Advance Access publication March 21, 2016

Political Analysis (2016) 24:211–225
doi:10.1093/pan/mpw002

A General Class of Social Distance Measures

Graham K. Brown
School of Social Sciences, University of Western Australia, Perth, WA 6009, Australia
e-mail: graham.brown@uwa.edu.au (corresponding author)

Arnim Langer
Centre for Research on Peace and Development, KU Leuven, Leuven 3000, Belgium

Edited by Jonathan Katz

Measures of diversity and disparity within a population are used for investigating a range of developmental outcomes, but often by employing "off-the-shelf" indicators that may not be theoretically appropriate for the hypotheses under investigation. In this article, we proposed a general class of social distance measures that both enables us to see the conceptual relationship between different existing measures of heterogeneity more clearly and is sufficiently flexible to allow for the development of tailored hypothesis-specific measures. We show how a range of existing aggregate measures of diversity and disparity fit within the general class and demonstrate illustratively how the measure can be used to develop more precise hypothesis-specific measures.

1 Introduction

In this article, we develop a conceptualization of "social distance" that encompasses dimensions of diversity (ethnicity, religion, and so forth) and disparity (income, education, asset ownership, etc.) and develops a general class of statistical measures to capture this conceptualization of social distance in hypothesis-specific ways. We demonstrate how a wide range of existing measures fit within this general class, and illustratively demonstrate how it can be used to develop new measures for hypothesis-specific testing.

Contemporary econometric analyses of political phenomena such as civil war, public good provision, and regime change are using an increasing range of country-level summary statistics that capture the extent to which the populations in question conform to different demographic and economic distributional patterns (see, e.g., Cederman, Gleditsch, and Buhaug [2011]; Selway [2011b]; and Cederman, Gleditsch, and Buhaug [2013] on civil war; and Lupu and Pontusson [2011] and Selway [2015] on public policy). Standard measures of inequality such as the Gini coefficient that were used in early econometric studies (see, e.g., Auvinen and Nafziger 1999) have been supplemented by studies using measures of "horizontal" inequality—inequality between ethnic or religious groups—and economic polarization (e.g., Østby 2008; Brown 2009; Baldwin and Huber 2010; Cederman, Weidmann, and Gleditsch 2011). Similarly, a relatively well-established measure of ethnic diversity, the Ethnolinguistic Fractionalization Index (ELF)—the default choice for econometric studies of demographic diversity since the pioneering study of ethnic diversity and economic growth in Africa by Easterly and Levine (1997)—has been supplemented by alternative measures of demographic diversity, including the demographic polarization index developed by Montalvo and Reyna-Querol (2005).

Attempts to "measure" ethnicity—whether in purely demographic terms or overlaid with other socio-economic distributional characteristics—have, however, met with considerable resistance from scholars who argue that ethnicity is a subjective, fluid, and essentially political phenomenon and, as such, the attempt to reduce it to a quantitative function of categorical attributes not only

Authors' note: Supplementary materials for this article are available on the *Political Analysis* Web site.

misses the theoretical point but potentially contributes to the kind of reification of ethnic identities that these scholars see as fundamental to the political problems of ethnicity (see, e.g., Kertzer and Arel 2002). Chandra and Wilkinson (2008) distinguish analytically between "ethnic structure" and "ethnic practice." "Ethnic structure," which they define as the "distribution of descent-based attributes" within a population, can be captured by such measures as ELF, while capturing "ethnic practice"—which, they suggest, is often more germaine to political outcomes than ethnic structure—requires the deployment of measures that capture the "act" of ethnicity, such as ethnic voting.

An alternative approach to nuancing the measurement of ethnicity is provided by Brown and Langer (2010), who argue for a conceptual framework in which notions of demographic diversity and socio-economic disparity are taken as sub-concepts within a broader conceptualization of "social distance." Within this framework, they argue that "horizontal" cultural characteristics like "ethnicity" and "religion," along with "vertical" socio-economic attributes such as income, education, and so forth, can be understood as *indicators* of social distance and that the kinds of summary statistics employed in quantitative analyses—whether based on population diversity, economic disparity, or a combination of both—should be interpreted as measures of particular configurations of social distance. The point here is that "ethnicity" in quantitative measurement should be taken as an *indicator* of social distance. Moreover, they concur with Cederman and Girardin (2007) that overcoming the conceptual obstacles to the quantivization of "ethnicity" requires the development of more hypothesis-specific measures, rather than the unquestioning use of generic "off-the-shelf" indicators (see also Kalyvas 2008).

In a parallel development, Bossert, D'Ambrosia, and La Ferrara (2011) propose a Generalized Fractionalization Index (GELF) that seeks to build upon the ELF diversity index by allowing for the incorporation of information about individuals across multiple dimensions, including identity aspects (e.g., ethnicity, religion) and also socio-economic and geographic dimensions. They suggest that their measure effectively overcomes the problem of ethnic salience in quantitative analysis "if one thinks that differences in income, or education, or any other measurable characteristic, may be the reason why ethnicity matters only in certain contexts ... [as] our *GELF* index already weighs ethnic categories by their salience" (Bossert, D'Ambrosia, and La Ferrara 2011, 3).

As we discuss further below, the GELF can be seen as a natural extension of the ELF: whereas the ELF can be intuitively interpreted as the probability that two randomly selected individuals from a population belong to different groups (see below), the GELF can be intuitively interpreted as the average expected level of dissimilarity between two randomly selected individuals or, in Brown and Langer's terminology, the average expected social distance between two randomly selected individuals. Although Bossert et al. provide a useful method of incorporating multiple dimensions into a single index of similarity, their measure is limited in that it is restricted to collapsing this information into an index of fractionalization. Yet, for a more hypothesis-specific approach to measurement, in certain circumstances we might be interested in other constellations of social distance.

In this article, we build upon the approaches of Brown and Langer and Bossert et al. to provide a general class of social distance measures. We show how existing measures of vertical disparity, horizontal diversity, and horizontal inequality can all be seen as particular instantiations of this general class and consider how the general class can be employed to develop more hypothesis-specific measures of social distance.

2 Social Distance

The abstract notion of "social distance" has been employed in a range of social scientific settings that vary subtly but importantly in their focus and interpretation. Within the economics literature, social distance has largely been expressed as a function of the quantity and, to a lesser extent, quality of interactions between individuals. Social distance in this interpretation is primarily important in the emergence of social norms and helps explain suboptimal (from a rational individual perspective) decision-making (see, e.g., Hoffman, McCabe, and Smith 1996; Akerlof 1997). Social distance from this perspective is understood as a characteristic either of individuals—in which case it is often used coterminously with the terminology of "social isolation"—or of dyadic relationship between individuals, albeit with macroscopic consequences. Social distance is hence strongly related to notions of trust, reciprocity, and "bounded" rationality. Broadly speaking, social distance from this perspective helps explain the obdurate refusal of real individuals to behave according to rational predictions (Ortmann and Gigerenzer 2000). In the most stylized but illuminating context, increased social distance is routinely found to correlate negatively with offers in both "laboratory" and real-world experimental dictator games (e.g., Brañas-Garza et al. 2010; Ligon and Schechter 2012; Binzel and Fehr 2013).

An alternative, more sociological, approach to "social distance" interprets it as a characteristic of populations as a whole and seeks to capture at the most abstract level the more intangible elements of heterogeneity in a population than relatively straightforwardly measurable phenomena such as income and education levels. Although often drawing on the individualist scale of social distance pioneered by Bogardus (1925), such studies are concerned with the society-level impacts of intermarriage, occupational mobility, and so forth (see, e.g., Pagnini and Morgan 1990). Studies such as this, however, are still primarily concerned with describing and evaluating social distance within a particular population.

If we are concerned with comparing the impact of aggregate social distance on developmental outcomes, it is useful to have an aggregate measure of social distance, but for different types of outcome (or different hypotheses about the same outcome), we might be interested in different *configurations* of aggregate social distance. It is here that Brown and Langer (2010) make a useful contribution. They argue that the range of measures that exist to capture dimensions of inequality, polarization, and fractionalization can be conceptualized as particular configurations of broader concepts of "diversity" and "disparity," which themselves can be interpreted within a broader notion of "social distance." Their conceptualization is schematically mapped in Fig. 1. This conceptualization "cuts" ethnicity in a slightly different way from Chandra and Wilkinson. Broadly speaking, Brown and Langer's concept of "horizontal diversity" maps onto Chandra and Wilkinson's notion of "ethnic structure." This concept of social distance does not, however, capture the "ethnic practice" dimension of Chandra and Wilkinson's framework, but it does allow for a more nuanced and sophisticated disaggregation of "descent-based" structures with socio-economic distribution. Indeed, the approach we take here, which our general class is designed to exploit, is that various forms of "ethnic action" might be the consequences of different configurations of social distance along both the diversity and disparity axes.

Although Fig. 1 provides a useful conceptual map of social distance, for analytical purposes we can cut the concept of aggregate social distance in a slightly different way. Any aggregate indicator of social distance needs to do two things. First, it needs to provide a metric for the dyadic pairwise comparison of different groups within a given society. Second, it needs to provide a way to aggregate these pairwise comparisons meaningfully into a single society-wide metric. For instance, the well-known Gini coefficient of income inequality, discussed further below, can be interpreted as the mean absolute relative difference in incomes between all the pairwise comparisons in a society. The metric for pairwise comparison here is absolute difference in incomes between each pair of individuals (relative to the overall mean), and because the "groups" in the Gini coefficient are individuals, the aggregation is the simple mean across all those pairings. This is a straightforward example, in part because the income differential in the comparison is so intuitive, and in part because the definition of groups on the individual level makes aggregation straightforward. Where we are interested in groups that constitute varying but significant proportions of the population, aggregation can be more problematic because of the need to weight for group size in the aggregation. And where we are interested in hypotheses that relate to more sophisticated relationship than simple income differentials, the population comparison dimension can be likewise more problematic. The general class we propose provides a systematic but flexible way to build such

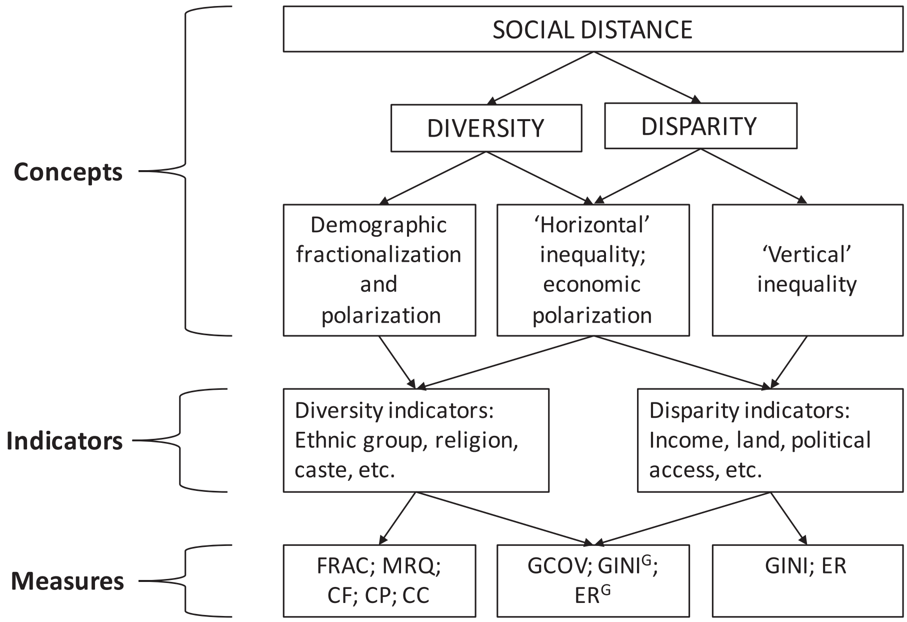

Fig. 1 Conceptual map of social distance.

hypothesis-specific measures based on this analytical decomposition of social distance into pairwise comparison and population aggregation.

3 The General Class

The general class of social distance measures we propose, which we term DIST, is a function of two pairwise matrices that correspond to the analytical distinction above:

- a dissimilarity matrix **D**, which is populated as the pairwise comparison of all groups within society on an dissimilarity function where d_{ij} is the extent of dissimilarity between groups i and j; and
- a population comparison matrix **Q**, which constitutes the aggregation calculus where q_{ij} is an function of the population proportions in groups i and j.

With two normalizing constants k and α (discussed further below), DIST is given by

$$\text{DIST}(k, \alpha, \mathbf{Q}, \mathbf{D}) = \left[k \sum_{i=1}^{m} \sum_{j=1}^{m} q_{ij} d_{ij} \right]^{\alpha}.$$

The DIST measure can be seen in some ways as an extension of the GELF developed by Bossert, D'Ambrosia, and La Ferrara (2011), and it is hence useful to begin with an analysis of that measure to see how DIST operates. The GELF measure is based on the construction of a similarity matrix **S** that compares all possible pairs of individuals for their degree of similarity. The characteristics that are used to measure similarity are not predetermined by the measure, nor is the way that they are computed into a similarity function, subject only to the condition that the similarity score ranges from a minimum of 0 to a maximum of 1. Maximum similarity can be intuitively restricted to situations where the two individuals are identical across all the observed characteristics. Minimum similarity is more problematic where non-categorical characteristics are to be incorporated. In a population of individuals with incomes ranging between 0 and 10 units, we might attribute minimum similarity to comparisons between 0 and 10, but what happens if the population is augmented by another individual with an income of 11? Alternatively, we might think in terms of proportions and attribute minimum similarity to a comparison between an individual who has all of the income in the population against other individuals who have none. But comparing across populations, this implies the same degree of similarity when the individual who has all the income has an income of 10 as when that individual has an income of 100. This is clearly also problematic. Bossert, D'Ambrosia, and La Ferrara (2011) discuss various methods of normalizing to ensure a similarity score within the range 0–1, which need not detain us here.

Where each pairwise similarity comparison between individual i and individual j yields a similarity score s_{ij}, the GELF measure is then defined as

$$\text{GELF} = 1 - \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} s_{ij}.$$

To give an empirical illustration, let us take a hypothetical population of six individuals with observable characteristics of years of education (YoE) and gender as follows:

1. Male, 12 YoE.
2. Male, 12 YoE.
3. Male, 6 YoE.
4. Female, 7 YoE.
5. Female, 12 YoE.
6. Female, 14 YoE.

It will be observed that individuals 1 and 2 are identical across both our observed characteristics; this will be important for what follows. We now need to determine the construction of the similarity function. In their own example, Bossert, D'Ambrosia, and La Ferrara (2011) use principal component analysis, but for illustration purposes here, we can take a simpler approach and additively attribute half of the similarity score to gender (same gender = 0.5; different gender = 0) and half to the absolute difference in YoE, assuming a minimum of 0 and a maximum of 14, that is, $0.5 - (|\text{YoE}_i - \text{YoE}_j|/28)$. This creates a similarity matrix **S** thus:

$$S = \begin{bmatrix} 1.00 & 1.00 & 0.79 & 0.32 & 0.50 & 0.43 \\ 1.00 & 1.00 & 0.79 & 0.32 & 0.50 & 0.43 \\ 0.79 & 0.79 & 1.00 & 0.46 & 0.29 & 0.21 \\ 0.32 & 0.32 & 0.46 & 1.00 & 0.82 & 0.75 \\ 0.50 & 0.50 & 0.29 & 0.82 & 1.00 & 0.93 \\ 0.43 & 0.43 & 0.21 & 0.75 & 0.93 & 1.00 \end{bmatrix}.$$

It can be observed that the leading diagonal is constituted entirely by 1s; every individual is entirely similar to themselves. Likewise, the 2×2 sub-matrix in the top-left corner is constituted entirely by 1s because individuals 1 and 2 are entirely similar to each other (across our observed characteristics). To derive the GELF index from this matrix, we simply take one minus the average cell score within the matrix, which in this case gives us a GELF of 0.36.

The GELF measure is premised on the comparison of individuals, each of whom weighs the same in the process of collapsing the matrix; hence, double summation term in the GELF equation is simply normalized by $1/N^2$. But where, as in this case, we have an internally homogeneous group, it is useful to be able to rewrite the GELF as a group-based measure, which then necessitates taking into account group size. We can do this by defining a group similarity matrix S^G in which internally homogeneous groups rather than individuals are compared pairwise and defining a second "population comparison" matrix **Q** that determines the weight to be given to each groupwise comparison in collapsing the similarity matrix. In a total population of N individuals clustered into m entirely homogeneous groups of n_k individuals ($k = 1, 2, \ldots, m$), each pair of which constitutes proportions

$p_k = n_k/N$ and $p_l = n_l/N$ of the population, $q_{kl} = p_k p_l$. Hence, we can redescribe our population above thus:

$$Q = \begin{bmatrix} 1/9 & 1/18 & 1/18 & 1/18 & 1/18 \\ 1/18 & 1/36 & 1/36 & 1/36 & 1/36 \\ 1/18 & 1/36 & 1/36 & 1/36 & 1/36 \\ 1/18 & 1/36 & 1/36 & 1/36 & 1/36 \\ 1/18 & 1/36 & 1/36 & 1/36 & 1/36 \end{bmatrix} S^G = \begin{bmatrix} 1.00 & 0.79 & 0.32 & 0.50 & 0.43 \\ 0.79 & 1.00 & 0.46 & 0.29 & 0.21 \\ 0.32 & 0.46 & 1.00 & 0.82 & 0.75 \\ 0.50 & 0.29 & 0.82 & 1.00 & 0.93 \\ 0.43 & 0.21 & 0.75 & 0.93 & 1.00 \end{bmatrix}.$$

It is important to note here that the matrices **Q** and **S** are functions of exactly the same set of population characteristics, but whereas **S** is defined as any function (within axiomatic restrictions) of the *values* that those characteristics take, **Q** is a specific function of the *prevalence* of those values within the population. For what follows, it is also useful to change the (group) similarity matrix to create a dissimilarity matrix **D**, whereby $d_{kl} = 1 - s^g_{kl}$. We can now rewrite the GELF index as

$$\text{GELF}^G = \sum_{k=1}^{m} \sum_{l=1}^{m} q_{kl} d_{kl}.$$

The groupwise form of the GELF we have derived here is useful because it forms the basis of the general class of social distance measures we propose. For our general class, however, we loosen three of the restrictions on the GELF^G index. First, for the GELF^G index, the computation of the elements q_{kl} in the population comparison matrix **Q** is restricted to the formula $q_{kl} = p_k p_l$. For our general class of social distance measures, we tolerate other forms of population comparison which, as we shall see, allows for the incorporation of alternative distributional patterns such as polarization. Second, for the GELF^G index, the population characteristics used to construct the matrices **Q** and **D** must be exactly the same. For a general measure of social distance, we may wish to examine the extent of dissimilarity in certain characteristics across groups constituted by other characteristics, for instance "horizontal" income inequality between ethnic groups; in order to do so, we need to allow for **Q** and **D** to be functions of different attributes, in this case ethnicity and income, respectively. Finally, we remove the restriction that s_{ij} and, hence, d_{ij} are bound by the range 0 to 1; instead, we impose the restriction $d_{ij} \geq 0$, that is to say we stipulate a minimum level of dissimilarity (two groups identical across all characteristics) but no maximum level of dissimilarity. This allows for the possibility that the final value of the measure in particular instantiations is not bounded by 0 and 1, as GELF is, but there is no *a priori* reason to impose such a restriction and, as we have seen, defining "minimum" similarity or "maximum" dissimilarity is problematic. Indeed, one of the advantages of basing our measure on dissimilarity rather than similarity is that it is mathematically neater to have a range with a single bound at the minimum (by 0) rather than at the maximum. As we shall see, however, certain instantiations of our general class do impose a maximum bound as well.

Let us turn now, then, to define our general class of social distance measures. Our population is composed of N individuals characterized by a vector of v observed characteristics $\mathbf{c} = (c_1, c_2, \ldots, c_v)$. Components of **c** can be categorical variables—for example, gender, or census-style ethnic categories; discrete variables—for example, YoE or age in years; or continuous—for example, income. We do not place any limitation on the number or nature of these characteristics although, as will become clear, certain subclasses and instantiations of the social distance measure are characterized by such limitations.

The N individuals in the population are clustered into m groups, each of which constitutes proportion p_i of the total population; groups may consist of a single individual. The clustering of individuals into groups is defined according to at least one, though not necessarily all, of the population characteristics. Hence, for instance, if a population is characterized by two attributes, age and gender, we could divide it into groups according to age, according to gender, or according to age and gender. Formally, the vector of characteristics **g** that determines the groups within the population is a subset of **c**. Groups are entirely homogeneous across those characteristics in the subset $\mathbf{g}$—indeed, this constitutes our definition of a "group"—but may vary internally across other characteristics $\{c \notin \mathbf{g}\}$. As above, we do not place further limitations upon the way in which individuals are clustered into groups, except to define particular subclasses of social distance.

The number of groups within our population m is now determinable as the total number of observed combinations of values of the components of $\mathbf{g}$. For each observed combination of characteristic values where c_{wi} is the characteristic value of component c_w $\{c_w \in g\}$ for the ith observed combination, we define a group i whereby the size of the group is given by

$$p_i = \frac{1}{N}\sum_{j=1}^{N}\left\{\begin{matrix} 1 & \text{if}\left\{c_w \in \mathbf{g} | c_{wj} = c_{wi}\right\} \\ 0 & \text{otherwise} \end{matrix}\right\}.$$

Similarly, we define a vector of individual characteristics $\mathbf{h}$ as a subset of $\mathbf{c}$ that constitute those components of $\mathbf{c}$ that are to be used to calculate the social distance *between* pairs of groups. In order to do so, we need to collapse the individual-level vector $\mathbf{h}$ into a group-level equivalent $\mathbf{h^g}$. For any characteristics in the union $c \in \mathbf{g} \cup \mathbf{h}$, this is straightforward, as by definition these characteristics will be unvarying within each group, hence $\mathbf{h^g} = \mathbf{h}$. For other characteristics, however, we will need to stipulate how within-group individual-level variation in the characteristics in $\mathbf{h}$ will be collapsed into the group level $\mathbf{h^g}$. An obvious contender function is the within-group mean, but we might also want to consider the within-group median, or even a more complex function that allows for the construction of hypothesis-specific social distance measures. We hence place no initial restraints upon this function.

We now define a population comparison matrix $\mathbf{Q}$ such that q_{ij} is a function of the proportion of the total population in group p_i and the proportion of the population in group p_j. This matrix defines the way in which population sizes are weighted in the comparison of dissimilarity. Thus, for instance, if a population is divided into two groups which constitute one-third and two-thirds of the population, respectively, and we apply the GELF^G function $q_{ij} = p_i p_j$, we would derive the following population comparison matrix:

$$\mathbf{Q} = \begin{bmatrix} 1/9 & 2/9 \\ 2/9 & 4/9 \end{bmatrix}.$$

In contrast to GELF, however, we allow for alternative formulae. For instance, with the same population we might choose a function $q_{ij} = p_i^2 p_j$, deriving the population comparison matrix:

$$\mathbf{Q} = \begin{bmatrix} 1/27 & 2/27 \\ 4/27 & 8/27 \end{bmatrix}.$$

We then define a dissimilarity matrix $\mathbf{D}$ where d_{ij} is a function of some of the group attributes in $\mathbf{h^g}$. Note that $\mathbf{Q}$ and $\mathbf{D}$ may be functions of the same population attributes or of different population attributes, but $\mathbf{Q}$ is a function of the prevalence of the attributes within the population, whereas $\mathbf{D}$ is a function of the value of the attribute in different groups.

Our generalized class of diversity measures DIST can now be given simply by

$$\text{DIST}(k, \alpha, \mathbf{Q}, \mathbf{D}) = \left[k\sum_{i=1}^{m}\sum_{j=1}^{m} q_{ij}d_{ij}\right]^{\alpha},$$

where k and α are normalizing constants. In what follows, we show how many of the existing measures of social distance can be expressed in this general format and, moreover, can be viewed as logical extensions of each other according to progressive loosening of restrictions upon $\mathbf{Q}$ and $\mathbf{D}$.

4 Existing Measures of Diversity and Disparity in the Generalized Form

In this section, we show how a range of existing measures of diversity and disparity can be formulated as particular instantiations of the generalized social distance measure DIST. Moreover, we show that within this form there is an important symmetry between the specific measures. As mentioned above, within the political economy literature on ethnicity, there are two principal measures used to capture demographic ethnic structure: fractionalization and (demographic) polarization. In the Chandra and Wilkinson (2008) framework, these are both unidimensional measures of ethnic structure. The fractionalization index, which we term here FRAC, is given by the formula

$$\text{FRAC} = 1 - \sum_{i=1}^{n} p_i,$$

where each group $i = 1, 2, \ldots, n$ constitutes proportion p_i of the population. Effectively a Herfindahl concentration index, FRAC gives the probability that two randomly selected individuals belong to different groups. Although the FRAC index has been subject to heavy criticism within as well as beyond the econometric literature, this has largely related to the data initially used by most scholars to compute the index, a heavily outdated and problematic atlas of linguistic diversity compiled by Soviet ethnologists in the 1960s. Attempts at improving the FRAC index have focused on collating more appropriate data (see, e.g., Alesina et al. 2002; Baldwin and Huber 2010) and estimating longitudinal changes in ethnic diversity over time and excluding "politically irrelevant" groups (Posner 2004). But these improvements retain the basic formula of FRAC; they simply improve the data input into the formula.

Montalvo and Reynal-Querol (2005), however, have argued that for certain socio-economic and political outcomes, the patterns of demographic structure picked up by FRAC may not be the most appropriate, theoretically or empirically. They propose an alternative measure of demographic *polarization* that captures the extent to which a population is divided into two equally sized groups. The formula for their measure, which we term here MRQ, is given by

$$\text{MRQ} = 1 - \sum_{i=1}^{n} \left(\frac{0.5 - p_i}{0.5}\right)^2 = 4\sum_{i=1}^{n} p_i^2(1 - p_i).$$

How do these two measures fit within the general class DIST? We can begin with the simplest description of a population within this schema—a population characterized by a single exhaustive and mutually exclusive horizontal characteristic, such as census ethnicity category. This corresponds, broadly, to the "structure" dimension of ethnicity in Chandra and Wilkinson (2008). Clearly, with such a population, **Q** and **D** will be functions of the same characteristic. Moreover, we can initially place the following intuitively reasonable limitations on **D** given the categorical nature of the population attribute:

- Each group is entirely similar to itself: $d_{ij} = 0$ if $i = j$; and
- Each group is entirely dissimilar to all other groups: $d_{ij} = 1$ if $i \neq j$.

This produces a dissimilarity matrix composed entirely of 1s, except with 0s along the leading diagonal; we refer to this for convenience as $\mathbf{D}^*$. (Note that if we defined a similarity matrix rather than dissimilarity, as per Bossert, D'Ambrosia, and La Ferrara [2011], we could equivalently derive for **S** the identity matrix **I**.) FRAC and MRQ fit within the DIST class thus:

- $\text{FRAC} = \text{DIST}(1, 1, p_i p_j, \mathbf{D}^*)$.
- $\text{MRQ} = \text{DIST}(4, 1, p_i^2 p_j, \mathbf{D}^*)$.

These simple measures of horizontal diversity or ethnic "structure" make use of one population characteristic that is assumed to be a mutually exclusive and entire dissimilar categorical attribute. But this is not particularly representative of sophisticated theories of ethnicity and identity politics more generally (Chandra and Wilkinson 2008; Brown and Langer 2010). Even if we define groups as categorical attributes, it is plausible to assert that some pairwise group comparisons are more similar to each other in cultural terms than other comparisons. Within the DIST, this can be achieved by loosening the stipulation that $\mathbf{D} = \mathbf{D}^*$ and allowing d_{ij} to vary between 0 and 1 based on the degree of cultural dissimilarity. When applied to FRAC, this approach results in a measure that was originally proposed by Greenberg (1956) for developing a measure of *linguistic* diversity, but has more recently been resuscitated for use in political science. Using the trees of language similarity that linguists have developed, Greenberg (1956) designates the linguistic "resemblance" between two languages as ρ_{ij} and gives the formula for linguistic diversity as $1-\sum_{i=1}^{N}\sum_{j=1}^{N}p_ip_j\rho_{ij}$.[1] Fearon (2003) and Baldwin and Huber (2010) use linguistic resemblance as a proxy for cultural distance, and use Greenberg's formula for a measure of "cultural fractionalization" (CF). CF fits within the DIST class as $\text{CF}=\text{DIST}(1,1,p_ip_j,1-\rho_{ij})$.

This demonstrates one of the analytical advantages of the DIST class. We have shifted from FRAC to CF by replacing the categorical dissimilarity matrix $\mathbf{D}^*$ with a gradated dissimilarity matrix where $d_{ij}=1-\rho_{ij}$. Within the context of DIST, there is no reason to limit the population comparison matrix to the form p_ip_j; we may wish to calculate a resemblance-based index of "cultural polarization" (CP). By performing the same substitution of the dissimilarity matrix on MRQ as expressed in the DIST format, we can derive this easily as $\text{CP}=\text{DIST}(4,1,p_1^2p_j,1-\rho_{ij})$.

Thus far, then, we have seen that three existing measures of demographic "horizontal diversity" or "ethnic structure"—FRAC, MRQ, and CF—can be expressed within the DIST class and, by extension, have derived a fourth measure of ethnic structure, "cultural polarization." We now turn to measures of vertical disparity. For simple measures of vertical disparity, we are also interested in populations characterized by one characteristic, but this is a discrete or continuous characteristic, such as income or YoE. $\mathbf{Q}$ and $\mathbf{D}$ will hence still be functions of the same attribute, but one that allows for greater variation in the distance matrix D than D^*. Because these measures do not incorporate any cultural or ethnic indicators, they fall outside the Chandra and Wilkinson (2008) schema entirely.

Consider the most common measure of vertical inequality, the Gini coefficient. There are various ways of representing and interpreting the Gini coefficient, but the most useful for our purposes here is as the "relative mean difference"—the average difference in income (education, etc.) between every possible pair of individuals, divided by the overall mean income. This fits within the generalized form as $\text{GINI}=\text{DIST}(1/2,1,p_ip_j,|r_i-r_j|)$, where r_i designates the income (education, etc.) level of group i relative to the overall average.[2] It is worth noting here that in our generalized format, the definition of the population comparison matrix Q in the Gini coefficient GINI and in the demographic fractionalization index FRAC is the same: p_ip_j. Whereas the dissimilarity matrix $\mathbf{D}$ in FRAC is restricted to the binary values in $\mathbf{D}^*$, in GINI it varies according to absolute relative difference. The constant k is halved.

This parallel between the two measures can be given an intuitive as well as mathematical form. One interpretation of the FRAC index is the probability that two randomly selected individuals from a population belong to different groups. In the same way, the relative mean difference interpretation of the Gini coefficient can be restated as the average expected difference between two randomly selected individuals from the population along a continuous attribute. In the sense, the "vertical" GINI can be thought of as a logical extension of the "horizontal" FRAC across a population characterized by one attribute that is continuous rather than categorical.

Given this, is it possible to find an equivalent vertical extension of the horizontal polarization measure MRQ? Unsurprisingly, the answer is yes; and perhaps even more unsurprisingly, that extension is an instance of the economic polarization measure first proposed by Esteban and Ray (1994, 1999; Duclos, Esteban, and Ray 2004), which we term here ER. Derived axiomatically, the ER index is a measure of how far distribution within a population "is grouped into significantly sized clusters such that each group is very similar in terms of the attributes of its members, but different clusters have members with very dissimilar attributes." Their measure is given by the formula

$$\text{ER}=k\sum_{i=1}^{n}\sum_{j=1}^{n}p_i^{1+a}p_j|y_i-y_j|,$$

[1]Greenberg uses r to designate resemblance; we use ρ here, as we use r for an alternative purpose below.
[2]See the Online Supplementary Appendix for full derivation.

for $k > 0$ and $a \in (0, a^*]$ where $a^* \simeq 1.6$. The value of a in the equation gives the degree of "polarization sensitivity", while k is a normalizing constant. (Note that in their formulation, the polarization sensitivity constant is given by α; here, we use a to avoid confusion with the α in our DIST equation.) Clearly, as Esteban and Ray (1994) themselves note, this equation is very similar to the Gini coefficient. Like the Gini coefficient, the measure is invariant to overall population means if the normalizing constant k takes the form $k' \cdot 1/\mu$. For our purposes, it is hence useful to assume that k takes this form and replace the y_is in the equation with their value relative to the mean, r_i. This form of the ER measure fits as a sub class of DIST where $\text{ER}(k, a) = \text{DIST}(k, 1, p_i^{1+a} p_j, |r_i - r_j|)$. From this, it is clear that finding a vertical "extension" of MRQ in an analogous way to the extension of FRAC to find GINI results in $ER(k = 2, a = 1)$. MRQ, it will be remembered, is equivalent to $\text{DIST}(4, 1, p_i^2 p_j, D^*)$. Extending MRQ in an analogous way to the FRAC–GINI link—replacing D^* with $|r_i - r_j|$ and halving the constant k—gives us $\text{ER}(k = 2, a = 1) = \text{DIST}(2, 1, p_i^2 p_j, |r_i - r_j|)$. Note that halving the constant k from 4 to 2 in extending MRQ, as we did in extending FRAC to GINI, also provides exactly the value of k which normalizes $\text{ER}(a = 1)$ into the range 0–1, reaching the maximum 1 where income is perfectly polarized into two evenly sized groups, one with zero income and the other with all the income (or other attribute).

Within our general class of social distance measures $DIST(k, \alpha, \mathbf{Q}, \mathbf{D})$, then, we have defined the following subclasses of measures:

- categorical horizontal diversity measures: Measures in which the population is characterized by a single categorical attribute that is mutually exclusive and entirely dissimilar;
- gradated horizontal diversity measures: Measures in which the population is characterized by a single categorical attribute that is mutually exclusive but characterized by gradated dissimilarity; and
- simple vertical disparity measures: Measures in which the population is characterized by a single continuous or discrete attribute, of which both **Q** and **D** are functions.

We have used two definitions of **Q**:

- $q_{ij} = p_i p_j$, which we can take as a function of population dispersion or fractionalization; and
- $q_{ij} = p_i^2 p_j$, which we can take as a function of population polarization;

and three definitions of **D**:

- a binary categorical definition $\mathbf{D}^*$;
- a gradiated categorical dissimilarity $d_{ij} = 1 - \rho_{ij}$; and
- a variable definition based on absolute difference relative to the population mean $d_{ij} = |r_i - r_j|$.

Each of these six measures corresponds to a different combination of these definitions with appropriate values for the normalizing constants k; α has been set to 1 in all cases.

We have seen, then, that four common measures of horizontal diversity and vertical disparity can be written as particular instantiations of the generalized class of social distance measures DIST and that we can define two broader subclasses of distance to measure simple horizontal diversity, and simple vertical disparity depending upon restrictions placed on the population attributes from which **Q** and **D** are constructed. Increasingly, however, scholars are becoming interested in the political dynamics of group-based inequalities, what Stewart (2000) terms "horizontal inequalities." Within our broader conceptualization of social distance, we are interested here in the subclass of simple measures that we can term "horizontal disparity": those measures that compare the performance of horizontally determined groups across a vertical attribute, such as income or education. This would include both measures of "horizontal inequality" and also other horizontal distributional patterns. Such measures, we should note, fall at the cusp of the Chandra and Wilkinson (2008) distinction between ethnic structure and ethnic action insofar as differences in socio-economic resource among ethnic groups is a good predictor of political outcomes, including ethnic violence.

The subclass of simple measures of "horizontal" diversity can now be easily defined in a similar manner as instantiations of DIST in which a population is characterized by two attributes, one categorical, of which **Q** is a function; and one continuous or discrete, of which **D** is a function. Hence, for instance, we might define a population according to ethnic group and income attributes; simple horizontal inequality measures on this population would be the subclass of measures in which **Q** is a function of the ethnic attribute and **D** a function of the income attribute. Alternative horizontal distributions from horizontal inequality that we might wish to map include, for instance, the notion of horizontal economic polarization.

Mancini, Stewart, and Brown (2008); Baldwin and Huber (2010); and Jayadev and Reddy (2011) discuss a range of possible measures of inter-group inequality. It is trivial to see that two of the common measures of horizontal inequality discussed in this literature fall within the DIST class because their equations are identical to instantiations already discussed. The Gini coefficient can be extended to measure inequality between groups—Mancini, Stewart, and Brown (2008) term this the "Group Gini" (GINI^G); Baldwin and Huber (2010) term the same measure "Between-Group Inequality" (BGI). Likewise, the vertical polarization index ER can be extended to a horizontal, group-based polarization index, ER^G. What is different here is the definition of the population attributes entered into **Q** and **D** rather than the manipulation of the matrices themselves. In the previous vertical examples GINI and ER, we were dealing with a single individual-level continuous variable that determined both the group size and the value to used to calculate d_{ij}, that is to say the vector of individual attributes determining group size and dissimilarity were identical $\mathbf{g} = \mathbf{h} = \{c\}$. We hence did not have to concern ourselves with collapsing $\mathbf{h}$ into $\mathbf{h^g}$. For these horizontal diversity measures, however, we now face a situation where individuals are characterized by two attributes: c_g, which determines their group and is hence used to define **Q**; and c_h, which will be used to determine the level of dissimilarity in **D**. Thus, we need to define the function $\mathbf{h^g} = f(c_g)$. For GINI^G and ER^G, as well as for the Group Coefficient of Variation (GCOV) measure discussed below, Mancini, Stewart, and Brown (2008) use the relative *group* mean, that is to say the mean income (etc.) of each group relative to the overall population mean. We define this here as $\bar{r}_i$. Hence, we can write

$$\text{GINI}^G = \text{DIST}(1/2, 1, p_i p_j, |\bar{r}_i - \bar{r}_j|)$$

and

$$\text{ER}^G = \text{DIST}(2, 1, p_i^2 p_j, |\bar{r}_i - \bar{r}_j|).$$

Mancini, Stewart, and Brown (2008), however, prefer a measure based on the population-weighted coefficient of variation in relative group incomes, which they term GCOV. This measure was originally proposed by Williamson (1965) as a measure of regional inequality, and can be expressed in the DIST class as[3]

$$\text{GCOV} = \text{DIST}(1/2, 1/2, p_i p_j, (\bar{r}_i - \bar{r}_j)^2).$$

Two observations are worth making of this measure. First, it is the first instantiation of DIST we have encountered with a value of α other than 1. Second, it is also the first instantiation of DIST we have encountered that does not have a maximum bound, although it retains the minimum bound of 0 where all groups have identical mean income. The two observations are to an extent linked. Although both the vertical and horizontal Gini indexes and the economic polarization index ER make use of the same distance attribute $\bar{r}_i$—group mean income (YoE, etc.) relative to the overall mean—it is the absolute distance between $\bar{r}_i$ and $\bar{r}_j$ that is taken, and the measure hence effectively bounds itself below 1 because all relative mean incomes above the overall mean will be necessarily exactly matched (proportionate to population size) by other groups with relative mean incomes below the overall mean. Taking the square of the difference in **D**, as GCOV does, violates this; larger distances contribute proportionately more to the measure than smaller distances, and while

[3]See the Online Appendix for the full, fun, algebraic derivation of this equivalence.

Table 1 Relationship between measures of horizontal diversity, vertical disparity, and horizontal disparity

		Population comparison matrix Q	
		Fractionalization: $q_{ij} = p_i p_j$	*Polarization:* $q_{ij} = p_i^2 p_j$
Dissimilarity matrix **D**	Binary categorical: $c_g = c_h$: $d_{ij} = \mathbf{D}^*$	$k = 1, \alpha = 1 \rightarrow$ FRAC	$k = 4, \alpha = 1 \rightarrow$ MRQ
	Gradiated categorical: $d_{ij} = 1 - \rho_{ij}$	$k = 1, \alpha = 1 \rightarrow$ CF	$k = 4, \alpha = 1 \rightarrow$ CP
	Variable characteristic, $c_g = c_h$: $d_{ij} = \lvert r_i - r_j \rvert$	$k = 1/2, \alpha = 1 \rightarrow$ GINI	$k = 2, \alpha = 1 \rightarrow$ ER
	Variable characteristic, $c_g \neq c_h$: $d_{ij} = \lvert \bar{r}_i - \bar{r}_j \rvert$	$k = 1/2, \alpha = 1 \rightarrow \text{GINI}^G$	$k = 2, \alpha = 1 \rightarrow \text{ER}^G$
	Variable characteristic, $c_g \neq c_h$: $d_{ij} = (\bar{r}_i - \bar{r}_j)^2$	$k = 1/2, \alpha = 1/2 \rightarrow$ GCOV	

taking the overall root (hence $\alpha = 1/2$) reduces this effect somewhat, it does not impose a mathematical upper bound.

The remarkable parallelism between the ways the measures discussed are realized in DIST is demonstrated in Table 1. By disaggregating a broad notion of "social distance" into two sets of pairwise comparison—group size in **Q** and group dissimilarity in **D**—the mathematical and logical relationship between these different measures is clarified. At this point, however, it is worth noting that these are not the only permissible definitions of **Q** and **D**, although they are particularly useful in the ease of their intuitive interpretations. The general ER formula, for instance, allows for the varying power a in the equation (distinct from the α in our generalized equation). As Esteban and Ray (1994) note in describing their measure, $ER(k = 0.5, a = 0)$ is equivalent to the Gini coefficient. But the ER index allows for values of a such that $0 \leq a \leq 1.6$. Values of a in the ER formula different from 0 or 1 would also fit into the generalized diversity class DIST through simple redefinitions of the power term in **Q**, with **D** remaining the same. Moreover, the power of this approach is that it makes it easy to derive new measures for hypothesis-specific testing. We have already derived one such novel measure CP by combining the "polarization" configuration of **Q** with the gradiated categorical comparison of CF in **D**. In the following section, we develop, illustratively, two additional measures based on DIST.

5 Extending the Approach

Thus, we have shown how the DIST measure provides a useful framework for examining and understanding the arithmetic relationship between different measures of social distance. The principal advantage of the DIST class, however, is that it allows us to tailor our measures to test specific hypotheses. By way of conclusion, we discuss illustratively some such ways in which it might be utilized.

5.1 *Cross-Cutting Cleavages*

A common auxiliary hypothesis about ethnic diversity and conflict potential is that conflict is less likely where groups have "cross-cutting" cleavages, for instance shared religious affiliation. Selway

(2011a) suggests a general measure of the "cross-cuttingness" (CC) of cleavages within a society based on the χ^2 statistics such that

$$\mathrm{CC} = 1 - \sqrt{\left[\sum \frac{(O-E)^2}{E}\right] / nm},$$

where O and E are, respectively, the observed and expected prevalence of one characteristic (e.g., religion) among another (e.g., ethnic group) and nm is a normalizing factor for comparison of different structures. We have not been able to find an exact mathematical equivalent for CC in DIST, but the advantage of DIST is that we can easily derive a similar measure tailored to a specific hypothesis.[4] Let us assume, as above, that we are interested in ethnic diversity mediated by the degree of religious overlap between ethnic groups. We can begin from FRAC, our simplest measure of ethnic diversity. In FRAC, the pairwise comparison in the dissimilarity matrix D takes the $D*$ form $d_{ij} = 1$ if $i \neq j$. How might we adjust d_{ij} to incorporate the degree of religious overlap between i and j? In order to populate the matrix **D**, we need a *dyadic* measure that captures the degree of religious overlap between each pair of ethnic groups. Brown (2009) proposes one such measure for examining the degree of ethnic overlap between geographical regions, which he terms the Ethnic Difference Measure; as we are interested in religious overlap between pairs of ethnic groups, we term it here the Religious Difference Measure (RDM). Where two ethnic groups i and j are divided into m religious groups ($k = 1, 2, \ldots, m$), which constitute proportion π_{ik} of group i and π_{jk} of group j, the measure can be given as

$$\mathrm{RDM}_{ij} = \frac{1}{2}\sum_{k=1}^{m} |\pi_{ik} - \pi_{jk}|.$$

Intuitively, the measure gives the proportion of one (either) ethnic group that would have to convert religious affiliation in order for the overall affiliation of both groups to match. We can now define an modified instantiation of the demographic fractionalization index with the dissimilarity matrix **D** populated by the equation $d_{ij} = \mathrm{RDM}_{ij}$, giving us DIST $(1,1,p_i,p_j, \mathrm{RDM}_{ij})$. Intuitively, this is a measure of ethnic demographic fractionalization mediated by religious overlap. Where a group is compared against itself ($i = j$), RDM_{ij} will be, by definition, 0. In other pairwise comparisons ($i \neq j$), RDM_{ij} will tend toward 1 the less religious overlap there is between the two groups. Hence, in the extreme case where no pairwise comparison has any degree of religious overlap, this measure will collapse into FRAC with $\mathrm{RDM}_{ij} = 1$ for all $i \neq j$. The power of the DIST measure is that we need not stop there, however. It would be possible to follow the same logic of religious overlap to "mediate" the ethnic polarization index MRQ with $\mathrm{DIST}(4, 1, p_i^2 p_j, \mathrm{RDM}_{ij})$. This instantiation reaches its maximum value 1 in contexts where a population is split into two evenly sized ethnic groups professing entirely different religions and tends toward zero both in ethnic distributions that tend toward zero in the MRQ index (one numerically dominant group; or many small groups); and in more ethnically polarized contexts that have high levels of shared religious affiliation.

5.2 *Within-Group Distribution*

A second way that DIST could be employed for more sophisticated hypotheses is to build on the "horizontal inequality" measures discussed above to incorporate within-group distributional characteristics. As already mentioned, a range of studies have examined the impact of "horizontal inequality" on ethnic conflict, and have found it a significant predictor of violent conflict utilizing a variety of different measures, including a "shortfall" measure of regional performance (Gates and Murshed 2005); the absolute income ratio between the largest two groups (Østby 2008); and the

[4]Because the CC measure is normalized around $1/nm$, it does not easily fit within the DIST structure. Selway (2011a) also discusses an earlier measure proposed by Taylor and Rae (1969). This measure would potentially fit within the DIST scheme.

GCOV measure discussed above (Mancini 2008). Each of these measures, however, only compares the *average* socio-economic performance of each group, whether against the national average (Gates and Murshed 2005); one other large group (Østby 2008); a logarithmized pairwise comparison (Cederman, Weidmann, and Gleditsch 2011); or a weighted comparison of all groups (Mancini 2008). This tells us nothing about *within-group* distribution beyond the simple mean.

Plausibly, however, the aggregate social distance *between* groups will be affected by the disparity *within* those groups. Horowitz (1985), for instance, is not concerned with mean difference between groups *per se*, but with how far a society is ethnically "ranked." One quantitative interpretation of this would be to measure the degree of *overlap* between the income distribution curves of the different groups. The logic here is that the less overlap there is between income distributions across groups, the more "ranked" that society is. The Bhattacharyya Coefficient (BC) provides a good approximation of overlapping distributions. Dividing the overall distribution into n quantiles, the coefficient is given by $\mathrm{BC} = \sum_{q=1}^{n} \sqrt{i_q j_q}$, where i_q and j_q are, respectively, the proportion of groups i and j in the qth quantile. The range of BC is 0, where there is no overlap between the two distributions, to 1, where there is complete overlap in distributions.

As with the religious overlap measure used above, we can now use this measure to "mediate" measures of demographic polarization and fractionalization with the extent of income overlap (rather than mean difference) between each pair of groups. Because BC increases with the degree of similarity, we need to invert it for the dissimilarity matrix $\mathbf{D}$ such that $d_{ij} = 1 - BC_{ij}$. Once again, however, the flexibility of DIST allows us to compile this with different population comparison matrices $\mathbf{Q}$. Hence, for instance, if we take the population comparison format for demographic fractionalization $q_{ij} = p_i p_j$, then we get $\mathrm{DIST}(1, 1, p_i p_j, 1 - BC)$. Because BC is bounded by 0 and 1 and is by definition equal to 0 where $i = j$, it can be seen that the maximum value this instantiation of DIST takes is the equivalent demographic fractionalization index for the same population (i.e., $\mathrm{DIST}(1, 1, p_i p_j, D^*)$) and that it will approach this value the less overlap there is in income distribution (or other characteristic) between groups, that is, as BC approaches 0. Hence, we have derived a measure of demographic fractionalization "mediated" by income overlap. By extension, however, it would also be possible to generate an equivalent measure of demographic polarization mediated by income overlap as $\mathrm{DIST}(4, 1, p_i^2 p_j, 1 - BC)$.

6 Conclusion

Measures of diversity and disparity within a population are used for investigating a range of developmental outcomes, but often by employing "off-the-shelf" indicators that may not be theoretically appropriate for the hypotheses under investigation. In this article, we have suggested that a broad concept of aggregate social distance provides a useful frame for relating these different aspects of population heterogeneity both conceptually and algebraically. We proposed a general class of measures DIST that both enables us to see the conceptual relationship between different existing measures of heterogeneity more clearly and is sufficiently flexible to allow for the development of tailored hypothesis-specific measures. Clearly, hypothesis-specific measures are likely to be "data-hungry" and the usefulness of DIST may, in that sense, be restricted by the level and nature of data availability. As data availability continues to improve, however, DIST provides a powerful way to explore the developmental impacts of heterogeneity with greater precision and flexibility.

Conflict of interest statement. None declared.

References

Akerlof, G. A. 1997. Social distance and social decisions. *Econometrica* 65:1005–27.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2002. *Fractionalization*. NBER Working Paper No. 9411.

Auvinen, J., and E. W. Nafziger. 1999. The sources of humanitarian emergencies. *Journal of Conflict Resolution* 43:267–90.

Baldwin, K., and J. D. Huber. 2010. Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *American Political Science Review* 104:644–62.
Binzel, C., and D. Fehr. 2013. Social distance and trust: Experimental evidence from a slum in Cairo. *Journal of Development Economics* 1:99–106.
Bogardus, E. S. 1925. Measuring social distance. *Journal of Applied Sociology* 9(2):299–308.
Bossert, W., C. D'Ambrosia, and E. La Ferrara. 2011. A generalized index of fractionalization. *Economica* 78:723–50.
Brañas-Garza, P., R. Cobe-Reyes, M. Paz Espinosa, N. Jiménez, J. Kovarik, and G. Ponti. 2010. Altruism and social integration. *Games and Economic Behavior* 69:249–57.
Brown, G. K. 2009. Regional autonomy, spatial disparity, and ethnoregional protest in contemporary democracies: A panel data analysis, 1985–2003. *Ethnopolitics* 8:47–66.
Brown, G. K., and A. Langer. 2010. Conceptualizing and measuring ethnicity. *Oxford Development Studies* 38(4):411–36.
Cederman, L.-E., and L. Girardin. 2007. Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies. *American Political Science Review* 101:173–85.
Cederman, L.-E., K. S. Gleditsch, and H. Buhaug. 2013. *Inequality, grievances, and civil war*. Cambridge: Cambridge University Press.
Cederman, L.-E., N. B. Weidmann, and K. S. Gleditsch. 2011. Horizontal inequalities and ethnonationalist civil war: A global comparison. *American Political Science Review* 104:478–95.
Chandra, K., and S. Wilkinson. 2008. Measuring the effect of "ethnicity". *Comparative Political Studies* 41:515–63.
Duclos, J.-Y., J. M. Esteban, and D. Ray. 2004. Polarization: Concepts, measurement, estimation. *Econometrica* 72:1737–72.
Easterly, W., and R. Levine. 1997. Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112:1203–50.
Esteban, J., and D. Ray. 1994. On the measurement of polarisation. *Econometrica* 62:819–51.
———. 1999. Conflict and distribution. *Journal of Economic Theory* 87:379–415.
Fearon, J. D. 2003. Ethnic structure and cultural diversity around the world: A cross-national dataset on ethnic groups. Department of Political Science, Stanford University.
Gates, S. M., and M. S. Murshed. 2005. Spatial-horizontal inequality and the Maoist insurgency in Nepal. *Review of Development Economics* 9(1):121–34.
Greenberg, J. H. 1956. The measurement of linguistic diversity. *Language* 32:109–15.
Hoffman, E., K. McCabe, and V. L. Smith. 1996. Social distance and other-regarding behavior in dictator games. *American Economic Review* 86:653–60.
Horowitz, D. L. 1985. *Ethnic groups in conflict*. Berkeley: University of California Press.
Jayadev, A., and S. G. Reddy. 2011. Inequalities between groups: Theory and empirics. *World Development* 39:159–73.
Kalyvas, S. N. 2008. Promises and pitfalls of an emerging research program: The microdynamics of civil war. In *Order, conflict, and violence*, eds. S. N. Kalyvas, I. Shapiro, and T. Masoud. Cambridge: Cambridge University Press pp. 1–14.
Kertzer, D. I., and D. Arel, eds. 2002. *Census and identity: The politics of race, ethnicity, and language in national censuses*. Cambridge: Cambridge University Press.
Ligon, E., and L. Schechter. 2012. Motives for sharing in social networks. *Journal of Development Economics* 99:13–26.
Lupu, N., and J. Pontusson. 2011. The structure of inequality and the politics of redistribution. *American Political Science Review* 105:316–36.
Mancini, L. 2008. Horizontal inequality and communal violence: Evidence from Indonesian districts. In *Horizontal inequality and conflict: Understanding group violence in multiethnic societies*, ed. Stewart F. Basingstoke: Palgrave pp. 106–135.
Mancini, L., F. Stewart, and G. K. Brown. 2008. Approaches to the measurement of horizontal inequalities. In *Horizontal inequalities and conflict: Understanding violent group mobilization in multiethnic societies*, ed. Stewart F.. Basingstoke: Palgrave pp. 85–105.
Montalvo, J., and M. Reyna-Querol. 2005. Ethnic diversity and economic development. *Journal of Development Economics* 76:293–323.
Ortmann, A., and G. Gigerenzer. 2000. Reasoning in economics and psychology: Why social context matters. In *Cognition, rationality, and institutions*, eds. M. E. Streit, U. Mummert, and D. Kiwit, 131–45. Berlin: Springer.
Østby, G. 2008. Polarization, horizontal inequalities, and violent civil conflict. *Journal of Peace Research* 45:143–62.
Pagnini, D. L., and S. P. Morgan. 1990. Intermarriage and social distance among U.S. immigrants at the turn of the century. *American Journal of Sociology* 96(2):405–32.
Posner, D. L. 2004. Measuring ethnic fractionalization in Africa. *American Political Science Review* 48:849–63.
Selway, J. S. 2011a. The measurement of cross-cutting cleavages and other multidimensional cleavage structures. *Political Analysis* 19:48–65.
——— 2011b. Cross-cuttingness, cleavage structure, and civil war onset. *British Journal of Political Science* 41:111–38.
——— 2015. *Coalitions of the well-being: How electoral rules and ethnic politics shape health policy in developing countries*. Cambridge: Cambridge University Press.
Stewart, F. 2000. Crisis prevention: Tackling horizontal inequalities. *Oxford Development Studies* 28:245–62.
Taylor, M., and D. Rae. 1969. An analysis of crosscutting between political cleavages. *Comparative Politics* 1:534–47.
Williamson, J. G. 1965. Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change* 13:1–84.