

# The Measurement of Cross-cutting Cleavages and Other Multidimensional Cleavage Structures

Joel Sawat Selway

Brigham Young University, 745 Kimball Tower, PO Box 25545, Provo, UT 84602

e-mail: joel\_selway@byu.edu

Many studies in economics and political science include the concept of ethnic diversity as a key independent variable in empirical studies. To date, however, only single-dimensional measures of ethnic diversity, such as ethnic fractionalization, have been available. In this paper, I define and measure three multidimensional characteristics of social structure—cross-cuttingness, cross-fractionalization, and subgroup fractionalization—and present a new cross-national data set comprised of indices along combinations of five cleavages: race, language, religion, region, and income. After addressing important definitional and measurement issues, I discuss the data and show how their inclusion in economic growth regressions reopens the theoretical debate regarding the influence of ethnic diversity, indicating the potential for these new indices to shed light on the role of ethnic diversity (and social structure more generally) with regards to a number of phenomena of concern to political scientists. The data set is publicly available from the Inter-University Consortium for Political and Social Research data archives and the Institute for Quantitative Social Science Dataverse Network.

## 1 Introduction

Over the past decade, an explosion of literature using measures for ethnic and religious diversity to explain a variety of phenomena such as economic growth (Easterly and Levine 1997; Montalvo and Reynal-Querol 2003, 2005a), civil war (Collier and Hoeffler 2000, 2004; Fearon and Laitin 2003; Montalvo and Reynal-Querol 2005b), the size of government (Annett 2001), public goods provision (La Porta et al. 1999; Kuijs 2000; Keefer 2005), and the number of parties in a political system (Ordeshook and Shvetsova 1994; Clark and Golder 2006) has occurred. By far the most common characteristic of social structure that appears in these studies is a measure of ethnic *fractionalization* first introduced by Taylor and Hudson (1972).

Fractionalization indices have come under heavy criticism for a number of reasons, ranging from accuracy and reliability to concept validity.<sup>1</sup> This paper confronts an important issue regarding fractionalization's ability to measure the concept of social diversity: unidimensionality. Unidimensionality gives primacy to a single social cleavage,<sup>2</sup> usually ethnicity, the choice of which often remains unjustified in many models.<sup>3</sup> In addition, scholars often fail to fully justify the use of the characteristic of fractionalization, which captures the number and relative size of groups.<sup>4</sup>

*Author's note:* I would like to thank Rob Franzese, Allen Hicken, Ken Kollman, Ashutosh Varshney, Bill Clark, numerous participants in the University of Michigan Annual Poster Session, 2006, three anonymous reviewers, and the editors of *Political Analysis* for their invaluable feedback and discussions. The author takes full responsibility for all final comments and work. Supplementary materials for this article are available on the *Political Analysis* Web site.

<sup>1</sup>For example, see the excellent critique by Laitin and Posner (2001).

<sup>2</sup>I use the terms *cleavage* and *dimension* interchangeably; *cleavage* refers to a "division on the basis of some criteria of individuals, groups, or organizations [between] whom conflict may arise. These criteria can be ascriptive, such as race, caste, ethnicity, language, or attitudinal, that is, ideology, preference, class, or religion" (Lane and Ersson 1994). Deegan-Krause (2007) provides a more strict definition of a cleavage, adding that structurally different groups also need shared attitudes and an organizational element to constitute a *full* cleavage. Although recognizing the value of this richer distinction, I maintain use of the term cleavage because of familiarity with the term *cross-cutting cleavages*, a focus of this article.

<sup>3</sup>For example, see Stoll's (2008) discussion on the lack of justification in using ethnic fractionalization indices in studies on the number of parties.

<sup>4</sup>One exception is Montalvo and Reynal-Querol (2003, 2005a, 2005b) who challenge the use of fractionalization, which captures the effective number of groups, in models of economic growth and civil war; instead, they offer a measure of *bipolarization*, which captures the degree there are two equally sized groups in society.

A prominent example illustrates how inadequate measurement has left the comparative literature in some way less developed, both theoretically and empirically. Easterly and Levine (1997) argue that “cross-country differences in ethnic diversity explain a substantial part of the cross-country differences in public policies, political instability, and other economic factors associated with long-run growth.” The measure they focus on, however, is the fractionalization of ethnolinguistic diversity. This conceptualization of ethnic diversity leads Easterly and Levine to then present an overview of several models (war of attrition, uncoordinated rent-seeking, common-pool problem, diversity of preferences, etc.) in terms of how the *number* of ethnic groups exacerbates these dynamics. The accompanying analysis lends empirical support to this theoretical groundwork. But consider two societies that are equally fractionalized, one in which ethnic groups each belong to their own religion and another in which ethnic groups are identically distributed among religions. Which society would we say is more likely to suffer from poor public policies and political instability, that is, which has more ethnic diversity? That the latter is clearly more diverse highlights the weakness of a unidimensional conceptualization and measure of the term diversity. The theoretical and empirical problems are more forceful when comparing a society with just two ethnic groups who belong to different religions, different socioeconomic classes, and live in different regions to a society with many ethnic groups that share a common religion and are evenly distributed both geographically and among social classes.

The move to multiple dimensions, thus, forces scholars to theorize more carefully regarding the relevant cleavages and characteristics and ultimately allows us to build richer theories relating social structure to political phenomena. As we proceed, however, we need to be aware that there are numerous characteristics that can describe social structure in multidimensional space. To date, multidimensionality has been addressed by simply including two fractionalization indices additively into statistical models. However, this approach ignores the way in which multiple dimensions are structured in relation to each other. In this paper, I discuss three distinct multidimensional characteristics of social structure: the first is a simple application of fractionalization to multidimensional subgroupings or *subgroup fractionalization*; second, I turn to a concept with a rich history in political science—*crosscuttingness*—which captures how identically distributed groups on one cleavage are on a second (or more) cleavages; lastly, I discuss *cross-fractionalization*, which is essentially a composite measure of subgroup fractionalization and cross-cuttingness.

Characteristics of social structure were the focus of political sociologists prior to the onset of the Institutional Revolution in the 1970's. Lipset (1960), Rae and Taylor (1970), Truman (1951), and others fuelled debate on how democracy could be established and sustained in the postcolonial world by focusing on the structure of those societies. Cross-cutting cleavages were central to these debates; but, lack of cross-country data to test many of these theories in an increasingly quantitative discipline, along with Lijphart's *consociationalism* that purported to make democracy possible regardless of underlying sociological considerations, put a nail in the coffin of this golden era of sociological theories. With the burgeoning of interest in cultural explanations of political and economic phenomena over the past decade and with sources now available with which to construct cross-national indices of crosscuttingness, the time is ripe for a thorough reexamination of crosscuttingness and other multidimensional characteristics of social structure.

In this paper, I first briefly review the use of multidimensional characteristics of social structure in the political science literature. I then carefully define and measure each concept before introducing the new data set, which consists of numerous indices of these three multidimensional characteristics along five key cleavages in society: race, ethnicity/language, religion, geographic region, and income group.<sup>5</sup> The data set is publicly available from the Inter-University Consortium for Political and Social Research (ICPSR) data archives and the Institute for Quantitative Social Science (IQSS) Dataverse Network.<sup>6</sup> Finally, I reestimate past studies on economic growth and show that these new indices significantly affect economic growth, demanding better theorizing on what we mean by “ethnic diversity” and what mechanisms are at play in this complex relationship. Specifically, I find that crosscuttingness and

<sup>5</sup>The selection of these cleavages is not arbitrary and reflects cleavages that commonly appear in sociological theories. Geographic region is the focus of several recent articles on the number of parties (Mozaffar, Scarritt, and Galaich 2003) and civil war (Cederman, Rød, and Weidmann 2006; Matuszeski and Schneider 2006) as well as in the classic political sociology literature (Rokkan 1967).

<sup>6</sup>The permanent URL's are <http://dvn.iq.harvard.edu/dvn/dv/joelselway> and <http://dx.doi.org/10.3886/ICPSR29602>.

cross-fractionalization both increase the percentage growth in real gross domestic product (GDP) per capita, whereas subgroup fractionalization decreases it.

## 2 Definitions and Measures

### 2.1 Cross-cuttingness Axiom and Subgroup Fractionalization

I first define the *cross-cutting axiom*, which both crosscuttingness and cross-fractionalization conform to, and that any composite measure that captures crosscuttingness must meet. To introduce this axiom, I first discuss, subgroup fractionalization, which does not conform to the axiom.

Let a subgroup be defined as a set of individuals identified by two (or more) groups on two (or more) cleavages to which they belong. For example, Latinos in the United States can be divided into subgroups according to race, for example, Afro-Latinos, white Latinos, or *indigenous* Latinos. More generally, a society with  $x$  groups on one cleavage and  $y$  groups on a second cleavage has  $x \cdot y$  potential subgroups:  $x_1y_1$ ,  $x_1y_2$ ,  $x_2y_1$ , and  $x_2y_2$ , for a  $2 \times 2$  society. Subgroup fractionalization is calculated using the familiar Herfindahl index, with  $p_i$  instead representing the proportion of each subgroup.<sup>7</sup>

To assess if subgroup fractionalization adequately measures crosscuttingness, consider two societies: Jupiter and Saturn. In each society, there are two ethnic groups, Blue and Green, and two religions, Alpha and Beta. In both countries, 50% of the population is Blue and 50% is Green, whereas 50% each are Alpha and Beta also. In Saturn, however, all the Blues are Alpha and all the Greens are Beta, that is, both cleavages, ethnicity and religion, reinforce each other. In contrast, in Jupiter, the two cleavages cross-cut: half of all Blues are Alpha, as are half of all Green, and half of both ethnic groups are Beta.

Figures 1 and 2 represent vertical tree diagrams with the race cleavage at the first node on the tree, branching out to two groups (Blue and Green). The numbers beneath each group are percentages, which sum to unity for each node in the tree diagram. The next level on the tree is the religion cleavage, which likewise branches into two groups (Alpha and Beta). It also has row percentages that sum to one for each of the two nodes at that level. At the bottom of the tree diagram in bold italics are the subgroup proportions. Figure 1 represents Jupiter and Figure 2 Saturn. In Figure 1, all subgroups—Blue Alpha, Green Alpha, Green Alpha, and Green Beta—constitute 0.25 (25%) of Jupiter's population. In contrast, in Saturn, the proportion of both Blue Alpha and Green Beta is 50%, whereas the proportion of Blue Beta and Green Alpha is 0.

Using the Herfindahl index of fractionalization, we can calculate subgroup fractionalization scores for each society. Thus, Saturn has a fractionalization score of  $1 - [(.5)^2 + (.5)^2] = 0.5$ , whereas Jupiter has a much higher score of  $1 - [(.25)^2 + (.25)^2 + (.25)^2 + (.25)^2] = 0.75$ . According to a common interpretation of fractionalization indices in relation to social conflict, a higher score represents more potential for conflict. Since the cleavages in Jupiter crosscut, whereas in Saturn they reinforce each other, this score may not accurately capture the potential for conflict in these two countries.

Next, by making a change to Saturn's social structure, we can further see how inadequate the subgroup-fractionalization measure captures crosscuttingness. In Figure 3, I have moved all the Blues in Saturn to the Beta religion.

Notice that, despite a significant change in Saturn's social structure, the subgroup fractionalization score remains the same, 0.5. We can thus see where subgroup fractionalization fails to capture the concept

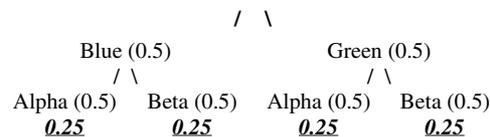


Fig. 1 Jupiter.

<sup>7</sup>The Herfindahl index gives fractionalization as  $F = 1 - \sum_{i=1}^N p_i^2$ .

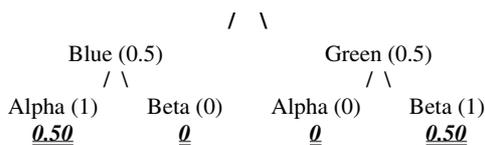


Fig. 2 Saturn.

of crosscuttingness: it is not sensitive to the identities of the groups that compose the subgroups. This leads to the cross-cutting axiom, which stated formally is

$$\frac{\text{Crosscuttingness axiom (subgroup identity sensitivity)}}{\text{Cross-cuttingness is changed whenever a subgroup } x_g y_k \text{ (group } g \text{ on cleavage } x \text{ and group } k \text{ on cleavage } y) \text{ changes in size.}}$$

2.2. A Measure of Crosscuttingness

Having established the axiomatic basis for cross-cuttingness measures, I now define crosscuttingness as follows:

*Crosscuttingness:* Group *i* on cleavage *x* is identically distributed among groups on cleavage *y* with all other groups on cleavage *x*.

Crosscuttingness is basically the concept of statistical independence, which tells us whether knowing to what group an individual belongs on *x* tells us anything about which group she belongs to on *y*. If knowing what group on *x* an individual belongs to tells us nothing what group on *y* she belongs to, then we have perfect crosscuttingness.<sup>8</sup>

Alternatively, we can think of independence as meaning that it does not matter what the distribution of members of *x<sub>i</sub>* is among groups on *y*, as long as they are distributed identically on *y* as all other groups *x<sub>-i</sub>* (*x* “not-*i*”). Accordingly, Figure 4 is perfectly crosscutting, despite the subgroup proportions not being identical.

Figure 4 shows us that in a perfectly cross-cutting society, ethnic groups are identically distributed among religious groups. This identically distributed property, however, also applies to the distribution of religious groups on the ethnic cleavage. Figures 5a and 5b show the same two perfectly cross-cutting societies of Jupiter and Mercury in Figures 1 and 4, respectively, the only difference being that the first branch displays the religion cleavage instead of the ethnic cleavage. In short, in a perfectly cross-cutting society, groups on the religion cleavage are identically distributed across the ethnic cleavage and ethnic groups are identically distributed on the religion cleavage.

My measure for crosscuttingness relies on the familiar chi-square test statistic for independence in basic cross-tabular analysis. The chi-square test statistic measures the deviation of observed proportions from expected proportions. It is given by:

$$\chi^2 \equiv \sum \frac{(O-E)^2}{E}. \tag{1}$$

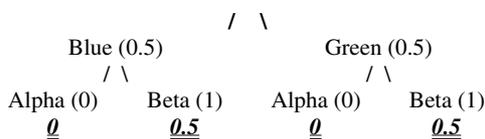
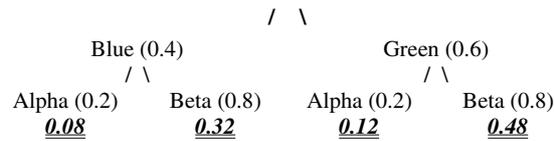


Fig. 3 Saturn, but with all Blues now Beta.

<sup>8</sup>In the language of probability, statistical independence means:  $P(A|B) = P(A)$



**Fig. 4** Mercury, another perfectly cross-cutting society.

For example, consider our two perfectly cross-cutting societies once more. Figure 6 is the same society as in Figures 4 and 5b (Mercury), with the information displayed in a contingency table rather than a tree diagram and now with actual group sizes rather than simply proportions.

The observed frequencies appear in the subgroup cells. Expected frequencies are calculated by multiplying the column and row percentages for each cell by the total sample size, which in this case is 1000. Thus, for Blue Alpha in 6b, we get an expected frequency of  $0.4 \times 0.2 \times 1000 = 80$ . The chi-square statistic is calculated as follows:

$$\chi^2 = \frac{[80-80]}{80} + \frac{[120-120]}{120} + \frac{[320-320]}{320} + \frac{[480-480]}{480} = 0. \quad (2)$$

To make crosscuttingness, now denoted by  $CC$ , comparable regardless of the size of the contingency table, I use the normalization of the chi-square statistic given by Cramer (Agresti 2002). So that higher values of  $CC$  imply higher crosscuttingness, I subtract Cramer's  $V$  from unity, as follows:

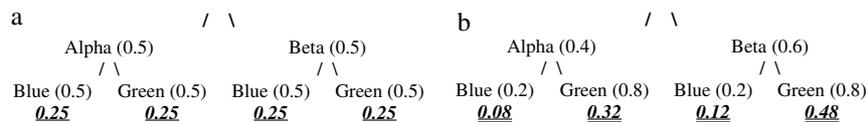
$$\equiv 1 - \sqrt{\left[ \frac{\sum \frac{(O-E)^2}{E}}{nm} \right]}, \quad (3)$$

where  $n$  is the sample size and  $m$  is the smaller of either the number of columns minus one or the number of rows minus one. Cramer's normalization is the appropriate chi-square measure where at least one of the variables is discrete in nature, and the other is either discrete or interval. For the society in Figure 6, we thus get:

$$= 1 - \sqrt{0/(1000 \cdot 1)} = 1. \quad (4)$$

### 2.3. A Measure of Cross-Fractionalization

Rae and Taylor (1970) derived a measure of cross-fractionalization, which I now formally define as “the extent to which individuals who are in the same group on one cleavage are in different groups on the other cleavage,” *which different groups they must also share with members of other groups from the first cleavage.*<sup>9</sup> The second part to this definition (in italics), I add to Rae and Taylor's original definition to reflect the fact that cross-fractionalization satisfies the cross-cutting axiom and can be thought of as capturing both subgroup fractionalization and pure crosscuttingness. Rae and Taylor measure cross-fractionalization as



**Fig. 5** (a, b) Cross-cutting societies, regardless of which cleavage you “begin” with.

<sup>9</sup>The authors refer to their measure as a measure of crosscuttingness and is the only cross-cutting measure that I could find in the political science literature, although its empirical use seems to have been confined to a couple of city attitudinal comparisons in the 1970's (Budge and O'Leary 1971, 1972).

	Alpha	Beta	Row %
Blue	80	120	0.2
Green	320	480	0.8
Column %	0.4	0.6	1.00

Fig. 6 Contingency table of the perfectly crosscuttingness society in 5b.

the proportion of all the pairs of individuals whose two members are in the same group of one cleavage but in different groups of the other cleavage. This leads to the following formula:

$$CF = \frac{A + B}{N(N-1)}, \quad (5)$$

where A is the number of pairs whose members are in the same group on the first cleavage but different groups on the second and B is the number of pairs whose members are in the different groups on the first cleavage but the same on the second cleavage. The denominator denotes the total number of pairs. Assuming that  $N$  is large, so that  $1/N(N-1)$  is approximately equal to  $1/N^2$ , Rae and Taylor show that  $CF$  can be expressed in the form:

$$CF = \sum_{x=1}^n p_x^2 + \sum_{y=1}^n p_y^2 - 2 \sum_{x,y} p_{xy}^2. \quad (6)$$

The three elements of equation (6) should look familiar as they are similar to the Herfindahl index. Specifically, the first element is simply 1 minus the fractionalization score for the groups on cleavage  $x$ , denoted  $F_x$ . The second element is  $1 - F_y$ , and the third element is 1 minus the subgroup fractionalization score. Thus, equation (6) can be written:

$$CF = 2F_C - F_x - F_y,$$

where  $F_C$  represents subgroup fractionalization and  $F_x$  and  $F_y$  the fractionalization scores for each of the individual cleavages.

To see how  $CF$  and  $CC$  differ, consider first the two societies in Figures 5a and 5b, Jupiter and Mercury. Both societies have the same number of groups on each cleavage, but the relative size of groups is bigger in Mercury. Since groups are identically distributed among groups on the other cleavage in both societies,  $CC$  returns the maximum value of 1 for both Jupiter and Mercury.  $CF$ , however, renders the score  $2 \times 0.75 - (0.5 + 0.5) = 0.5$  for Jupiter and  $2 \times 0.6464 - (0.48 + 0.32) = 0.49$  for Mercury. Although these scores are only slightly different (we will see in the empirical section below that they can differ quite significantly), they do reflect the fact that each ethnic group in Jupiter is divided into exactly two religious groups, whereas each ethnic group in Mercury is effectively divided into 1.5 religious groups.

Next consider another perfectly cross-cutting society, Venus, which is displayed in Figure 7. Since Venus is perfectly crosscutting,  $CC$  returns the maximum value of 1;  $CF$ , however, renders the score  $2 \times 0.83 - (0.67 + 0.67) = 0.33$ , reflecting the fact that each ethnic group is now divided into three religious groups. We thus see more clearly now that  $CF$  is sensitive to the fractionalization (number and relative size of groups) of the individual cleavages.  $CC$  can thus be considered a “pure” measure of cross-cuttingness, whereas  $CF$  is most appropriate for measuring the characteristic of cross-fractionalization. A series of simulations in which I vary the number of groups on and the fractionalization of each dimension and which more fully exhibits the different properties of crosscuttingness and cross-fractionalization is given in the supplementary material.

Consider once more the hypothetical society of Mercury in Figures 4 and 5b. If I make a simple change to this society by switching the proportion of the Alpha and Beta religions under the Blue ethnic group, note first that subgroup fractionalization does not change. However, cross-fractionalization does change because that switch has altered the proportions of religious groups in Mercury. Alpha now comprises 44% of society (previously was 20%) and Beta now comprises 56% (previously 80%). Since the religious fractionalization score is a major component of cross-fractionalization, Mercury’s  $CF$  score now falls to 0.32.

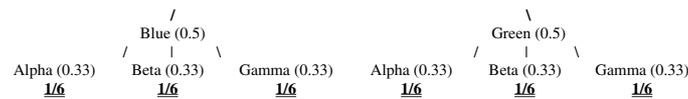


Fig. 7 Venus.

This much lower score reflects the fact that the majority (80%) of the Blue ethnic group now belongs to a different religious group than the majority (also 80%) of the Green ethnic group. Indeed, testing *CF* under a variety of different hypothetical societies, I found it to always conform to the cross-cutting axiom. A detailed simulation is given in the supplementary material, which is accompanied by an Excel file with a *CC* and *CF* calculator. Moreover, in the empirical section below, I demonstrate that the two measures do indeed seem to be capturing different characteristics.

#### 2.4. Limitations and Assumptions of *CC*

My measure for crosscuttingness, *CC*, makes several assumptions, which I wish to make plain here. First, the measures assume that group categories within cleavages are mutually exclusive. Assuming that an individual cannot share identity in multiple categories of a single cleavage becomes especially problematic in melting-pot countries such as the United States. However, researchers might want to relax this assumption, though this extension would be some undertaking. Relaxing the assumption certainly has implications for the underlying logic of crosscuttingness theories—conflicting loyalties. If an individual can belong to more than one group within a cleavage, then certainly her level of conflicting loyalties increases.

A second assumption *CC* makes is an equal level of polarization (cultural distance) among groups within cleavages. For example, the difference between being Black and White in a given society may not be the same as the difference between being Black and Asian or White and Hispanic. Fearon (2003) offered an innovative measure of cultural distance—structural distance between languages—as a proxy for the cultural distance between ethnolinguistic groups.<sup>10</sup> However, to my knowledge, similar classification systems do not yet exist for other cleavages such as race and religion.

A third assumption *CC* makes is that all cleavages are of equal political salience. This is neither a necessary nor an unrestrictive assumption. The measure could be adjusted, however, to allow for the weighting of cleavages. The *salience* (importance of dividing line and its associated conflicts to those involved) of different cleavages in a society has been the focus of work by Posner (2004), who has suggested detailed surveys to capture true attitudes toward different group identities within a country. Posner (2004) himself, as well as Wilkinson (2000), rely on content analysis of sampled newspapers to measure the salience of cleavages in Africa and India, respectively. Alternatively, Mozaffar, Scarritt, and Galaich (2003) attempted to quantify salience by identifying groups that are “politicized.” Their Ethnopolitical Group Fractionalization is based on five “objective” criteria with which they scan secondary academic resources to determine politically salient groups. Such efforts may, in the future, be combined with the indices of crosscuttingness presented in this paper.

Finally, my measures of cross-fractionalization and cross-cuttingness are also currently limited to two dimensions. In future versions of this data set, I plan to extend the cross-fractionalization and cross-cuttingness measures to three and four dimensions.<sup>11</sup>

### 3 The Data Set

I introduce several new indices of crosscuttingness, cross-fractionalization, and subgroup fractionalization across the following five cleavages: race, ethnicity/language, religion, geographic region, and income. I also calculate single-dimension indices (fractionalization and bipolarization) for each cleavage. In total, there are 69 new indices for a total of 128 current countries:<sup>12</sup> 9 cross-cutting indices,

<sup>10</sup>The measure uses the linguistic classification system, which categorizes each language into an increasingly broader group, branch, family, stock, and phylum (the broadest linguistic class).

<sup>11</sup>Discussion of future plans in this direction are given in the supplementary material.

<sup>12</sup>A few of these countries have very few indices (Yemen, Palestine, Lebanon, Kuwait, Kazakhstan, Iceland, Malta, and Fiji). There are also a handful of former countries and some provinces (Hong Kong, Northern Ireland, and Puerto Rico) in addition to these 128 countries.

9 cross-fractionalization indices, 24 fractionalization/subgroup fractionalization indices, and 24 bipolarization/subgroup bipolarization indices.<sup>13</sup> Following, I first describe the data sources and summarize some of the technical decisions regarding the construction of the indices. Then, I describe the indices themselves in terms of geographic spread and GDP-per-capita variation. Next, I compare the empirical content of these indices with each other to demonstrate that they do actually capture distinct features of a society. Finally, I compare the single-dimension indices with existing indices of ethnic and religious fractionalization and discuss issues of reliability.

### 3.1 Sources

The problem with compiling indices for multidimensional characteristics of social structure in the past has always been the lack of information on subgroup proportions. One possible source of information entails surveys wherein individuals are asked to identify themselves along the cleavages of interest. The most accurate type of survey would be a country census, wherein the entire population is polled. However, many countries do not collect information on ethnicity or race, and for those that do the raw data is difficult to obtain. An alternative strategy is to extract similar information from nationally representative surveys, such as public opinion surveys. Accordingly, I compiled my indices from nine sources: The World Values Survey,<sup>14</sup> European Values Survey,<sup>15</sup> The Eurobarometer,<sup>16</sup> the Afrobarometer,<sup>17</sup> Arab Barometer, the Latin American Public Opinion Project,<sup>18</sup> the Asian Barometer,<sup>19</sup> the Comparative Study of Electoral Systems,<sup>20</sup> and a survey performed by the World Health Organization.<sup>21</sup> The use of several surveys also allowed me to test the robustness of my scores for many countries that appeared on two or more surveys.

### 3.2 Decision Rules

The data set, which is available at the author's Web site, includes the contingency tables for each country survey used in the construction of the indices. These contingency tables allow researchers considerable freedom to select the appropriate level of aggregation for their project. The main principle underlying my index is to identify levels of aggregations for each cleavage that reflect *latent* macro social structure in each country. The major implication of latent indices is the inclusion of all categories picked up in the survey, even if only a single individual belongs to the group. Nevertheless, since I have several sources for most countries, I average the scores for all surveys undertaken in each country. More detailed information on the decision rules are in the supplementary material.

Briefly, however, I eliminated surveys if they significantly affected representativeness along any dimension, disagreed with other information sources (encyclopedias and existing ethnicity data sets), or hampered comparability across countries. For example, unless only a single survey was available, I only included surveys that required respondents to place themselves in income deciles.

<sup>13</sup>Devised by Montalvo and Reynal-Querol (2003), bipolarization (*B*) captures the degree there are two equally sized groups in society and is measured by the formula:

<sup>14</sup><http://www.worldvaluessurvey.org>.

<sup>15</sup><http://www.europeanvaluesstudy.eu/>.

<sup>16</sup>See Antonis Papacostas, *Eurobarometer 63.4: European Union Enlargement, the European Constitution, Economic Challenges, Innovative Products and Services*, May–June (2005); R. Soufflot de Magny, *Eurobarometer 63.3 Lifelong Learning in the Ten New European Union Member Countries and Consumer Rights in Poland*, March–April (2005); Karlheinz Reif and George Cunningham, *Central and Eastern Euro-barometer 3: Political Disintegration*, October–November (1992).

<sup>17</sup><http://www.afrobarometer.org>.

<sup>18</sup><http://www.vanderbilt.edu/lapop/HOME>.

<sup>19</sup>Takashi Inoguchi et al. *AsiaBarometer Survey Data* (2003 or 2004) [computer file]. AsiaBarometer Project (<http://www.asiabarometer.org/>) [producer and distributor] (accessed December 2006).

<sup>20</sup><http://www.cses.org>.

<sup>21</sup>Gary R. Andrews, *World Health Organization Collaborative Study on Social and Health Aspects of Aging in Fiji, Korea, Malaysia, and the Philippine* (1983–5); *Burma, Indonesia, North Korea, Sri Lanka, and Thailand* (1990); and *Bahrain, Egypt, Jordan, and Tunisia* (1991).

**Table 1** Ethnoreligious subgroups in Switzerland

<i>Lang. Rel.</i>	<i>Buddhist</i>	<i>Hindu</i>	<i>Muslim</i>	<i>Orthodox</i>	<i>Other</i>	<i>Protestant</i>	<i>Catholic</i>	<i>Row Total</i>	<i>Row (%)</i>
French	1	0	0	1	2	135	200	<b>339</b>	<b><i>0.31</i></b>
German	0	0	1	2	16	297	209	<b>525</b>	<b><i>0.48</i></b>
Italian	0	0	0	1	1	1	185	<b>188</b>	<b><i>0.17</i></b>
Spanish	0	0	0	0	0	2	7	<b>9</b>	<b><i>0.01</i></b>
Other	0	1	3	3	0	7	13	<b>27</b>	<b><i>0.02</i></b>
Column Total	<b>1</b>	<b>1</b>	<b>4</b>	<b>7</b>	<b>19</b>	<b>442</b>	<b>614</b>	<b>1088</b>	<b><i>0.31</i></b>
Column (%)	<b><i>0.001</i></b>	<b><i>0.001</i></b>	<b><i>0.004</i></b>	<b><i>0.006</i></b>	<b><i>0.017</i></b>	<b><i>0.406</i></b>	<b><i>0.564</i></b>		

Source: World Values Survey, 1996.

Note. Bolded values are row/column totals. Bolded and italicized values are row/column percentages.

### 3.3 Example Calculation

So that the reader can get a sense of how *CC* is calculated for an actual case, and to increase familiarity of the group categories garnered from the surveys, I take the reader through the calculation of ethnoreligious crosscuttingness for Switzerland. Table 1 shows the distribution of ethnic (linguistic, primary language in the home) and religious groups in Switzerland. We can see that the German and French ethnic groups are quite evenly spread over two religious denominations: Protestantism and Catholicism. Germans tend to identify slightly more with Protestantism, whereas French identify more with Catholicism. Nevertheless, ethnicity and religion are quite crosscutting over these groups. We can further see that the Italian ethnic group is almost entirely Catholic. This brings Switzerland's score down from what would otherwise be a fairly high score.

Table 2 shows the expected values for each cell. These expected frequencies are calculated by the *Row%* multiplied by the *Column%* multiplied by the total number of observations. Thus, for French Buddhists, the expected frequency is  $.001 \times .31 \times 1088 = 0.31$ , which is the value of the top-left cell in Table 2. The expected frequencies in Table 2 are then subtracted from the observed frequencies in Table 1, squared, and divided by the expected frequencies again. Thus, for the top left cell, we get  $(1 - .31)^2 / .31 = 1.56$ . We repeat this for the remaining 34 cells and sum them. This gives us the chi-square statistic of 381.20. Next, we divide this statistic by the product of the total number of observations in the sample (1088) and the number of groups minus one in the least fractionalized of the two cleavages (4). Finally, we subtract this from unity, giving us a crosscuttingness score of .704.

### 3.3 Representativeness

Of the 168 countries with populations over 400,000, I was able to compile scores for 124 countries or 74% of current countries.<sup>22</sup> In general, the indices have good regional coverage. Surveys were most readily available in Europe (West and East), the Americas, and the Asia-Pacific. Africa and the Middle East are the most underrepresented regions in my indices, with West Africa having available surveys for just 7 of the 20

**Table 2** Expected values of ethnoreligious subgroups in Switzerland

	<i>Buddhist</i>	<i>Hindu</i>	<i>Muslim</i>	<i>Orthodox</i>	<i>Other</i>	<i>Protestant</i>	<i>Catholic</i>
French	0.31	0.31	1.25	2.18	5.92	137.72	191.31
German	0.48	0.48	1.93	3.38	9.17	213.28	296.28
Italian	0.17	0.17	0.69	1.21	3.28	76.38	106.10
Spanish	0.01	0.01	0.03	0.06	0.16	3.66	5.08
Other	0.02	0.02	0.10	0.17	0.47	10.97	15.24

<sup>22</sup>At the time of this article being accepted for publication, work is already underway to increase the number of countries using the *Demographic and Health Surveys*. See <http://www.measuredhs.com>.

**Table 3** Representativeness of world geographic regions and of economic development

	Average GDP Per Capita (US\$)	
	All region	Missing countries
22/22 East Europe	12,039	n/a
18/27 Middle East	10,721	13,344
25/25 Western European	35,017	n/a
18/24 Asia-Pacific	10,120	2830
7/20 West Africa	3039	3333
13/26 South, East, and Central Africa	3487	3186
23/25 Latin America	7688	5336
124/168 All countries	12,148	6518

Note: n/a, not applicable

countries. The figures for the remaining regions are as follows: 25 of 25 countries in Western Europe, 22 of 22 in Eastern Europe, 18 of 24 in Asia-Pacific, 23 of 25 in Latin America, 18 of 27 in the Middle East and North Africa, 20 of 46 in Sub-Saharan Africa.

Since this index will likely be used to study phenomena where the level of economic development may be an important determinant or confounding variable, it is important to ensure that the sample of countries I have in my index are representative in terms of country wealth. The average GDP per capita of countries not included in the cross-cutting and cross-fractionalization measures are given in Table 3.

The difference in economic development between in- and out-sample countries is just over \$5000. However, this disparity in levels of economic development is not seriously pronounced in the regions most underrepresented in my indices. Indeed, in the Middle East and West Africa, out-sample countries actually have a higher GDP per capita. In the rest of Sub-Saharan Africa, the GDP per capita of out-sample countries is only slightly lower. Nevertheless, this difference in average GDP per capita is certainly something to be aware of when evaluating results used with these indices.

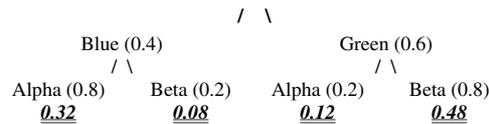
A final aspect of representativeness concerns the exclusion of conflict countries where survey researchers are less likely to have access. Using the Major Episodes of Political Violence data on civil wars and violence, I calculate that there was at least one civil or ethnic war onset between 1946 and 2008 in 61 countries. Of these 61 countries, 37, or 60%, appear in my data set. Of the 99 total onsets, I have data for 70. In addition, there were 100 countries that experienced a major episode of political violence, including both civil wars and lower magnitude conflicts. Of these 100, I have data for 66, and of the 236 onsets of major political violence, I have data for 177 or 75%. In sum, although researchers should be aware of these limitations, the coverage of these indices is fairly good and comparable to the study by Montalvo and Reynal-Querol (2005b) study on civil war using their new bipolarization index.

### 3.4 Are Empirical Values of Multidimensional Measures Different?

Figure 8 shows a scatterplot of linguistic-religious crosscuttingness and cross-fractionalization. Both measures are on a 0–1 scale, though the maximum value of *CF* is 0.794 (Guyana). Although there is a moderate positive relationship (0.32 Pearson correlation), it is clear that for the majority of observations, which lie in the right half of the graph, a country's level of crosscuttingness does not predict its level of cross-fractionalization. That the few observations in the left half of the graph are more strongly correlated supports my earlier assertion that conceptualizations of crosscuttingness in past literature seem to agree on the definition of the minimum level of crosscuttingness (reinforcingness). Indeed, from a *CC* value of about 0.5 upward, *CF* ranges almost the entire spectrum. Thus, *CC* and *CF* are clearly very different characteristics.

I turn now to compare crosscuttingness to the single-dimension characteristics of ethnic and religious fractionalization. Figure 9a–9d show the relevant scatterplots. The cloud of observations in both figures implies that crosscuttingness has a very low level of correlation with ethnic/linguistic fractionalization (0.308) and religious fractionalization (0.04).

Finally, I compare Linguistic-Religious Cross-Fractionalization to Linguistic-Religious Subgroup Fractionalization. There are a few countries for which the scores are similar for these two characteristics.



**Fig. 8** Mercury, with proportions of alpha and beta switched for the Blue ethnic group.

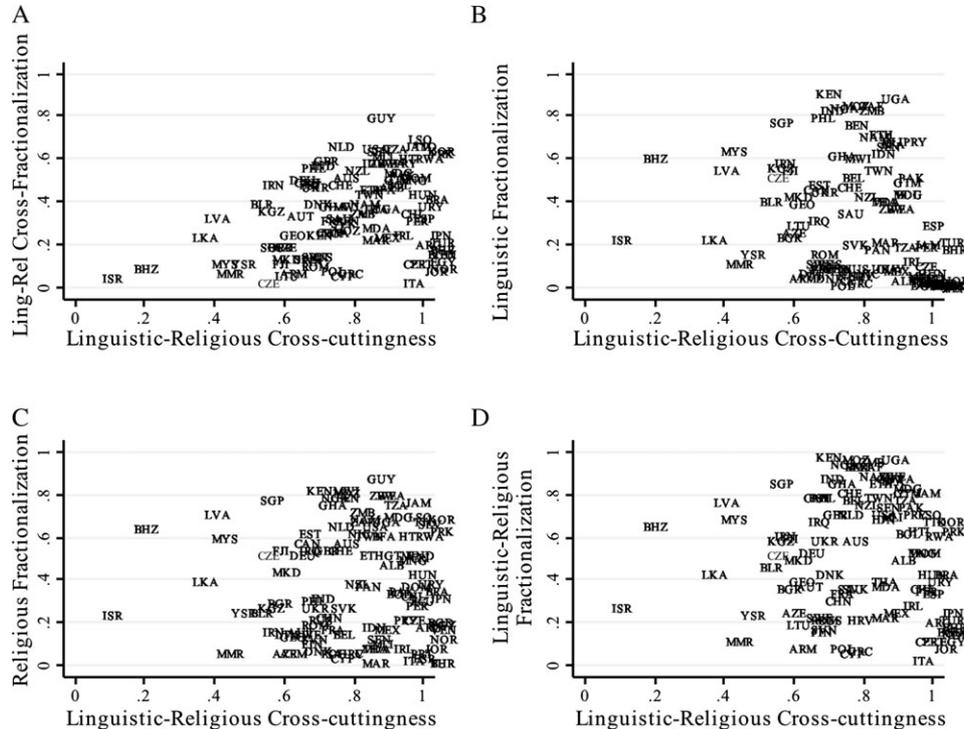
This is reflected in a somewhat high correlation of .691. Nevertheless, we see that a significant number of countries lie off the diagonal, supporting my theoretical prediction that these are indeed different characteristics of social structure.<sup>23</sup>

### 3.5 Robustness of Data

I computed ethnic and religious fractionalization and bipolarization scores and compared them with existing indices: Reynal-Querol's (2002), Fearon's (2003) Ethn-linguistic fractionalization index, and Alesina et al. (2003). The data are highly correlated with these data sets for both ethnic (as high as .89) and religious fractionalization (.71). The accompanying appendix details the full correlation tables. I also compared all measures within countries before averaging. This allowed me to assess the extent that different surveys were producing the same scores for each country. The average SDs range from .003 to .076, depending on the measure. Information on SDs and full contingency tables that allow researchers to make their own decisions on inclusion of certain surveys are available in the master data set.

## 4 Crosscuttingness across the World

Do regions of the world differ significantly across these multidimensional characteristics of social structure? Table 4 compares regions of the world along the ethnoreligious crosscuttingness index.<sup>24</sup> Latin America is by far the most ethno religiously crosscutting, with an average score of .991 compared to



**Fig. 9** (A–D) Scatterplots of fractionalization, crosscuttingness, and cross-fractionalization.

<sup>23</sup> Additional scatterplots are presented in the supplementary material.

<sup>24</sup> Additional indices compared across regions are presented in the supplementary material.

a world average of .72 and .775 for the next most cross-cutting region. This high level of ethnoreligiously crosscuttingness stems from the overwhelming strength of Catholic religious identity in the region. Africa is the second most ethnoreligiously cross-cutting region, stemming from the multiethnic proselytizing efforts of both Christianity and Islam, one of which usually dominates in a particular country. Nigeria and Ghana are exceptions to this fairly high level of ethnoreligiously crosscuttingness because both Islam and Christianity constitute major proportions of the population. Western Europe, Asia, and the Middle East have similar medium levels of ethnoreligious crosscuttingness, indicating that there is a higher tendency for different ethnic groups to belong to their own religions than in Latin America and Africa. Eastern Europe is the least ethnoreligiously crosscutting meaning that ethnic groups are highly defined by religious affiliation. This low score reflects the fact that the dividing line between Catholicism and Orthodoxy runs through the center of this region. Ethnic minorities that spill over into neighboring countries, therefore, also belong to a different sect of Christianity. Islam also penetrated the region, and the Balkans and peripheral areas of Russia have concentrated Muslim populations. The country with the lowest level of ethnoreligious crosscuttingness is Israel, where Arabic Speakers are overwhelmingly Muslim and Hebrew Speakers are almost exclusively Jewish.

## 5 Empirical Significance

As a first effort to demonstrate the potential empirical impact of these new indices, I reevaluate studies on ethnicity and economic growth. I show that past studies have presented an incomplete picture of the effect of ethnic fractionalization on economic growth. Indeed, I show that the effect of ethnic fractionalization on economic growth is conditional on levels of crosscuttingness and cross-fractionalization: at high levels of ethnoreligious CC and CF, ethnic fractionalization has no significant effect on growth rates. This result underscores the potential for new understanding of the relationship between social structure and not just economic growth but a variety of other phenomena in comparative political economy.

Ethnic diversity is theorized to affect economic growth via three mechanisms, recently categorized by Habyarimana et al. (2007) into three families: preferences, technology, and strategy selection.<sup>25</sup> But do more ethnic groups necessarily result in more polarized and indivisible *preferences* regarding the nature of the policies provided and the perceived beneficiaries of public goods? Do more ethnic groups necessarily increase difficulty in communication in a policymaking setting (*technology* mechanism) or decrease levels of trust (*strategy selection* mechanism)? Although Easterly and Levine (1997) present strong evidence to this effect, Montalvo and Reynal-Querol (2005a) alternatively offer a different proxy, *bipolarization*, or the extent that a country is divided into two equally sized ethnic groups, as a better predictor of economic growth. But it is not clear that this alternative proxy is an improvement in concept validity, the extent that an operationalization captures the concept it purports to measure, in this case ethnic diversity.

Both proxies assume that the groups are highly polarized, which may not be true whether we are considering many groups or two evenly sized groups. By polarization, I am referring to how different ethnic

**Table 4** Ethnoreligious crosscuttingness by region of the world

	<i>West Europe, North America, and Oceania</i>	<i>East Europe, Former Soviet Bloc</i>	<i>Middle East and North Africa</i>	<i>Sub-Saharan Africa</i>	<i>Asia-Pacific</i>	<i>Latin America and Caribbean</i>
Count	28	22	11	21	18	22
Mean	0.71	0.61	0.73	0.8	0.72	0.9
Min	0.2	0.15	0.05	0.64	0.32	0.76
Max	1	0.93	1	0.95	1	1
Variance	0.03	0.03	0.09	0.01	0.05	0

<sup>25</sup>There is a further intervening mechanism by which ethnic diversity leads to the underprovision of policies that sustain economic growth, such as the building of infrastructure, the provision of education, and vital financial measures like reining in the black market premium and restraining government consumption. In addition, ethnically fractionalized countries are more likely to experience civil conflict and government instability.

groups are in terms of their preferences and modes of communication and how low their levels of trust are. The multidimensional characteristics of social structure discussed in this paper may better capture the likelihood that ethnic groups are polarized by capturing how ethnic groups differ along other socially salient cleavages. In terms of preferences, if ethnic groups are concentrated in their own region, it is easier to demand targeted goods to one's ethnic group. Likewise, if ethnic identity is reinforced by income, ethnic groups are more likely to have strong preferences over the distribution of government resources. Ethnic groups that do not share religions may have distinct preferences regarding types of schooling, such as the inclusion of religious education in the curriculum. In short, this implies that lower levels of crosscuttingness would result in lower levels of economic growth.

Even if we think that the number of groups in society is what matters, does the logic extend to subgroups? Does a higher number of ethnoreligious, ethnogeographic, or ethno-income groups entail more diverse preferences, more difficulty in communication, or lower levels of trust? Perhaps an ethnic group divided among several religions—which have their own “language,” set of norms, ontologies, and moral/ethical values that in many cases are more clearly articulated and specifically applied (often in written form) to politics, policies, and public goods—will find its religious subgroups polarized. One could make a good case, then, for subgroup fractionalization being a better proxy of social diversity than the one-dimensional ethnic fractionalization measure. If an ethnic group is internally divided by a second cleavage, does it matter if it shares membership with other ethnic groups on that second cleavage or just that it is heavily divided? This main distinguishing factor between subgroup fractionalization and cross-fractionalization—the sharing of memberships with other ethnic groups—would result in a different hypothesized direction between ethnic diversity and economic growth for each proxy: a negative relationship for subgroup fractionalization, but a positive one for cross-fractionalization.

Although recognizing the need to develop a more in-depth theory regarding social structure and economic growth, I nevertheless estimate a cross-national growth regression substituting various multidimensional indices in place of ethnic fractionalization. In addition, I estimate a multiplicative model between ethnic fractionalization and two of the multidimensional measures presented in this paper: crosscuttingness and cross-fractionalization. In short, I test four hypotheses based on the brief preceding discussion:

1. Hypothesis 1: An increase in subgroup fractionalization will result in a decrease in economic growth.
2. Hypothesis 2: An increase in crosscuttingness will result in an increase in economic growth.
3. Hypothesis 3: An increase in cross-fractionalization will result in an increase in economic growth.
4. Hypothesis 4: An increase in ethnolinguistic fractionalization will result in a decrease in economic growth but only at low levels of crosscuttingness and cross-fractionalization.

### 5.1 Data and Variables Description

I test my hypotheses in a sample of approximately 100 countries between 1972 and 2003.<sup>26</sup> My dependent variable is taken from the World Bank Development Indicators and is the percentage growth in real GDP per capita. I also control for several variables common in past studies. The first four are from World Bank data: the level of GDP per capita (*GDP*) as a measure of a country's level of development;<sup>27</sup> the log of gross domestic investment as a percentage of GDP (*Invest*); the log of consumption spending as a percentage of GDP (*Spend*); and the log of net inflows of foreign direct investment (*FDI*). Next, I include a dummy variable for the occurrence of a major conflict, defined by UCDP/PRIO Armed Conflict Dataset as greater than 1000 deaths per year (*Conflict*). I include three variables to capture public goods provision: from United Nations Development Program data, I include the log of the literacy rate as a measure of

<sup>26</sup>Note both the number and the actual countries in each regression differ based on the availability of data for each of my indices.

<sup>27</sup>Note in pooled samples, this variable is usually the average GDP of the initial period (GDPI). The squared term is also included in the model (GDPI<sup>2</sup>Q). The results are similar in the random effects model substituting GDPI and GDPI<sup>2</sup>Q for GDP.

**Table 5** Ethnic diversity and long-run growth (growth of per capita real GDP)

	(1)	(2)	(3)	(4)	(5)	(6)
GDP	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
FDI	0.56*** (0.09)	0.58*** (0.09)	0.57*** (0.09)	0.58*** (0.10)	0.57*** (0.12)	0.60*** (0.13)
Invest	2.65*** (0.48)	2.61*** (0.51)	2.87*** (0.48)	2.77*** (0.52)	2.79*** (0.89)	2.59*** (0.89)
Spend	-0.86 (1.32)	-0.88 (1.38)	-0.41 (1.33)	-0.57 (1.41)	-0.97 (2.61)	-1.14 (2.81)
Lit	1.16 (0.77)	1.30 (0.80)	1.43* (0.78)	1.53* (0.82)	1.23 (0.77)	1.37* (0.82)
Life	3.75* (2.02)	2.63 (2.12)	3.95* (2.06)	2.82 (2.15)	3.75 (3.83)	2.39 (3.68)
Tel	-1.11*** (0.25)	-1.03*** (0.26)	-1.12*** (0.25)	-1.11*** (0.27)	-1.09*** (0.27)	-1.02*** (0.29)
CW	-1.45*** (0.54)	-1.35** (0.62)	-1.58*** (0.57)	-1.42** (0.63)	-1.54* (0.92)	-1.30 (0.87)
Brit	0.13 (0.53)	0.40 (0.58)	0.19 (0.55)	0.07 (0.59)	0.21 (0.55)	0.35 (0.55)
Dem	-0.00 (0.03)	-0.00 (0.03)	-0.01 (0.03)	0.01 (0.03)	-0.01 (0.04)	-0.00 (0.04)
Africa	2.94*** (1.04)	3.05** (1.20)	1.64 (1.08)	2.19* (1.17)	2.24* (1.35)	3.25** (1.41)
Latin	0.98 (0.67)	1.43* (0.78)	0.40 (0.73)	1.32 (0.85)	0.61 (0.77)	1.44 (0.96)
Asia	2.88*** (0.80)	2.85*** (0.85)	2.19*** (0.80)	2.50*** (0.87)	2.71** (1.06)	3.12** (1.15)
MidE	3.27*** (1.03)	2.73** (1.08)	3.02*** (1.05)	3.12*** (1.14)	3.18*** (0.78)	3.01*** (0.82)
LF	-2.60*** (0.96)				-7.54* (4.14)	-6.02*** (2.29)
LRFC		-2.85** (1.16)				
LRCC			2.88** (1.39)		0.03 (1.54)	
LRXC				-1.79 (1.44)		-3.94** (1.98)
LF × LRCC					7.92 (5.34)	
LF × LRXC						9.70** (4.94)
Constant	-13.59 (10.71)	-9.45 (11.31)	-20.47* (10.80)	-12.93 (11.43)	-13.88 (18.46)	-7.15 (18.45)
N	1814	1649	1781	1617	1781	1617
No. countries	98	90	96	88	96	88

FDI, foreign direct investment. Random effects model with SEs in parentheses.

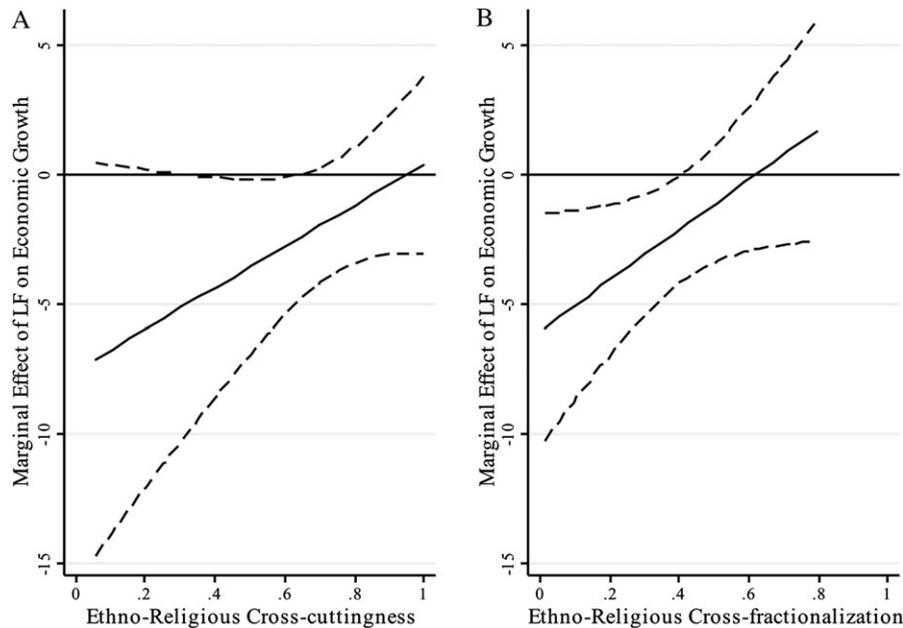
\* $p < .1$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

investment in schooling (*Lit*) and the log of life expectancy (*Life*), and from the Banks data set, I include the log of per capita users of telephones, as a measurement of investment on infrastructure (*Tel*). I also include a set of dummy variable for regions, former British colonies, and decade. Finally, I include a measure of how democratic a country is from the Polity IV data set.<sup>28</sup>

## 5.2 Estimation Strategy

I estimate the models using a time series cross-sectional (TSCS) regression with random effects. The substantive results are robust to two other estimations: TSCS with vector-decomposed fixed effects (Plümpert and Troeger 2007), which accounts for the “stationary” nature of the social structure measures, and

<sup>28</sup>Various models were run with lags of the independent variables (where deemed appropriate), but because they do not produce significantly different results, I simply report the unlagged variables in this paper. The results are also robust to choices of logging certain variables. I also estimated models with different proxies for conflict, including the UCDP/PRI0 data on number of conflicts and numerous measures of social conflict (government crises, riots, guerilla warfare, and the composite measure) from the Banks data set. I substitute gross school enrollment for the literacy rate. Although some of the social structure variables lose significance in some models, the main findings do not significantly change.



**Fig. 10** (a, b) Marginal effect of ethnic fractionalization on economic growth conditional on ethno-religious cross-cuttingness and cross-fractionalization.

a Seemingly Unrelated Regression (SUR) estimation with the data pooled by decade, which many past studies on economic growth have used.

### 5.3 Results

Table 5 presents the results from the ethnic and religious cleavages.<sup>29</sup> In accordance with several previous studies (Easterly and Levine 1997; La Porta et al. 1999; Posner 2004; Alesina and LaFerrara 2005), I find that ethnolinguistic fractionalization (LF) has a negative effect on economic growth (Table 5). Indeed, going from zero fractionalization (LF = 0) to complete fractionalization (LF = 1) would result in a 2.6% decrease in economic growth. Taken literally, the coefficient in regression (1) implies that if Nigeria had the sample mean value of LF (0.29) instead of its actual value of 0.83, it would have experienced a growth rate just under 50% higher than its average of 2.9% per annum for the 29 years in the sample.<sup>30</sup>

Turning now to subgroup fractionalization, I find weak support for Hypothesis 1. The linguistic-religious subgroup fractionalization index takes on a negative coefficient.<sup>31</sup> This result is suggestive of subgroups exhibiting some of the same mechanisms (preferences, technology, and strategy selection) as ethnic groups, opening up an avenue for further research on diversity within ethnolinguistic groups. More analysis needs to be done to see exactly what types of societies are producing this empirical relationship. It is possible that much of this observed correlation could result from countries with one very large ethno-religious group split among numerous religious affiliations (Western Europe with its many Christian denominations and atheists) or in countries with one religion with numerous ethnic groups (Afghanistan, Pakistan). Regardless, better theorizing needs to be done to understand how linguistic-religious groups form preferences over public policies and how they interact in society.

In accordance with Hypothesis 2, the coefficient on linguistic-religious cross-cuttingness (LRC) takes on a positive sign and is statistically significant.<sup>32</sup> Moreover, the magnitude of LRC's effect is slightly higher than that of ethnolinguistic fractionalization. Taken literally, this means that the coefficient in

<sup>29</sup>The remaining analysis can be found in the supplementary material.

<sup>30</sup>Though I do not present the results in this paper, I also find that religious fractionalization, ethnolinguistic and religious bipolarization, and linguistic-religious bipolarization have a significant negative effect on economic growth.

<sup>31</sup>The other indices all take on negative coefficients but are not statistically significant.

<sup>32</sup>Three of the four other indices take on negative coefficients but are not statistically significant.

regression (10) implies that if Myanmar had the sample mean value of LRC (0.76) instead of its actual value of 0.32, it would have experienced a growth rate over one-third higher than its actual average of 3.8% per annum for the 28 years in the sample. Again, more theoretical precision is needed before firm conclusions can be drawn. Hypothesis 3 is not supported by the results in Table 5, however. Although it is not statistically significant, linguistic-religious cross-fractionalization takes on a negative sign contrary to our predictions. However, the coefficients of the other four cross-fractionalization indices (see supplementary material) were all positive and statistically significant.

Finally, I turn to Hypothesis 4, which is tested in Models 5 and 6 of Table 5. Both models show a negative coefficient on ethnolinguistic fractionalization and a positive coefficient on the interaction terms. These coefficients seem to support the idea that as crosscuttingness and cross-fractionalization increase, the effect of LF decreases. However, in order to fully assess this relationship, I follow recent convention (Brambor, Clark, and Golder 2006; Kam and Franzese 2007) in plotting the marginal effects, Figure 10a and 10b. These plots indeed give us a more accurate picture as they show that at high levels of LRCC and LRXC, the effect of LF is statistically insignificant (confidence intervals fall either side of the zero line). Figure 10b, especially, provides strong support for Hypothesis 4 since at low levels of LRXC, LF has a negative and significant effect on economic growth. Figure 10a provides weaker support for Hypothesis 4: the effect of LF is only significant at medium–low levels of LRCC. At very low levels of LRCC, the effect of LF is again insignificant.

## 6 Conclusion

This preliminary analysis has suggested that a much richer theory of ethnic and social diversity is needed. If we think that the structure of ethnicity vis-à-vis other salient cleavages in society is predictive of the likelihood these social groups will be able to coordinate governance and develop policies in favor of economic growth, these new multidimensional indices go a long way in honing in on precise theoretical predictions and intervening mechanisms. The results provide strong preliminary evidence that the effect of ethnic fractionalization is conditional on other features of a society's social structure. Although I have only presented evidence regarding how ethnicity is structured in relation to religion (LRCC and LRXC), the indices also allow us to develop and test theories regarding the geographic distribution of ethnic groups, ethnic income inequality, regional income inequality, etc. In short, there are numerous avenues of research relating ethnic structure, not to mention other salient cleavages such as religion and race, to economic performance.

Several active research programs in the economics and political science disciplines rely on measures of ethnic and religious diversity for empirical evaluation. The indices presented in this paper represent the first attempt at moving beyond single-dimension measures and can be used in combination with existing single-dimension measures such as ethnic fractionalization and polarization. Just as there are numerous single-dimension characteristics of social structure, in this paper, I have presented three multi-dimensional characteristics of social structure with accompanying measures and cross-national data.<sup>33</sup> I have shown that empirically these measures are capturing different aspects of social structure. Finally, I have demonstrated their potential contribution in an analysis of ethnic diversity and economic growth.

Before concluding, I note two shortcomings of the data set that will be addressed in the second phase of development. First, the underrepresentation of Sub-Saharan Africa and the Middle East demands attention to filling in missing countries of these indices. A second prominent shortcoming is the current restriction of the crosscuttingness and cross-fractionalization indices to two dimensions. The extension of subgroup fractionalization to multiple dimensions was fairly easy to undertake. However, the statistics literature is not clear on the best approach to measuring independence for more than two random variables, and attention to this debate should precede the extension of crosscuttingness to multiple dimensions. In addition, the measure of cross-fractionalization, which was devised by two Political Scientists, has currently not been extended to more than two dimensions. Phase two of this data set will attempt to tackle

<sup>33</sup> Again, I actually compile data on a fourth multidimensional measure, subgroup bipolarization, though I only make brief mention of it in this paper.

these issues and make crosscuttingness and cross-fractionalization measures available in three and four dimensions and across a larger number of countries.

There has been somewhat of a trend away from structural variables in the political science literature to factors with lower causal priority, that is, one step closer to the dependent variable in the causal chain (Gerring 2005). Chandra and Wilkinson (2008) categorize these recent approaches as proxies for “ethnic practice,” and their own data set on the ethnicization of party systems is a good example of how ethnic differences are manifested in the political arena. Another example is the measure by Wimmer et al. (2009) of “ethnic power relations,” which identifies all politically relevant ethnic groups and their access to state power. Both these examples move directly to the coordination of policies made by political elites, which is a welcome new direction since it gets at the underdeveloped intervening mechanisms in the literature, not just on economic growth, but on other phenomena of interest, such as civil war and public goods provision. But does macro social structure determine one of or both of these two forms of ethnic practice? If so, we have a more generalizable theory, an important characteristic of causal propositions (Gerring 2005). Thus, my hope is that the cross-national data set presented in this paper promises to revive discussion on the relationship between macro social structure and political and economic outcomes. As we move to multiple dimensions, however, I have emphasized in this paper that crosscuttingness is just one way two (or more) cleavages can be structured in relationship to each other and may not be the most appropriate characteristic for the research question under consideration. Going forward, I urge researchers to think carefully about the use of measures of social structure in their work and as a community of scholars continue to improve upon existing conceptualizations and measures. The data set is publicly available from the ICPSR data archives and the IQSS Dataverse Network.<sup>34</sup>

## References

- Agresti, Alan. 2002. *Categorical data analysis. Wiley series in Probability and Statistics*. New York: Wiley-Interscience.
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2003. Fractionalization. *Journal of Economic Growth* 8:155–94.
- Alesina, Alberto, and Eliana LaFerrara. 2005. Ethnic diversity and economic performance. *Journal of Economic Literature* 63:762–800.
- Annett, Anthony. 2001. *Social fractionalization, political instability, and the size of government* IMF Staff papers Vol. 48 (No. 3).
- Brambor, Thomas, William Clark, and Matt Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14:63–82.
- Budge, Ian, and Cornelius O’Leary. 1971. Cross-cutting cleavages, agreement and compromise: An assessment of three leading hypotheses against Scottish and Northern Irish Survey responses. *Midwest Journal of Political Science* 15:1–30.
- . 1972. Attitudinal and background cross-cutting: Further evidence from Glasgow and Belfast. *Midwest Journal of Political Science* 16:712–22.
- Cederman, Lars-Erik, Jan Ketil Rød, and Nils Weidmann. 2006. *Geo-referencing of ethnic groups: Creating a new dataset*. In GROW Workshop Peace Research Institute, Oslo, Norway (PRIO).
- Chandra, Kanchan, and Steven Wilkinson. 2008. Measuring the effect of “ethnicity”. *Comparative Political Studies* 41:515–63.
- Clark, William, and Matt Golder. 2006. Rehabilitating Duverger’s theory: Testing the mechanical and strategic modifying effects of electoral laws. *Comparative Political Studies* 39:679–708.
- Collier, Paul, and Anke Hoefler. 2000. *Greed and grievance*. Civil War World Bank Policy Research Working Paper No. 2355.
- . 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56:563.
- Deegan-Krause, Kevin. 2007. New Dimensions of Political Cleavage. In *Oxford Handbook of Political Behaviour*, eds. Russell Dalton and Hans-Dieter Klingemann. Oxford: Oxford University Press.
- Easterly, W., and R. Levine. 1997. Africa’s Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112:1203–50.
- Fearon, J. D. 2003. Ethnic and cultural diversity by country. *Journal of Economic Growth* 8:195–222.
- Fearon, J. D., and D. D. Laitin. 2003. Ethnicity, insurgency, and civil war. *American Political Science Review* 97:75–90.
- Gerring, John. 2005. Causation: A unified framework for the social sciences. *Journal of Theoretical Politics* 17:163.
- Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2007. Why does ethnic diversity undermine public goods provision? *American Political Science Review* 101:709–25.
- Kam, Cindy D., and Robert J. Franzese. 2007. *Modeling and interpreting interactive hypotheses in regression analysis*. Ann Arbor, MI: University of Michigan.
- Keefer, Philip. 2005. Democracy, public expenditures, and the poor: Understanding political incentives for providing public services. *The World Bank Research Observer* 20:1–27.

<sup>34</sup>The permanent URL’s are <http://dvn.iq.harvard.edu/dvn/dv/joelselway> and <http://dx.doi.org/10.3886/ICPSR29602>.

- Kuijjs, Louis. 2000. *The impact of ethnic heterogeneity on the quantity and quality of public spending*. IMF Working Paper No. 00/49.
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1999. The Quality of Government. *Journal of Law, Economics, and Organization* 15:222–79.
- Laitin, David, and Daniel Posner. 2001. The implications of constructivism for constructing ethnic fractionalization indices. *APSA-CP* 12.
- Lane, Jan-Erik, and Svante O. Ersson. 1994. *Politics and society in Western Europe*. London: Sage.
- Lipset, Seymour Martin. 1960. *Political man: The social bases of politics*. Garden City, NY: Doubleday.
- Matuszeski, Janina, and Frank Schneider. 2006. *Patterns of ethnic group segregation and civil conflict*. Job Market Paper, Harvard University.
- Montalvo, J. G., and M. Reynal-Querol. 2003. Religious polarization and economic development. *Economics Letters* 80:201–10.
- . 2005a. Ethnic diversity and economic development. *Journal of Development Economics* 76:293–323.
- . 2005b. Ethnic polarization, potential conflict, and civil wars. *American Economic Review* 95:796–816.
- Mozaffar, Shaheen, James R. Scarritt, and Glen Galaich. 2003. Electoral institutions, ethnopolitical cleavages, and party systems in Africa's emerging democracies. *American Political Science Review* 97:379–90.
- Ordeshook, Peter, and Olga Shvetsova. 1994. Ethnic Heterogeneity, district magnitude, and the number of parties. *American Journal of Political Science* 38:100–23.
- Plümpner, Thomas, and Vera E. Troeger. 2007. Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political Analysis* 15:124–39.
- Posner, Daniel N. 2004. Measuring ethnic fractionalization in Africa. *American Journal of Political Science* 48:849–63.
- Rae, Douglas W., and Michael Taylor. 1970. *The analysis of political cleavages*. New Haven, CT: Yale University Press.
- Reynal-Querol, M. 2002. Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution* 46:29–54.
- Rokkan, Stein. 1967. Geography, religion and social class: Cross-cutting cleavages in Norwegian politics. In *Party systems and voter alignments*, eds. S. M. Lipset and S. Rokkan. New York: Free Press.
- Taylor, C., and M. C. Hudson. 1972. *The world handbook of political and social indicators*. 2nd ed. New Haven, CT: Yale University Press.
- Truman, David Bicknell. 1951. *The governmental process: Political interests and public opinion, Borzoi books in political science*. New York: Knopf.
- Wilkinson, Steven. 2000. *Ethnic mobilization and ethnic violence in post independence India*. Paper read at APSA Annual Convention, August 31–September 3, Washington, DC.
- Wimmer, Andreas, Lars-Erik Cederman, and Brian Min. 2009. Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review* 74:316–37.