem of ethnic and other deep divisions is greater in countries that are not yet democratic or not fully democratic than in the well-established democracies. On [this point], I cannot think of any expert who disagrees...” Birnir (2007) argues that in early stages of the democratization process, ethnicity - which typically survives during periods of authoritarian rule – provides a particularly reliable cue to voters about which party will best serve their interests. We should therefore expect ethnic voting and ethnic parties to be most prevalent in societies that are least democratic. The measure of political development is the Polity 2 score. If democratic development diminishes the salience of ethnicity in politics, the Polity variable should have a negative coefficient.

Given the different wording used for the “vote” questions across the surveys, the models include indicator variables for three of the four surveys (with the CSES survey as the omitted category). 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).

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, p. 288). For a given group,  $g$ , define

$$I_g = \sum_{i=1}^n \left( \frac{p_g^i}{P_g} \cdot \frac{p_g^i}{T_i} \right), \quad (6)$$

where  $i$  is a region,  $n$  is the total number of regions,  $p_g^i$  is the population of group  $g$  in region  $i$ ,  $P_g$  is the total population of group  $g$  in the country, and  $T_i$  is the total population in region  $i$ .

Holding all else constant,  $I_g$  will increase as a group comes more concentrated in a region or as a group becomes larger (because this makes it more likely that individuals interact primarily with individuals from their own group). When all members of a group live in the same region (or regions) and this region has no members from other groups, the variable will take its maximal value of 1. The minimum (0) occurs when each individual in a region is his or her own group (and thus each group has only one member).

We can weight each  $I_g$  by group size to develop an isolation score for the country:

$$I_k = \sum_{g=1}^G \left( I_g \cdot \frac{P_g}{T} \right), \quad (7)$$

where  $k$  is a country,  $G$  is the total number of groups in country  $k$  and  $T$  is the population of country  $k$ . To calculate  $I_k$ , I use the region variable that exists in each of the surveys. Isolation should have a positive coefficient if ethnic voting increases when groups are more geographically isolated from each other.

*Results.* Table 6 presents the results when Ethnic Voting is the dependent variable. The first three models include one of the three measures of electoral laws, and all of the other variables described above. The fourth model interacts ELF with the electoral law. The regressions estimate robust standard errors and the continuous right-hand side variables are standardized to have a mean of 0 and a standard deviation of 1 to facilitate easy comparisons of the magnitudes of the estimated coefficients. The parentheses provide the p-values for the estimated coefficients.

The control variables are grouped at the bottom of the table. Looking across the four models, ELF is of course positive and very precisely estimated. Among the survey indicator variables, the WVS surveys are associated with slightly higher levels of ethnic voting, but the precision of these estimates vary with model specification. The results also show a relatively robust relationship between GDP and Ethnic Voting, but contrary to the expectation that the politicization of ethnicity is most likely to occur in less developed countries, the results show that ethnic voting levels are

Table 6: Variation in Ethnic Voting

DV: Ethnic Voting	(1)	(2)	(3)	(4)
Avg.DM (ln)	-0.020* (0.071)			
Prop. MMD		-0.026** (0.024)		
PR			-0.062** (0.019)	-0.055** (0.028)
ELF*SMD				0.135*** (0.000)
ELF*PR				0.093*** (0.000)
Polity2	-0.015 (0.450)	-0.010 (0.583)	-0.016 (0.391)	-0.018 (0.310)
Federalism	-0.050* (0.100)	-0.064** (0.045)	-0.061* (0.063)	-0.050 (0.110)
Geo. Isol.	0.026 (0.115)	0.028* (0.084)	0.027* (0.096)	0.029* (0.060)
(ln)GDP	0.048* (0.051)	0.045* (0.058)	0.049** (0.043)	0.055** (0.020)
ELF	0.112*** (0.000)	0.117*** (0.000)	0.115*** (0.000)	
WVS	0.047* (0.079)	0.044 (0.101)	0.038 (0.152)	0.044* (0.098)
Afrobarometer 2	-0.024 (0.601)	-0.035 (0.457)	-0.052 (0.316)	-0.028 (0.571)
Afrobarometer 3	0.004 (0.923)	-0.003 (0.937)	-0.022 (0.618)	-0.003 (0.937)
Constant	0.146*** (0.000)	0.154*** (0.000)	0.202*** (0.000)	0.181*** (0.000)
R-squared	0.379	0.401	0.411	0.434
N	67	67	67	67

Note: OLS models with p-values based on robust standard errors in parentheses. All continuous variables are standardized to have a mean of 0 and standard deviation of 1. \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$

higher in richer countries. Finally, as expected, there is a positive relationship between Ethnic Voting and the geographic isolation of groups. In countries where individuals are least likely to come into contact with members of other groups, ethnic voting scores are highest.

Next consider the electoral law. For each of the three measures of electoral law – average district magnitude (model 1), proportion of legislators from multimember districts (model 2), and the proportional representation indicator variable (model 3) – the coefficients are negative and rather precisely estimated. The effects are also relatively large, as can be seen by considering the PR indicator in model 3. The coefficient for PR is  $-.062$ , which means that the absolute value of the effect is about that of a one-standard deviation change in ELF itself, and it is larger in absolute value than the effect of a one standard deviation change in GDP (logged) or residential segregation.

The final model in Table 6 explores how the relationship between ethnic diversity and ethnic voting is mediated by the electoral law. If PR facilitates ethnic voting more than majoritarian electoral laws, then the impact of ethnic diversity on ethnic voting should be stronger in PR systems than in majoritarian ones. To test this possibility, rather than including ELF, model 4 includes ELF interacted with majoritarian electoral laws and ELF interacted with PR. The indicator for PR remains negative and significant,<sup>9</sup> and the effect of ELF on ethnic voting is about 50 percent larger in majoritarian systems.

Finally consider political development and decentralization. In each of the four models, the coefficients on Polity 2, though negative as expected, are all estimated with substantial error. We thus find no meaningful relationship between political development and ethnic voting. The federalism indicator variable, by contrast, is precisely estimated, and it has a negative sign, indicating that ethnic voting is highest in the more centralized countries. This result supports those who contend that decentralization can depoliticize ethnicity.

Table 7 presents the same models but using Ethnic Parties as the dependent variable. As noted

---

<sup>9</sup>A Wald test indicates that the probability that  $SMD * ELF$  and  $PR * ELF$  have the same coefficients is only .05.

previously, the measures are closely related empirically. It is therefore not surprising that the results are very similar to those in Table 6. In particular, the results show that Ethnic Parties has lower values in systems with PR, and that the effect of ELF on ethnic parties is larger in SMD than PR systems.<sup>10</sup> The results also show a negative relationship between federalism and ethnic parties. And the coefficient for Polity 2 is negative and is somewhat more precisely estimated compared with the Polity2 coefficients in Table 6 (though is still not statistically significant using standard thresholds).

## **5 Why does PR have lower levels of ethnic voting?**

Scholars have been divided regarding the effects of decentralization on ethnic politics, and the results in the previous section suggest that federalism depoliticizes ethnicity. The results also suggest that the general belief that ethnicity is most salient to politics in less well-established democracies is very weakly supported by the data. Perhaps the most surprising results are those concerning proportional representation. It is widely believed that by making party formation low cost, PR should make ethnicity more salient because elites can appeal to even small groups based on ethnicity. But regardless of whether one measures Ethnic Voting or Ethnic Parties, PR is associated with less ethnic behavior, and the impact of ethnic diversity on ethnic behavior is weaker in PR systems than in majoritarian ones. Why might the common belief be wrong? We can gain some insight by looking more closely at voting patterns in some majoritarian and PR systems.

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 US, where there is a large majority group and a relatively small (but non-trivial) minority group that is geographically dispersed. In Bangladesh, a country that has struggled to establish free and

---

<sup>10</sup>A Wald test indicates that the probability that the coefficients for the two variables interacting ELF with the electoral law variables are the same is .08.

Table 7: Variation in Ethnic Parties

DV: Ethnic Parties	(5)	(6)	(7)	(8)
Avg.DM (ln)	-0.023** (0.045)			
Prop. MMD		-0.023* (0.074)		
PR			-0.064** (0.025)	-0.057** (0.040)
ELF*SMD				0.157*** (0.000)
ELF*PR				0.118*** (0.001)
Polity2	-0.038 (0.132)	-0.032 (0.185)	-0.038 (0.112)	-0.040* (0.091)
Federalism	-0.082** (0.040)	-0.092** (0.023)	-0.093** (0.028)	-0.083* (0.053)
Geo. Isol.	0.045* (0.069)	0.048* (0.055)	0.047* (0.062)	0.048** (0.048)
(ln)GDP	0.077** (0.017)	0.072** (0.024)	0.076** (0.018)	0.083*** (0.008)
ELF	0.136*** (0.000)	0.139*** (0.000)	0.138*** (0.000)	
WVS	0.071* (0.071)	0.066* (0.091)	0.061 (0.127)	0.066* (0.100)
Afrobarometer 2	-0.070 (0.108)	-0.079* (0.074)	-0.099** (0.047)	-0.076 (0.133)
Afrobarometer 3	-0.027 (0.567)	-0.031 (0.515)	-0.053 (0.313)	-0.035 (0.497)
Constant	0.169*** (0.000)	0.176*** (0.000)	0.227*** (0.000)	0.208*** (0.000)
R-squared	0.327	0.327	0.347	0.358
N	67	67	67	67

Note: OLS models with p-values based on robust standard errors in parentheses. All continuous variables are standardized to have a mean of 0 and standard deviation of 1. \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$

fair elections, the Muslims (93 percent of respondents in the WVS 1999) are the dominant group and the Hindus (7 percent) are the small minority group. The top of Table 8 provides 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 support. The story in the US 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 (in the 1996 CSES survey, 84 percent of blacks said they voted Democratic), and it is virtually impossible to imagine the Democrats winning national elections without this black support.

In both the US and Bangladesh, there are no ethnic parties, but there is a minority ethnic group that “votes ethnically,” and that can swing elections through their support for a center-left party that supports the minority group on a number of policy issues important to the group. More generally, in majoritarian systems, there can be strong incentives for a minority group to vote together, even if this group is geographically dispersed and even if there is no competitive ethnic party. This is because issues other than ethnicity can divide the majority group, providing an opening for a cohesive minority group to influence 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 8 (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 percent of Hausa support the APP, 63 percent of Yoruba support the

Table 8: 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

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

PDP, and 40 percent 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, a country where ethnicity is held to play an important role in politics. There are two parties that are “ethnic” in the sense that one can predict reasonably well a person’s ethnicity based on knowledge that they support the party. For the Republican Left of Catalonia (ERC, which advocates inde-

pendence), 81 percent of supporters are Catalan, and for the Convergence and Union party (CiU, which does not advocate independence), 65 percent of supporters are Catalan. But only 50 percent 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 seven percent.<sup>11</sup> So the Catalans are a small group, about half of them vote for non-ethnic parties, and those that 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 Catalans is the Galacians, who are even less inclined to vote cohesively. Only about 5 percent report supporting the Galacian 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 non-ethnic parties to compete for ethnic votes, and for multiple ethnic parties to compete for the same ethnic party.

The politicization of ethnicity can obviously be large in PR systems, as the cases of Belgium and Macedonia make clear. But perhaps the electoral 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 heterogenous preferences, in PR systems they should often conclude that supporting ethnically oriented parties is not the most effective way of advancing their own interest because they should have attractive options among non-ethnic parties. In Spain, for example, the Catalans can support either a moderate ethnic party as opposed to a secessionist one, or they can support the PSOE, which has traditionally defended regional autonomy. Such choice diffuses the cohesiveness of group voting behavior, depoliticizing ethnicity when compared with, say, the blacks in the US, the

---

<sup>11</sup>The data in this example come from the 2004 CSES survey. The survey has few respondents who are Basque, making discussion of this group impossible.

Hindus in Bangladesh, or the Yoruba in Nigeria.

## 6 Conclusion

This paper develops measures that make it possible to test beliefs about the relationship between the macro political context and the politicization of ethnicity. Ethnic Voting considers how easy it is to predict voting behavior based on knowing only an individual's ethnicity. Ethnic Parties considers how easy it is to predict an individual's group based only on knowledge of the party the individual supported. Though the two measures are closely related, Ethnic Parties will take a higher value in situations where groups support a variety of parties that receive little or no support from outside the group. The analysis reveals that the Ethnic Voting and Ethnic Parties scores are very similar empirically, and that at any level of ethnic diversity there is considerable variation in both measures.

The empirical analysis shows that decentralization is associated with less rather than more ethnic voting, supporting those who contend that depoliticization can stabilize ethnically divided societies. The empirical analysis also shows that levels of political development are not correlated with ethnic voting levels, contradicting the widely held belief that ethnic differences fade in importance when societies develop mature democratic institutions. And perhaps most interestingly, there is a strong and robust effect of the electoral system, but the negative relationship between PR and ethnic voting is the opposite of what is typically assumed in the literature. There are reasons this result for PR should be unsurprising. Majoritarian systems provide clear incentives and opportunities for ethnic voting to occur, even by small groups and even when groups are not geographically concentrated. In fact, the risk of majoritarianism is not that it alienates groups by excluding them, but rather that it encourages the politicization of ethnicity. By contrast, PR systems create ample opportunities for parties to divide the vote of members from the same group, depoliticizing ethnicity.

These results suggest an important avenue for future research. Since many common beliefs about how ethnic diversity becomes politicized are unsupported by the data, it is not clear how ethnic diversity might affect governance outcomes. Does the effect of ethnic diversity on outcomes depend on the politicization of ethnicity? One could explore this question by examining the relationship between ethnic voting and governance outcomes. Does ethnic voting influence things like the survival of democracy, inter-group conflict, or the quality of public goods provision? And does the impact of ethnic voting on governance depend on the institutional context from which ethnic voting occurs?

## 7 References

- Baldwin, Kate and John D. Huber. 2010. "Economic versus cultural differences: Forms of ethnic diversity and public goods provision." Typescript, Columbia University, forthcoming in *American Political Science Review*.
- 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): 165-176.
- Birnie, Jóhanna Kristín. 2007. *Ethnicity and Electoral Politics*. New York: Cambridge University Press.
- Brancati, Dawn. 2008. "The Origins and Strength of Regional Parties." *British Journal of Political Science* 38(1): 135-159.
- Brancati, Dawn. 2009. *Peace by Design: Managing Ethnic Conflict through Decentralization*. New York: Oxford University Press.
- Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge: Cambridge University Press.
- Chandra, Kanchan. 2005. Ethnic Parties and Democratic Stability. *Perspectives on Politics* 3(2): 235-252.
- Cox, Gary M. 1997. *Making votes count: strategic coordination in the world's electoral systems*. New York: Cambridge University Press.
- Dunning, Thad and Lauren Harrison. 2010. "Cross-cutting cleavages and ethnic voting: An experimental study of cousinage in Mali." Forthcoming, *American Political Science Review*.
- 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: 335-1.
- Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley, California: University of California Press.
- Horowitz, Donald. 1991. *A Democratic South Africa? Constitutional Engineering in a Divided Society*. Berkeley, CA: University of California Press.
- Johnson, Joel W. and Jessica S. Wallack. "New Electoral Systems Dataset: Electoral Systems and the Personal Vote." Typescript, UCSD.

- Kymlicka, W. 1998. "Is Federalism a Viable Alternative to Secessionism?" In *Theories of Secession*, pp. 111-50, Percy B. Lehnig (ed.). New York: Routledge Press.
- Laitin, David D. *Identity in formation: The Russian-speaking populations in the new abroad*. Ithaca, NY: Cornell University Press.
- Lijphart, Arend. 1977. *Democracy in Plural Societies*. New Haven: Yale University Press.
- Lijphart, Arend. 1999. *Patterns of Democracy: Government forms and performance in 35 countries*. New Haven: Yale University Press.
- Lijphart, Arend. 2002. "The wave of power-sharing democracy." In Andrew Reynolds (ed.), *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy*. Oxford: Oxford University Press.
- Massey, Douglas S. and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67(2): 281-315.
- Ordeshook, Peter C. and Olga V. Shvetsova. 1994. "Ethnic Heterogeneity, District Magnitude and the Number of Parties." *American Journal of Political Science* 38(1):100-23.
- Posner, Daniel N. 2004a. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi." *American Political Science Review* 98(4): 529-545.
- Posner, Daniel N. 2004b. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48 (4): 849-863.
- Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press.
- Reilly, Ben and Andrew Reynolds. 1999. *Electoral Systems and Conflict in Divided Societies*. Washington, DC: National Academy Press.
- Sisk, Timothy D. and Andrew Reynolds. 1998. *Electoral Systems and Conflict Management in Africa*. Washington, DC: US 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." Typescript: UCLA.
- Tsebelis, George. 1990. "Elite Interaction and Coalition Building in Consociational Democracies." *Journal of Theoretical Politics* 2: 5-29.
- Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge: Cambridge University Press.