JICA Research Institute

JICA-RI Working Paper

Ethnic Diversity and Economic Instability in Africa:
Policies for Harmonious Development

Conceptualizing and Measuring Ethnicity

Graham K. Brown and Arnim Langer

No. 9

March 2010

JICA Research Institute

JICA Research Institute

JICA Research Institute
10-5 Ichigaya Honmura-cho
Shinjuku-ku
Tokyo 162-8433 JAPAN
TEL: +81-3-3269-3374
FAX: +81-3-3269-2054

Conceptualizing and measuring ethnicity

K. Brown[*] and Arnim Langer[**]

Abstract

In this paper, we examine the ways in which the notion of 'ethnicity' has been operationalized for quantitative research. We argue that there is a conceptual mismatch between modern theories of ethnicity and the current way in which ethnicity is employed statistically, but that this mismatch can be overcome if we reconceptualise econometrics as using 'ethnicity' as an indicator of a broader notion of 'social distance'. We conceptually subdivide 'social distance' into measures of 'diversity' and 'disparity', and review the strengths and weaknesses of existing measures employed in the literature. We then explore how far different measures of ethnicity are correlated both across stylized distributions and using real world data from districts in Ghana, Uganda, and Indonesia. We find that even measures that purport to pick up very different underlying distributions are, in fact, very highly correlated, and suggest that this should caution us to be more careful in the interpretation of econometric results. We conclude that quantitative research using theoretically more sophisticated measures of ethnicity that link particular distributions with particular political outcomes, combined with qualitative research that allows a closely examination of causal processes, may be the appropriate way forward for empirical research on ethnicity.

Keywords: ethnicity, measurement, methodology, inequality, social distance

[*] Centre for Development Studies, University of Bath (g.k.brown@bath.ac.uk)
[**]Oxford Department of International Development, University of Oxford (arnim.langer@qeh.ox.ac.uk)

Introduction

Although widely studied by anthropologists and political scientists since the late 1960s, ethnicity and by extension ethnic groups and ethnic diversity remain highly contested concepts. Despite the 'slippery' (Fearon 2003, 195) and 'problematic nature of ethnicity' (Nagel 1994, 154), an increasing number of political scientists and economists have started to incorporate ethnic groups into their (empirical) analyses or use ethnic diversity as an independent variable for explaining a range of social and economic outcomes, most notably relating to violent group mobilization, secessionist movements and economic growth differences (see, for example: Collier and Hoeffler 1998; Easterly 1997; Rodrik 1999; Sambanis 1999). However, other scholars have remained sceptical of using ethnicity as an explanatory variable. Martin Doornbos has strikingly captured this position as follows: 'Ethnicity does not explain anything, it needs to be explained' (Doornbos 1991, 19). While we take the view that ethnic groups and ethnic diversity can be important explanatory variables with regard to a range of social and economic phenomena, the concept of ethnicity faces a number of conceptual, measurement and practical problems and limitations which can potentially have a major impact on the validity of the research results. The paper aims to discuss these obstacles and propose different ways in order to address or at least minimize their potentially negative impact on research outcomes and results.

The paper proceeds as follows. In Section 1 we will discuss different definitions and approaches to ethnicity. Section 2 discusses how ethnicity is generally operationalized in the quantitative literature, explores the tensions between these approaches and the theoretical perspectives on ethnicity highlighted in Section 1, and proposes an alternative conceptual framework for quantitative analyses which, we suggest, substantially overcomes these tensions. Section 3 presents and discusses a range of measures that can be used to quantify and compare the degree of ethnic diversity and disparity across different countries and societal contexts. Section 4 then explores the relationship between some of these measures of ethnic diversity

and disparity in three countries, Ghana, Indonesia, and Uganda. It is worth noting here that the objective of this exercise is not to test a particular hypothesis, but rather to show how different diversity measures can have different results, which may lead to different –and sometimes contradictory- conclusions. In the final section, we will draw some conclusions with regard to the usability and limitations of ethnicity and ethnic diversity as explanatory variables in social and economic research, with a particular focus on policy relevance.

1. What is 'ethnicity'?

While there is no universally accepted definition, ethnicity is generally characterized as a sense of group belonging, based on ideas of common origins, history, culture, language, experience and values (see, for example: Anderson 1983; Bates 2004; Glazer and Moynihan 1976; Horowitz 1985; Varshney 2001). In addition to real or putative common descent, most definitions of ethnicity emphasize 'the sharing of a "culture", the most notable aspect of which is language' (Bates 2004, 5). Three analytical approaches have emerged from the debates on the character of ethnicity: primordialism, instrumentalism and constructivism.

There appears to be a clear distinction between the ways in which people in non-scholarly and scholarly circles think about ethnicity and ethnic differences. The view that is popular in non-academic circles has been termed 'primordialism'. The primordialist view emphasizes the 'affective properties of ethnicity' and asserts that people's ethnic consciousness is 'deeply embedded in the constitution of the self' (Young 2003, 13). Ethnicity is regarded as a *natural* result of biological differences or a long historical process (see, in particular, Van den Berghe 1978). Primordialists generally claim that ethnicity is unchanging and unique. While newspapers and magazines as well as politicians often take a primordialist view in order to explain the emergence of ethnic conflicts, among academics this is increasingly discredited. Consequently, as Kanchan Chandra notes, 'it is now virtually

impossible to find a social scientist who *openly* defends a primordialist position' (Chandra 2001, 8; emphasis added).

In the academic literature, the primordialist view is usually associated with Edward Shils and Clifford Geertz (see Geertz 1963; Shils 1957). Yet, as Robin Cohen argues, Shils and Geertz's 'primordial' argument should not be equated with the political *practice* of primordialism and is hence rather more sophisticated and difficult to refute than might appear at first glance (see Cohen 1999). Geertz understood primordialism thus: 'By a primordial attachment is meant one that stems from the 'givens' of social existence … congruities of blood, speech, custom and so on, *are seen* to have an ineffable, and at times overpowering, coerciveness in and of themselves. One is bound to one's kinsman, one neighbour, one fellow believer, *ipso facto* [by that very fact]; as the result not merely of personal affection, practical necessity, common interest, or incurred obligation, but at least in great part by virtue of some unaccountable absolute import attributed to the very tie itself' (Geertz 1963 quoted in Cohen 1999, 4, emphasis added). Thus, as Cohen points out, Clifford Geertz does not assert (as is often argued by anti-primordialists) that common blood, kinship and beliefs are primordial givens of social existence, but instead he argues that these things 'are seen' to be primordial.

From the beginning of the 1970s, the primordialist view of ethnicity was challenged by so-called 'instrumentalists' or 'functionalists'. The instrumentalists' main criticism focused on the primordialists' inability to see that certain episodes of apparently ancient hatred were 'selectively retrieved by the knowledge elite, ignoring the many instances of cooperation and coexistence' (Varshney 2001, 4812). Moreover, most ethnic groups interact in a cooperative and peaceful rather than violent way (Fearon and Laitin 1996). Another weakness of the primordial view is its inability to explain why ethnic groups change over time.

The instrumentalist view emphasizes the importance of ethnicity and ethnic affiliations in social, political and economic competition. Instrumentalists particularly focus on the role of elites in this respect. According to instrumentalists, ethnicity is a *resource* used by

elites to define group identity, regulate group membership and boundaries, and make claims and extract (state) resources (see, for example: Brass 1985; Glazer and Moynihan 1976; Hardin 1995; Rabushka and Shepsle 1972). Another reason why ethnic elites sometimes use and accentuate ethnic differences is 'to transfer potential hostility from inequalities and power disparities *within* their communities to the elites and subjects of other communities' (Nafziger and Auvinen 2002, 159, emphasis added).

Since the early-1980s, the constructivist perspective has probably become the most influential perspective in the study of ethnicity. In contrast to primordialists, constructivists tend to focus on the processes through which ethnic groups have emerged and have gained social significance (Ukiwo 2005). However, as Robert Bates argues the constructivist approach to ethnicity combines elements of both the primordial and instrumentalist approaches: 'In keeping with the primordialists, constructivists view ethnic identities as a cultural endowment; but in keeping with instrumentalists, they view ethnic identities as malleable. Distinguishing their position is the belief that while identities can be reshaped, they can be altered only at significant cost' (Bates 2004, 5).

In the process of identity construction, a crucial role is played by so-called 'cultural entrepreneurs who codify and standardize a language, equipping it with a written form, create an ethnos-centred historical narrative, populated with internal heroes and external villains, and build a literary tradition' (Young 2003, 14). In many African countries, the colonial regimes played an important role in the 'promotion', 'systematisation' and in some cases the actual 'invention' of ethnic groups and identities (see Ranger 1983; Young 1985).

While ethnic identities once constructed tend to endure and can arouse deep attachments from the people involved, the boundaries of ethnic groups remain fluid and situational. As Ralph Premdas puts it, 'the maintenance of boundaries is situationally determine, may shift over time and context, and generally serves to differentiate members dialectically and oppositionally from other groups in terms of "we-they" antipathies' (Premdas

1995, 4). However, the 'delineation' of ethnic groups is often quite arbitrary and can be the result of external processes or decisions taken by people who are not part of a group. Ukoha Ukiwo notes that 'ethnic groups may be "objectified" by recognition and classification by the state, religious institutions and the intelligentsia as well as by self-identification and recognition by others' (Ukiwo 2006, 21).

2. How do we operationalize 'ethnicity'?

There is, then, a clear academic consensus that the extreme primordialist position that views ethnic affiliation as fixed as unchanging is no longer tenable, if indeed it ever was thought to be so. What is less clear is the extent to which this renders quantitative inquiries that utilize 'measures' of ethnicity problematic. This debate is rendered more difficult to navigate because it often falls along disciplinary and methodological lines. Opposition to 'measuring' ethnicity is often bundled together with a broader opposition to quantitative methods and to positivist epistemologies in social science. In introducing their insightful collected volume examining the politics of census classifications from an anthropological perspective, for instance, David Kertzer and Dominique Arel (2002, 19) assert that the 'notion that cultural categories can be reduced to an objective core... is dangerously close to the primordialist notion of timeless identities'.

It is clearly beyond the scope of this paper to resolve these larger disputes, but before proceeding to discuss different specific measures of ethnicity, in this section we defend the proposition that while the any measurement of ethnicity diversity for quantitative inquiry may be *practically difficult*, it is not *conceptually impossible*. We can draw a parallel here with the quantitative measurement of democracy. As Gary Goertz (2006) discusses extensively, there are a range of measures of democracy available for quantitative inquiry, but what is important from a methodological perspective is what he terms 'concept-measure consistency'. Thus, for instance, while the commonly polity dataset may be a useful measure of democracy if we

concerned with an conceptualization of democracy that places emphasis on party competitiveness and balance of powers, it is less useful as a measure if we are concerned with the emancipatory dimension of democracy. This is because this latter dimension is not well coded in the polity dataset – Switzerland, in the polity dataset, is coded with a full 10 points on democracy since 1848, despite not granting women the vote at the federal level until 1971.

From this perspective, the charge against quantitative measures of ethnicity is that they are employing data (usually census or survey responses) that are 'dangerously close' to primordialism to measure a concept ('ethnicity') that is generally agreed to be socially constructed – there is not good concept-measure consistency, in Goertz' terminology. Our claim in this section is that we *can* reconcile the concept and measurement in a consistent way, but that it requires critical and explicit attention to this relationship. In fact, such critical attention is present to some degree in the contemporary quantitative literature. In this section we first review the steps that have been taken to improve concept-measurement consistency in dealing with ethnicity and then proceed to make a more general argument about how improve this consistency.

Until quite recently, these problems have been insufficiently acknowledged and addressed in most quantitative studies focusing on the implications of ethnic diversity on different social and economic outcomes. An important issue that illustrates this lack of methodological rigour is the fact that many scholars who have analysed statistically the relationship between ethnic diversity and violent conflict or economic growth have failed to define the concept of ethnicity, even though they were generally quite careful in defining other terms and concepts (Green 2004). Thus, for instance, William Easterly and Ross Levine, in their influential article *Africa's Growth Tragedy: Policies and Ethnic Divisions*, offer no definition of ethnicity and ethnic diversity is simply 'operationalized' by using the Ethno-Linguistic Fractionalization (ELF) dataset (see section 0 below) (see Easterly and Levine 1997). Similarly, in Collier and Hoeffler's quantitative studies of violent conflict and civil war,

ethnicity also remains undefined and again the ELF index is used to 'operationalize' ethnic diversity (see Collier and Hoeffler 1998; Collier and Hoeffler 2002). In both these studies the ELF index is only discussed in passing, without linking the composition of the dataset to a particular perspective on ethnicity. Yet, as Elliot Green rightly notes, 'while many authors do not offer their own definition of ethnicity, they do, however, draw upon data sets of ethnic groups that needed to be based upon some objective criteria' (Green 2004, 16).

While some scholars ignore the definitional aspects of the concept of ethnicity altogether, others pay lip-service to a constructivist approach to ethnicity. Rogers Brubaker has argued in this respect that while the mainstream scholarship on ethnic politics has by and large moved away from the kind of ethnic primordialism associated with the works of Clifford Geertz this has resulted nonetheless in the kind of '*clichéd constructivism*' in which broadly constructivist ontological pronouncements are followed by 'groupist' empirical analysis, which take the ethnic group as an essentially primordial given (Brubaker and Cooper 2000). While Brubaker's critique is valid and very important, we argue that his solution -i.e. that we should analyze 'ethnicity without groups' and focus on the social *processes* of ethnicization, social mobilization and organization, and *identification* (rather than 'identity') seems to throw out the child with the bathwater.

While quantitative studies of ethnic diversity are inherently problematic because this requires the reduction of ethnicity into exhaustive and mutually exclusive ethnic groups (something sophisticated theories of ethnicity militate against), as long as the interpretation of results is cognizant of the limitations of this kind of categorization, quantitative analysis can provide a useful systematic form of comparison. Indeed, within the econometric and quantitative literature on ethnic conflict, there appears to be a shift toward taking such concerns into consideration through more nuanced 'measures' of ethnicity (Posner 2004). There are a number of other problems which make the construction of a cross-national dataset of ethnic groups an inherently tricky exercise and complicate the interpretation of significant

correlations between a particular measure of ethnic diversity (based on these datasets) and some other 'dependent' variable (like the emergence of political violence and positive economic growth). The two main problems relate to the following questions: 1) how can we identify the (main) ethnic groups of a country? and 2) how do we deal with changes in ethnic demography over time?

With regard to the first problem, James Fearon has emphasized the difficulty of coding ethnic groups and finding the 'right' list of ethnic groups: 'Anyone with primordialist leanings should be quickly disabused of them by undertaking to code "ethnic groups" in many different countries. It rapidly becomes clear that one must make all manner of borderline-arbitrary decisions, and that in many cases there is no single right answer to the question "What are the ethnic groups in this country?" Constructivist or instrumentalist arguments about the contingent, fuzzy, and situational character of ethnicity seem amply supported' (Fearon 2003, 197). Take, for example, Ghana. What is the most appropriate ethnic classification for this country? Should we distinguish between the four major ethnolinguistic groups (i.e. the Akan, Ewe, Ga-Dangmes and Mole-Dagbanis) or should we also take into account sub-divisions within these groups (the Akan, for example, consist of twenty smaller ethnic sub-groups of which the Ashantis and Fantis are the demographically the most important ones)? Clearly, the approach we take in this respect will have a major impact on the 'measured' level of ethnic diversity, and consequently may have an important affect on the relationships we aim to test. Moreover, considering that there might be more than one 'plausible' ethnic classificatory scheme in certain countries, 'we must be careful that we do not, in effect, choose the coding that best supports our theory, after the fact' (ibid., 197).

While scholars like James Fearon have addressed a pragmatic problem in 'listing' ethnic groups, however, there remains a concern whatever 'level' one codes at, the ethnic groups that are coded for are those which have already been bestowed *political* significance through (usually, in the developing world) colonial processes of identity formation,

bureaucratisation and census definition. The charge here is one of selection on the dependent variable – in using existing 'bureaucratic identities' to code ethnicity, we are using those identities that are already politicized to some degree.

With regard to both this issue and the issue of demographic changes over time, Daniel Posner identifies two different kinds of ethnic demographic changes; the first one relates to a change in the salient 'ethnic cleavage', while the second one involves 'changes in the number and relative sizes of groups' (Posner 2005, 2-3). With respect to the first type of ethnic demographic change, it is important to note that 'societies contain multiple dimensions of ethnic cleavage, each of which can be salient in different settings and historical moments' (ibid., 2-3). India's 'ethnic' landscape, for example, looks quite different when defined in terms of ethnic, religious, language or caste differences (ibid.). With respect to the second type of ethnic demographic change, it is important to note that changes in the number and relative sizes of ethnic groups can be the result of differences in natural growth patterns across groups, 'internal redefinitions' (as with the redefinition of Isaaqs and Hawiyes as separate ethnic groups in Somalia) and migration flows (ibid.). The way one deals with these issues (for example, whether one decides to enumerate groups on the basis of language, religion or caste differences) will again have a major impact on the 'measured' degree of ethnic diversity in a particular country.

Scholars like Fearon and Posner, then, are usefully problematizing the quantification of ethnicity and, in doing so, contributing towards bridging the concept-measurement divide. Here, however, we wish to propose a more general claim about the way in which we can conceptualize quantitative enquiries into 'ethnicity' in order to accommodate both the criticisms of data primordialism and the existing concept-measurement divide. Our claim is that in quantitative analysis of 'ethnicity', the concepts we are dealing with should *not* be conceived of as 'ethnicity' per se, but as 'diversity' and 'disparity' – which can also usefully be seen together second-level conceptual dimension of a broader top-level concept of 'social

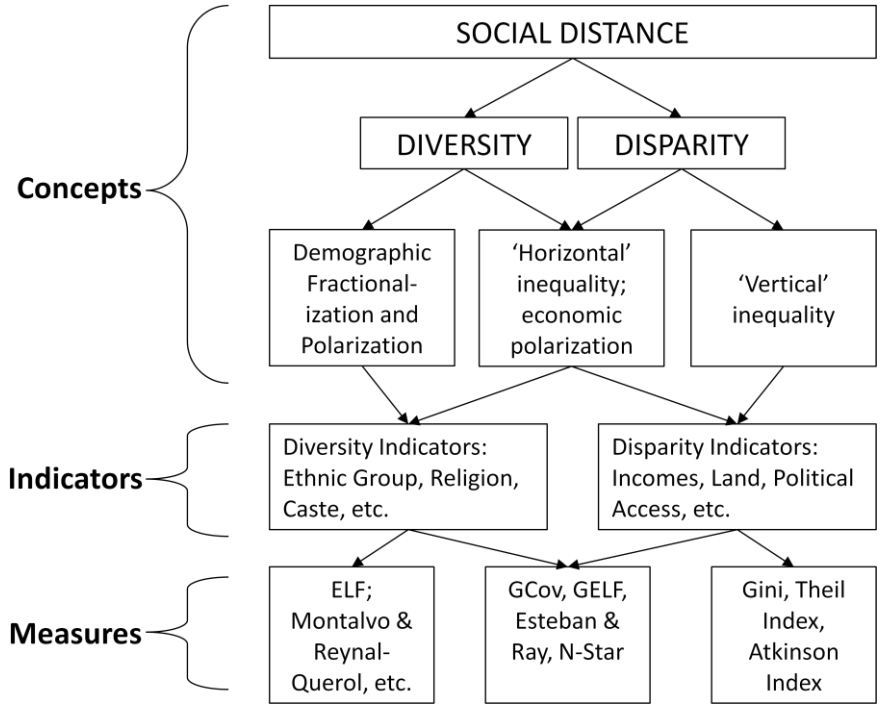
distance'. This perspective that the concepts we are interested in quantitatively are not 'ethnicity' or 'religion' but more abstract notions of social distance in fact more-or-less explicit or implicit in many of the proposed measures we discuss in the next section which are concerned with concepts such as 'fractionalization' and 'polarization'. These latter concepts can in turn be seen as third level conceptualization of particular *forms* of social distance.

If the concepts we are dealing with quantitatively can be brought together under the broad notion of 'social distance', what role *does* 'ethnicity' play? Our suggestion is that in this scheme 'ethnicity' is an *indicator* that can be usefully operationalized, along with other forms of identity cleavage such as 'religion', 'gender' and so forth and along with indicators of disparity such as distribution of socio-economic and political resources, for the measurement of social distance. This does not mean that theoretical, conceptual and empirical enquiry into the nature of 'ethnicity' itself is not necessary or useful, but that in so far as we are interested in examining quantitatively the relationship between 'ethnicity' and different economic and political outcomes such as the provision of public goods, violent conflict, and economic growth, a theoretical stance that posits 'ethnicity' as an indicator of the concept of 'social distance' appears both to be closer to the way in which these measures are already implicitly constructed, and to resolve many of the theoretical objections to the use of purportedly primordial datasets for the construction of these measures. To expand upon this latter claim, if we reconceive of ethnicity as an indicator of social diversity and disparity, the fact that the data we typically employ in the measurement of this indicator are neither as fluid nor as multidimensional as theorists of ethnicity *qua* ethnicity contend is less problematic because we take this as an issue of measurement bias and measurement error. Bias, because the categories we use are acknowledged to be those legitimized by historical state practices. Error, because the exclusive and exhaustive codings employed do not fully capture the nuance of ethnic identity. It may seem perverse to contend that some of the theoretical problems with measuring 'ethnicity' are resolved by acknowledging measurement bias and measurement

error, but our contention is that this is less problematic that reconciling 'primordial' data with constructivist concepts. Measurement error and measurement bias are inherent features of virtually all quantitative variables – even something as apparently uncontentious as GDP is subject to bias – the exclusion or under-estimation of the informal sector and other productive activities such as housework – and measurement error – as witnessed by the often large statistical discrepancy between the different methods of calculating GDP. One of the advantages of quantitative methods is that the accumulation of multiple observations allows us to minimise problems of measurement error in the examination causal relationships and correlations.

To summarize, then, within a Goertz-style schema we suggest that the various measures discussed below be understood as measures of different sub-concepts of diversity and disparity – for instance, fractionalization or horizontal inequality – and that these two in turn can be seen as sub-concepts of a broader notion of social distance. Crucially, however, we treat 'ethnicity' in this formulation as an *indicator* of diversity in the same way that land-ownership or income is an indicator of disparity. This conceptual map is depicted schematically in Figure 1.

Figure 1. Conceptual map of social distance



3. How do we measure diversity and disparity?

In the previous sections, then, we defended conceptual notions of diversity and disparity and argued that they were consistent with sophisticated theories of ethnicity as long as we treat the categorical 'ethnic' data employed in these measures as indicators of social distance rather than as ethnicity *qua* ethnicity. In this section, we review the range of measures of 'ethnicity' available in the quantitative literature. In accordance with our conceptual mapping, however, we treat these as different measures of social distance which may combine aspects of diversity and of disparity. We hence distinguish between 'pure' measures of diversity that measure (ethnic) diversity alone – that is, the measure has only one type of input: the number or population proportion of each group – and what we term 'synthetic' measures – measures that incorporate information on ethnic diversity with other information on between-group *disparity*, particularly geographical location, economic resources, and political power. Clearly, a third class of measures are possible which

incorporate information about disparity but not about diversity; into this class of measures would fall various traditional indicators of inequality such as the Gini coefficient and the Atkinson index, which are extensively discussed and debated elsewhere. It is important here to note that many of the measures we discuss in the synthetic section are flexible enough to be used in other ways – the indices of geographical segregation that we discuss in section 3.2.1, for instance, can easily be adapted as measures of occupational or educational segregation in the economic sphere.

3.1 Pure Diversity Measures

Leaving aside classificatory schemes (Bangura 2001) or 'dummy' variables for certain patterns of ethnic diversity such as Collier and Hoeffler's 'ethnic dominance' dummy, there are two main measures employed in the econometric literature on ethnic diversity: the ELF, discussed briefly above, and the polarization index developed by Montalvo and Reynal-Querol, which we refer to here and throughout as demographic polarization in order to clarify the distinction with the index of (economic) polarization developed by Esteban and Ray.

The standard measure of group diversity employed in econometric analyses of economic growth and political conflict is based on the concept of fractionalization. Effectively a Herfindahl concentration index, the measure represents the probability that two randomly selected individuals in the population belong to a different group. Where each group – ethnic, religious or other – constitutes proportion p_i of the total population, the measure is given by

$$FRAC = 1 - \sum_{i=1}^{N} p_i^2$$

The measure scores zero where in a perfectly homogenous population (i.e. all individuals belong to the same group) and reaches its theoretical maximum value of 1 where an infinite population is divided into infinite groups of one member. The index is intuitively

interpretable as the probability that two randomly selected individuals from the population will belong to different ethno-linguistic groups. The data originally used to construct the ELF measure, the *Atlas Naradov Mira* was based on a worldwide survey of ethnic diversity by a group of Soviet ethnologists in the early 1960s and published as an atlas (Bruk and Apenchenko 1964). As James Fearon further notes: 'The Soviet team mainly used language to define groups, but sometimes included groups that seems to be distinguished by some notion of race rather than language, and quite often used national origin (e.g. Anglo-Canadians are listed in the United States)' (Fearon 2003, 196).

While the ELF is easy to use, different scholars have pointed out serious faults and weaknesses of the index (see, in particular, Chandra 2001; Green 2004; Laitin and Posner 2001; Posner 2004). While some of these criticisms are specific to the ELF index, others focus more generally on the limitations of measuring ethnic diversity (Posner 2004). First, the ELF index has been criticized for not accurately representing ethnic diversity for at least three reasons: i) it is actually predominantly based on linguistic categories, and it therefore misses 'multiple dimensions of ethnic identity in all countries' (Laitin and Posner 2001, 14); ii) it is out-of-date and thus unrepresentative both of demographic changes over the past four decades and the realignment of country boundaries over the same period, and iii) it 'also suffers from a number of basic coding inaccuracies' (Posner 2004, 850).

Second, and more fundamentally, it is argued that the use of fractionalization in conflict analysis is problematic as it assumes a linear relationship between increasing fractionalization and increasing conflict vulnerability. This latter critique gave birth to the concept of (demographic) polarization. Daniel Posner identifies a number of other problems with summarizing a country's 'ethnic landscape with a single index of fractionalization', including the problem that the Herfindahl formula (which underlies the ELF index) is 'insensitive to a great deal of potentially relevant variation in the ethnic landscapes of the countries being compared', the problem 'that ethnic fractionalization indices such as the ELF

fail to incorporate potentially relevant information about the spatial distribution of groups around the country', and, the problem that fractionalization indices 'convey no information about the *depth* of the divisions that separate members of one group from another' (Posner 2004, 851). The serious shortcomings of the ELF index have driven a number of scholars to attempt to construct new measures and indices (see, for example, Alesina, et al. 2003; Fearon 2003; Posner 2004). While these new ethnic diversity measures, such as Fearon's Index and Posner's Politically Relevant Ethnic Groups (PREG), are very good at addressing some of the major faults of the ELF index, a major 'problem' with them is that they are no longer 'pure' demographic measures. Thus, for instance, Posner's PREG index is 'an index of fractionalization that reflects the groups *that are actually doing the competing over policy*, not the ones that an ethnographer happens to identify as representing distinct cultural units' (Posner 2004, 853; original emphasis).

Dissatisfied with this assumption of linearity, Montalvo and Reynal-Querol (2002; 2003; 2005) have developed an index of demographic *polarization* rather than fractionalization. The assumption behind their measure is that whilst the fractionalization-conflict matrix rightly attributes a low chance of ethnic conflict to homogenous population, highly fractionalization societies are also less likely to be conflictual as no group has the 'critical mass' necessary for conflict. Instead, they postulate that conflict is more likely the more a population is polarized into two large groups. The index given by the formula

$$MRQ = 4\sum_{i=1}^{N} p_i^2(1 - p_i)$$

Like *FRAC*, this measure scores zero in a homogenous population, but also tends to zero in highly fractionalized populations, reaches its maximum instead at the point of even distribution between two groups (i.e. i=2, p_1=0.5, p_2=0.5). The relationship between fractionalization and polarization is thus broadly quadratic, taking the form of an inverse U-

curve, with greater variation at the peak of the curve. Comparing their measure against the fractionalization index, Montalvo and Reynal-Querol (2003) find a statistically significant relationship between religious polarization and conflict, but no such relationship with religious fractionalization.

3.2 Synthetic Measures of Ethnic Disparity

3.2.1 Geographical Distribution and Ethnic Diversity: Segregation Indices

The measures of ethnic diversity discussed thus far have related to national-level functions of population distribution. Yet there is reason to suppose that the way in which groups are geographically concentrated may have important ramifications for ethnic politics, including conflict and economic development. There is intuitively a much greater risk of political instability if a 5% minority group is concentrated in one particular region of the country than if it were dispersed evenly across the country.

While segregation indices have not been used in the study of ethnicity in the developing world, within the sociological literature on race relations in the United States extensive use has been made of such indices of segregation focusing, for instance, on educational establishment. These indices measure how far different ethnic groups are concentrated across particular 'units' – schools, geographical regions, occupations, etc. (Coleman, et al. 1982; Reardon and Yun 2001). There is not one established or dominant index, however. Reardon and Firebaugh (2002) review the available sets of segregation indices in depth and distinguish four different analytical components of segregation, which are catered for in vary combinations by existing indices, two of which are of particular importance:

- Disproportionality – segregation as 'a function of the disproportionality in group proportions across organizational units' (ibid., 39). Thus, for instance, if Group A

constitute 10% of the overall population but 40% of one particular unit (e.g. geographical region), that group is disproportionately represented in the unit.

- Association – segregation as 'association between groups and organizational units', producing measures related to basic statistics such as the χ^2.

In addition, Reardon and Firebaugh identify a number of axiomatic principles that a segregation index should satisfy, including the transfer principle which states that the segregation index should show a reduction if a person is moved from one unit to another unit in which the person's group is less predominant that in their original unit. This principle turns out to be of particular pragmatic importance because only one measure, which they term *H*, satisfies this principle; *H* also incorporates both the disproportionality and the association components of segregation. Derived from the Theil information index, the formula for *H* (here termed *SEGH*) is given by

$$SEGH = \sum_{m=1}^{M} \sum_{j=1}^{J} \frac{t_j}{TE} \pi_{jm} \ln \frac{\pi_{jm}}{\pi_m}$$

where

$$E = \sum_{m=1}^{M} \pi_m \ln \frac{1}{\pi_m}$$

and

- *j* and *m* index units and groups respectively;
- t_j is the number of individuals in unit *j*;
- π_{jm} is the proportion of unit *j* constituted by group *m*;
- π_m is the proportion of group *m* in the overall population; and,
- *T* is the total population.

To operationalize the segregation index for measuring geographic segregation, we need to identify the appropriate 'unit' across which group distributions are to be measured. An intuitively plausible option is the Primary Administrative Divisions (PADs) of the country

in question – usually states in federal systems (e.g. the USA, Nigeria, etc.), and provinces (e.g. Indonesia, Canada), regions (UK), *départements*, in unitary states.

Using the Primary Administrative Divisions of a country as the basis for computing an ethnic segregation index presents two interlinked problems, however. Firstly, the size of PADs varies considerably from country to country, both in relative terms – the average proportion of the total population in each PAD – and in absolute terms – the absolute number of inhabitants in each PAD. Indeed, such size variation can occur within the same country, as well as across countries. To give an extreme example, the Indonesian province of East Java contains around 35 million inhabitants, or 17 per cent of the national populations; in contrast, the province of North Maluku has around 800,000 inhabitants, less than 0.5 per cent of the national population. Clearly, the larger the proportion of the total population contained with one subdivision, the larger the amount of segregation that could be 'hidden' within that province. We can demonstrate this with the two simple examples below. In scenario 1, three groups are distributed across three states, each group slightly concentrated in a different state, producing a segregation index (using *SEGH*) of 0.054. Scenario 2 describes the same population, but with States B and C merged to form state D. By combining these two states, the segregation measure is reduced by a half to 0.028, even though the 'actual' physical segregation of the groups has not altered, merely the administrative calculus.

Table 1. Scenarios of geographical segregation

Scenario 1

	Group 1	Group 2	Group 3
State A	10	10	20
State B	20	10	10
State C	10	20	10

Segregation 0.0536

Scenario 2

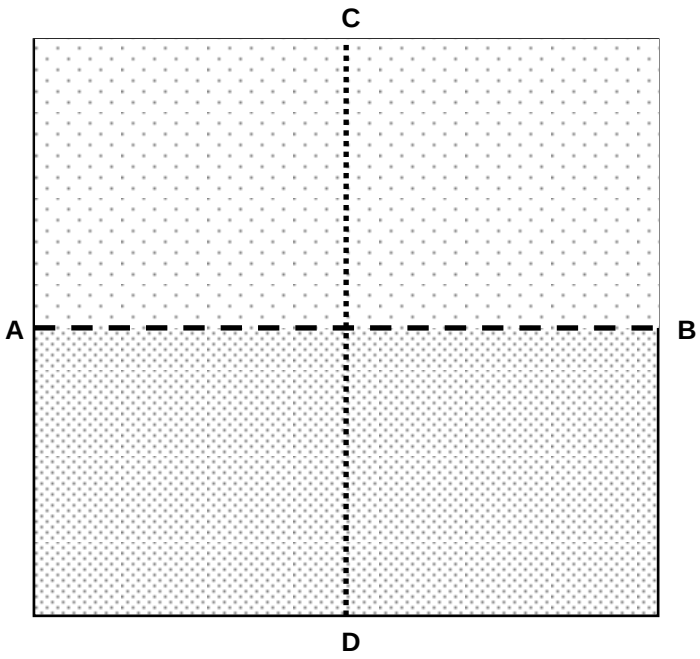
	Group 1	Group 2	Group 3
State A	10	10	20
State D	30	30	20

Segregation 0.0278

A second problem relates to the *way* in which countries are sub-divided. Assuming that the 'actual' spatial distribution of (ethnic) groups is relatively segregated, administrative boundaries can be drawn in a way that either accentuates or minimizes the calculation of segregation. Administrative boundaries drawn 'around' (ethnic) groups will accentuate the measure of segregation; boundaries drawn 'through' ethnic groups will minimise the measure. An extreme, simplified version of this problem is illustrated in Figure 2, which depicts a 'country' physically segregated into two groups, one of which occupies the 'north' of the country, the other the 'south'. If the country is divided into two administrative zones along the 'east-west' axis (line AB), segregation is complete – such a situation would score 1 on all six segregation measures. If, on the other hand, the country is divided along the 'north-south' axis (line CD), the *administrative* segregation of the country would score 0; each zone would be equally balanced between the two groups.

Segregation indices have not been extensively employed in development literature on ethnic diversity, but have clear potential applications in investigating regional dynamics of economic growth, service provision and, potentially, ethnic conflict. Moreover, in line with our observation above about the flexibility of synthetic measures, segregation indices may prove useful in exploring economic and sociological ramifications of ethnic segregation across non-geographical units, such as industrial sector.

Figure 2. Schematic diagram of segregation



3.2.2 Economic Distribution and Ethnic Diversity

3.2.2.1 Horizontal Inequality

There is a growing literature on different measures of 'horizontal' inequality between groups within society. Much of this literature relates to technical decomposition exercises that allow standard measures of individual 'vertical' inequality to be decomposed into 'between group' and 'within group' contributions (e.g. Anand 1983; Kanbur and Zhang 1999; Maasoumi 1986). Mancini et al. (2008) argue, however, that such decomposition measures are less useful when one is interested primarily in the extent of inequality between groups *qua* groups because they conflate the two different dimensions of inequality. After a review of different available 'direct' measures of horizontal group inequality, they conclude that the best available measure, which they term the GCov, is the simple population-weighted coefficient of variation in group means to the overall average – one of the measures proposed by Williamson in his early study of regional inequality (Williamson 1965). Algebraically:

$$GCov = \left(\sum_{i=1}^{n} p_i (y_i - \bar{y})^2 \right)^{1/2}$$

where n ethnic groups have a mean income (years of education, etc.) y_i relative to the overall mean of $\bar{y}$, and which constitute proportion p_i of the total population.

3.2.2.2 Economic Polarization

Section 3.2.3 above discussed the measure of demographic polarization introduced by Montalvo and Reynal-Querol. An alternative, synthetic, measure of polarization, developed by Duclos, Esteban and Ray (Duclos, et al. 2004; Esteban and Ray 1994), includes a dimension of economic distribution along with population distribution. Esteban and Ray argue that economic polarization, defined as the extent to which a populations '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' (Esteban and Ray 1994, 819), is an important and both qualitatively and quantitatively different dimension of distribution than inequality. They go on to argue that polarization may be high even when inequality is low and hence may provide a better correlate of conflict.

Derived axiomatically, the Esteban and Ray measure is given by the formula

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

where n groups constitute p_i proportion of the total population and have an economic endowment (income, assets, etc.) of y_i. The constant k is a normalizing function and can be set to $1/\mu$ to allow for comparability across populations. The other constant, α, is the required degree of 'polarization sensitivity' (Esteban and Ray 1994, 834) and is bounded by the range $0 \leq \alpha \leq 1.6$. Where $\alpha=0$, the formula exactly equivalent to the Gini coefficient; as α rises, the measure place more weight on how far the population is polarized into two large economically distant groups. Note also that when the economic aspect $|y_i - y_j|$ is disregarded and α is set to 1, the measure is equivalent to the *MRQ* demographic polarization index. The measure is thus in effect a combination of demographic polarization and group inequality and, in as much, has

strong intuitive potential as a correlate of conflict. Ezcurra (2009) finds that income polarization is negatively correlated with growth across the regions of the European Union, but more systematic evidence on the relationship between income polarization and economic growth is less widely available.

3.2.2.3 The Generalized Fractionalization Index

Bossert et al. propose a measure that expands upon the ELF by creating a Generalized Fractionalization Index (GELF) which is able to incorporate non-ethnic dimensions, including economic variables such as income, education level, and so forth. Bossert et al. derive the GELF index by defining a similarity matrix $\boldsymbol{S}$ which compares each individual within a society to each other individual and assigns them a similarity rank which ranges from 0 (the two individuals are not similar across any dimension) to 1 (the individuals are exactly similar on every dimension). Hence for example, in a population of three people, the similarity matrix might look like this:

$$\boldsymbol{S} = \begin{Bmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix}$$

Obviously, every person is exactly identical to themselves, producing the diagonal series of 1s in the matrix (person 1 compared with person 1; person 2 compared with person 2; etc.). In addition, Bossert et al. stipulate reasonably that similarities must be transitive – person 1's similarity to person 2 must be the same as person 2's similarity to person 1. Hence, the matrix is symmetrical along the diagonal. In this example matrix, person 1 and person 2 have a similarity score of ½, while person 3 is completely different from both person 1 and person 2. Bossert et al. then define their Generalized Fractionalization Index as

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

where s_{ij} is the degree of similarity between person *i* and person *j*, i.e. the respective value in the similarity matrix ***S***. As defined by Bossert et al., the GELF is calculated on the individual level, but it is easy to see that the index can be modified to take account of groups of different sizes that are internally homogenous. Defining t_{gh} as the similarity index between two internally homogenous groups that constitute proportion p_g and p_h of the total population ($p_g+p_h+...+p_m=1.00$), the GELF can be rewritten as

$$GELF = 1 - \sum_{g=1}^{m} \sum_{h=1}^{m} p_g p_h t_{gh}$$

Whereas the ELF can be interpreted as the probability that two randomly chosen individuals belong to different ethnic groups, then, the analogous intuitive interpretation of the GELF is simply the average expected level of dissimilarity between two individuals drawn at random from the population.

The great flexibility of the GELF measures is that potentially any number of dimensions – cultural, geographical, socio-economic, political, etc – could be incorporated into the measure. In addition, it does not dictate *how* entries into the similarity matrix are to be computed, although one could easily stipulate certain intuitive limitations upon the calculation, e.g. that an increase in 'distance' between two individuals on any one dimension should, ceteris paribus, result in a decrease in 'similarity'. As we explore below, however, this also renders the measure arguably *too* flexible.

3.2.3 Political Distribution and Ethnic Diversity

Less well developed in the literature, but nonetheless worth reviewing briefly, are measures of ethnic diversity that have been developed to incorporate *political* information rather than geographic or socio-economic. These measures may have particular importance for understanding the dynamics of ethnic diversity and growth, however, because of the emerging consensus in the literature that the influence of ethnic diversity on economic growth

operates indirectly through its impact on policy choices and public goods provision. Intuitively, the *political* dynamics of ethnic diversity are hence likely to be of importance. Indeed, as mentioned above, Daniel Posner has criticized the ELF fractionalization index in such terms for incorporating too much spurious information. For Posner, what is important in explaining political outcomes is the population distribution of *political relevant* ethnic groups, or PREGs. Posner hence re-estimates the ethnic diversity of African countries based on criteria of political relevance over four decade-long periods and computes a new fractionalization index, using the same formula as above (section 3.1). It is important to distinguish between Posner's PREGs and other attempts to re-estimate the ELF, such as that of Alesina et al. (2003). Alesina et al. are concerned with the *quality* of data going into the ELF formula, but still describe a 'pure' measure of fractionalization. By excluding political irrelevant groups, Posner's PREG constitute what a 'synthetic' measure in so far as there is a non-demographic (i.e. political) dimension in the computation of the index, even if this is simply what is 'left out', so to speak.

Lars-Erik Cederman and Luc Girardin (2007) develop a synthetic measure of ethnic diversity aimed at capturing both the importance of *dyadic* relationships and the importance of political power, which they term N* (N-star). Primarily aimed at mapping ethnic conflict rather than ethnic politics and developmental performance, Cederman and Girardin envisage ethnic politics as a 'star-like' constellation in which the overall level of diversity or polarization is less important than the interaction between the group(s) that control the state, and group(s) that do not control the state. Their measure stipulates the attribution of Ethnic Groups in Power (EGIPs) and assumes that those groups lie at the centre of a 'star-like' configuration, whereby the total risk of conflict is a multiplicative function of the risk of conflict between the EGIP and <u>each</u> other group separately. Algebraically, their measure is defined thus:

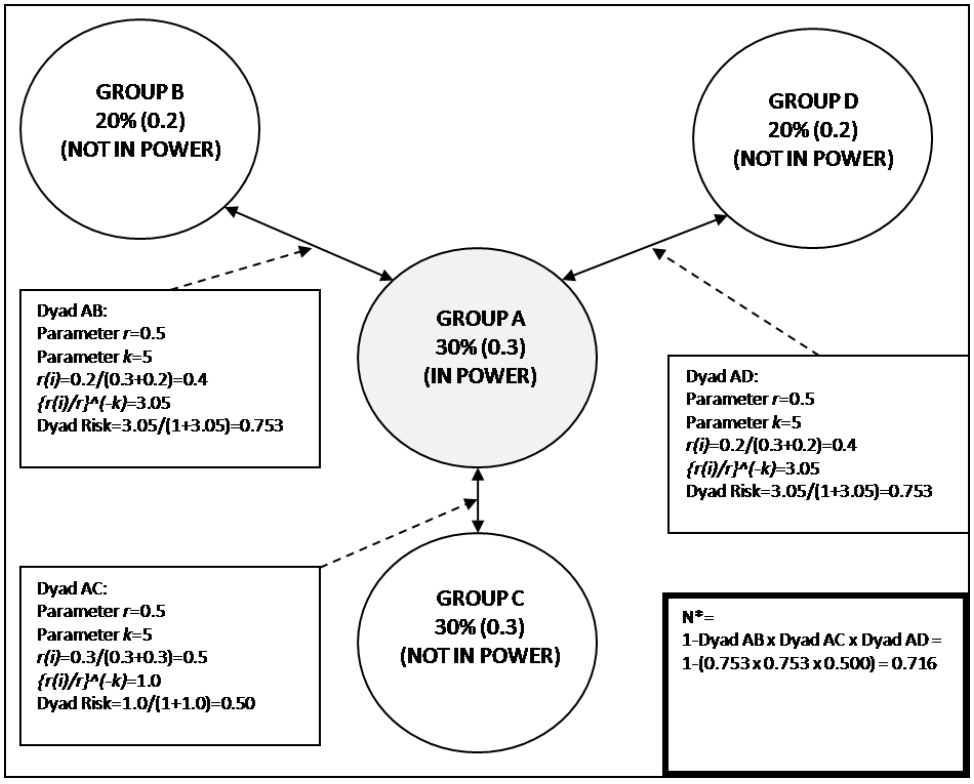
$$N^*(r,k) = 1 - \prod_{i=1}^{n-1} \frac{\{r(i)/r\}^{-k}}{1 + \{r(i)/r\}^{-k}}$$

where

- *r(i)* is the share of the combined population share of each dyadic relationship due to the non-EGIP group;
- parameter *r* is a 'threshold value' which 'stipulates at what demographic balance the odds are even for a challenge'. Cederman and Girardin recommend 0.5;
- parameter *k* controls how fast the risk of conflict increases. Cederman and Girardin recommend 5.

Figure 3 schematically depicts the calculation of the N*index for a hypothetical country with one group in power than constitutes 30% of the population (A), and three other groups, B-D, which constitute respectively 20%, 30%, and 20% of the population.

Figure 3. Graphical representation of Cederman and Girardin N* measure



The N* measure thus has a number of intuitively attractive features.

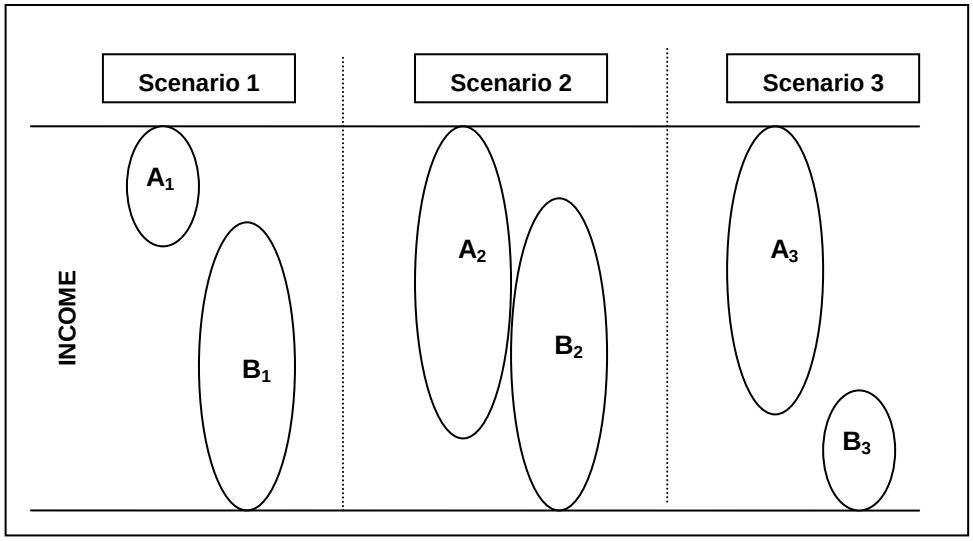
4. Does it matter how we measure 'ethnicity'?

In this section, we compare the *performance* of the 'pure' and synthetic measures of socio-economic diversity and disparity reviewed above. We do not compare the political measures because of a simple data incompatibility – the PREG data are only available for Africa, whereas the N-Star data is only available for non-African countries. Moreover, even socio-economic data is hard to come by systematically across countries. We thus instead limit ourselves to using within-country variation at the subnational level, drawing on raw census data from Ghana, Indonesia, and Uganda. In any case, at this stage our intention is simply to evaluate the extent to which these different measures actually pick up significant differences in underlying distributions. Our title for this section hence deliberately echoes Frances Stewart and her collaborator's (Stewart, et al. 2003) comparison of definitions of poverty. Whereas Stewart et al. sought to show how the term 'poverty' is bandied around rather too freely given that different definitions pick up very different sets of people, however, our broad argument in this section is largely the inverse: despite the fact that these different measures purport to capture very *different* distributional relationships, the measures produces remarkably highly correlated results, making it difficult – and from a policy perspective potentially dangerous – to draw firm conclusions about the relationship between ethnic diversity and development, at least on the basis of straightforward econometrics.

The three measures described above all synthesize dimensions of socio-economic disparity with measures of demographic distribution, but purport to capture different relationships – inequality, polarization, and fractionalization respectively. In an inter-group context, however, none of these has an immediately intuitive interpretation. To elucidate this, Figure 4 depicts three stylized distributional scenarios. In scenario 1, a relatively small group, A_1 is much richer on average than the larger group, B_1. Scenario 2 shows two evenly sized groups, one of which is slightly richer than the other. Finally, scenario 3 reverses scenario 1, with a large wealthy group and a small poorer one. How would be intuitively rank these in

terms of horizontal inequality, economic polarization, and generalized fractionalization (assuming, in this latter case, no within group heterogeneity)? Scenario 2 clearly has the highest level of *demographic* polarization and would score a maximum 1.0 on the demographic measure discussed above, but how does it rank in terms of economic polarization? This clearly depends to an extent on how much 'weight' we give to the different dimensions of polarization – economic and demographic – in the computation of the measure, a task performed by the α variable in the equation, as well as on how big the actual economic distance is. Similarly, we could reasonably assert that scenario 2 is the least unequal, but selecting which of scenarios 1 and 3 are the most unequal is more difficult. Moreover – and this point we will return to in some depth below – whichever measure we choose, we need to be aware that different scenarios such as those laid out below might have radically different implications in terms of political economy – for instance, group competition versus cooperation, as well as conflict potential – which cannot necessarily be neatly ordered by a single synthetic measure.

Figure 4. Scenarios of inter-group disparity



In the first place, however, we compare how these different measures rank different distributions. We do this those comparing stylized distributions and actual data from Ghana, Uganda and Indonesia. To compare stylized distributions, we systematize the type of scenario depicted above by calculating each different measure in two group scenarios with five different income ratios between the richer and the poorer group: 2, 4, 6, 8, and 10. We calculate each measure for all possible demographic distributions, from the richer group constituting 1% of the population to 99% of the population. To give some idea of what these ratios 'mean' in the real world, in South Africa at the height of Apartheid in 1970, the dispersion ratio of average incomes between Whites and Blacks was 13:1 (van den Berg and Louw 2004), while the political sensitive Chinese-Malay income ratio in Malaysia has varied between around 3:1 and 2:1 since 1970.

For the GCov, we need no further information to compute the measure. For the economic polarization index ER, we need to set the free variable α, which we set near its maximum 1.5, following other implementations of the measure (e.g. Ezcurra 2009); we discuss later the impact of varying α. For the GELF measure, we need to define the similarity matrix. With two variables of interest – 'ethnic' group and income – we operationalize the matrix by giving equal weight (0.5) to each dimension. The ethnic dimension is a simple dummy – 'individuals' of the same group score the full value (i.e. 0.5); those of different groups score the minimum value 0. For the income dimension, we take the reciprocal of the income ratio. Again, we assume no income variation within group for the moment. Hence, for instance, assuming the richer group constitutes 25% of the population and has an average income ratio of 4 times the poor group, we create the similarity matrix

$$S(0.25,4) = \begin{Bmatrix} 1 & 1 & 1 & 1/8 \\ 1 & 1 & 1 & 1/8 \\ 1 & 1 & 1 & 1/8 \\ 1/8 & 1/8 & 1/8 & 1 \end{Bmatrix}$$

Figure 5. Scenarios of horizontal inequality in a two group context

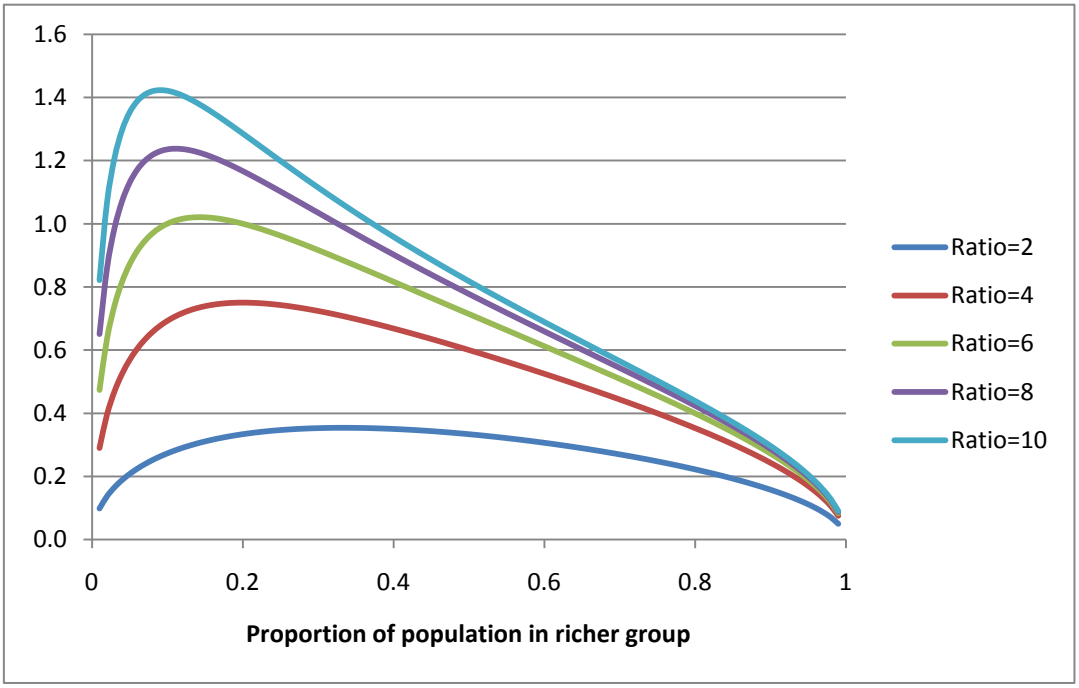


Figure 6. Scenarios of economic polarization in a two group context

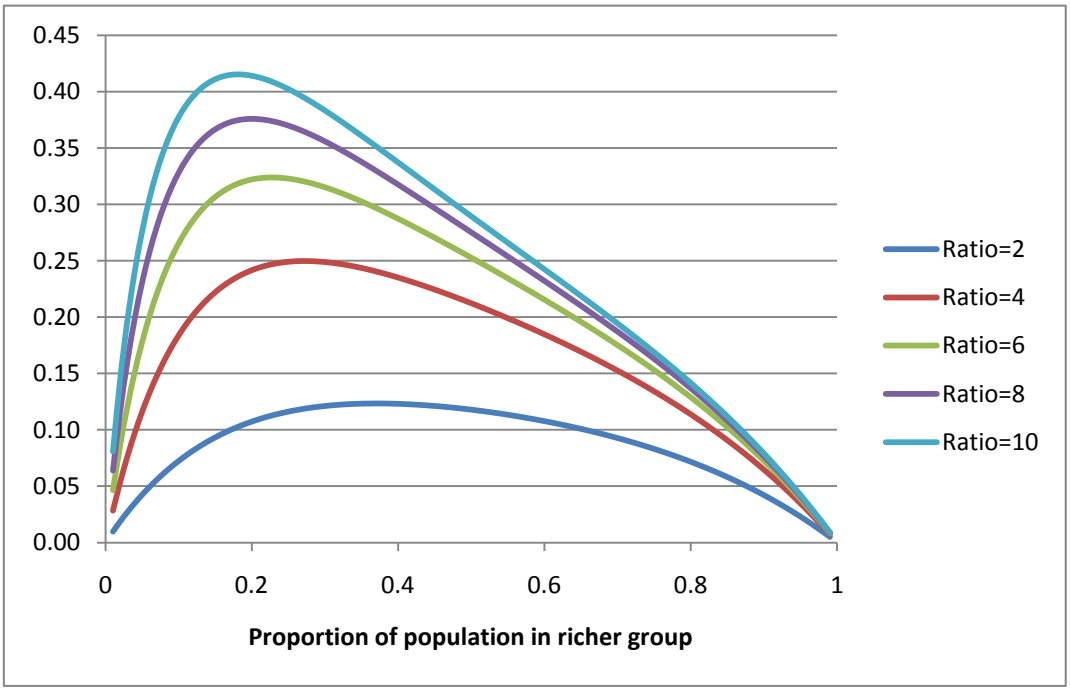


Figure 7. Scenarios of generalized fractionalization in a two group context with no within-group inequality

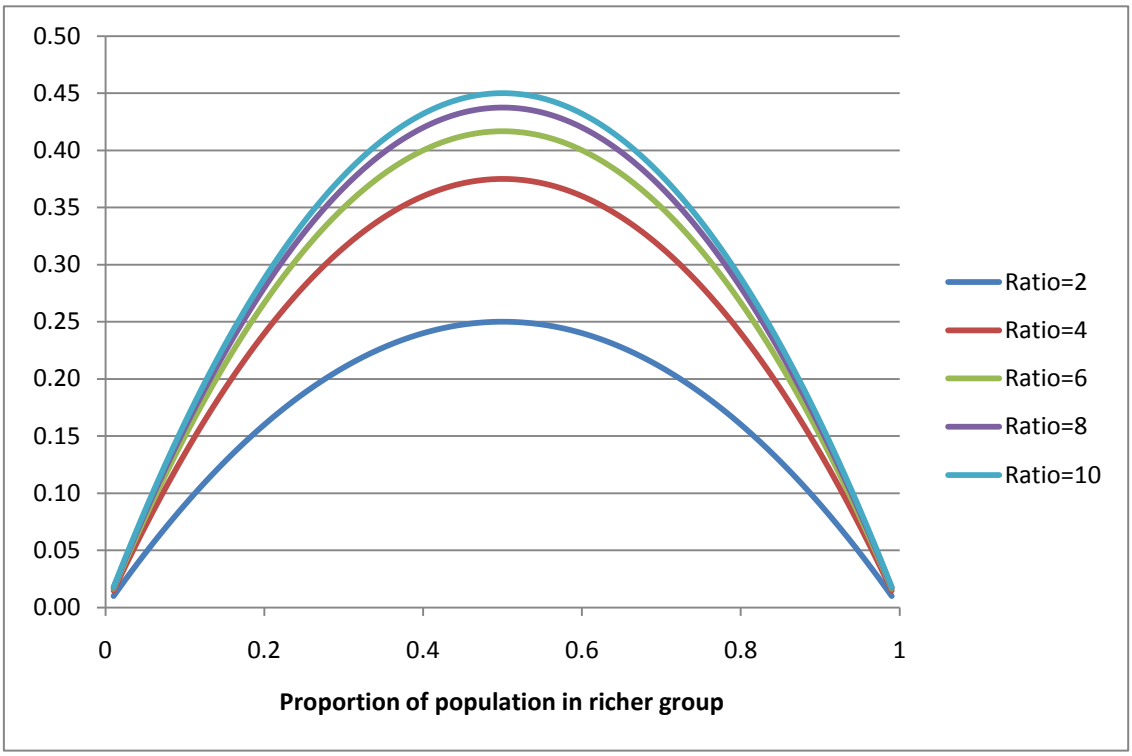


Figure 5 to Figure 7 report the results of these scenario calculations. At the lowest income ratio of 2:1, the curve traced by all three measures as we increase the proportion of the population in the richer group is broadly similar, tracing out an inverse-U curve. While the GELF measure produces an exactly symmetrical curve, with its maximum point at the 50:50 distribution, the other two measures are skewed towards higher results where the richer group is less than half the population; the GCov reaches its maximum where the rich group constitutes 1/3 of the population; the ER where it constitutes slightly over a third at 36.666%. As we increase the between group income ratio, the GELF measure remains symmetrical with the peak tending asymptotically towards 0.50 – where there is a theoretical complete economic as well as cultural distance between the two groups, the measure would collapse into the ethnic fractionalization index, which gives a result of 0.50 for a population distribution of 50:50. For the other two measures, however, as the disparity ratio increases, the peak of the curve becomes increasingly skewed towards reporting higher levels of inequality/polarization

in the population distributions with a smaller rich group: at the highest income ratio calculated here, 10:1, the GCov reaches it maximum where the rich group constitutes around 10% of the population; the ER polarization index at around 18%. While the GCov and ER measures clearly behave differently from the GELF at high inequality ratios, then, it is less clear that they are picking up qualitatively different distributions as claimed by Esteban and Ray. Moreover, the differences between these curves and the GELF measure only really emerge at high level of inequality which we would rarely see in the real world, although it must be borne in mind that the GELF measure is also able to incorporate dimensions of within-group heterogeneity, which we have presumed to be zero here.

To explore further the relationship between these different measures, we utilize real world data for subnational tiers from three countries: 110 districts in Ghana, 304 *kabuptan* in Indonesia, and 56 districts in Uganda. Our data is drawn from samples of the 2000 Census in each of the African countries and, in the Indonesian case, from the complete data of the 1995 Inter-Censal Survey. As reliable income data is not available in any of these surveys, we consider educational inequalities, using years of schooling completed as our socio-economic indicator. Further, we use religion rather than ethnicity as the identity marker of interest because of coding problems in the latter variable. As above, we initially set α to 1.5 for the ER economic polarization index. With raw data availability, we are able to compute the GELF measure on the individual basis; we weight it as above, 50% on a dichotomous variable of religious similarity and 50% on educational distance. Hence, for instance, two co-religionists among whom one has no education and one has the full complement of educational years available would score a similarity matrix entry of 0.50; two people of different religions but who had the same years of education would also score 0.50; two co-religionists, one of whom had half the available years of education and the other of which had none would score 0.75.

Table 2. Correlation matrix for different measures of religious and educational diversity, Ghana, Indonesia, and Uganda

	ELF	RMQ	GCOV	ER(1.5)	GELF(0.5,0.5)
ELF	1.000				
RMQ	0.855*	1.000			
GCOV	0.405*	0.338*	1.000		
ER(1.5)	0.261*	0.358*	0.920*	1.000	
GELF(0.5,0.5)	0.990*	0.833*	0.325*	0.176*	1.000

Note: Starred(*) results significant at the 1% level.

Figure 8. Scatter matrix of different measures of religious and educational diversity, Ghana, Indonesia, and Uganda

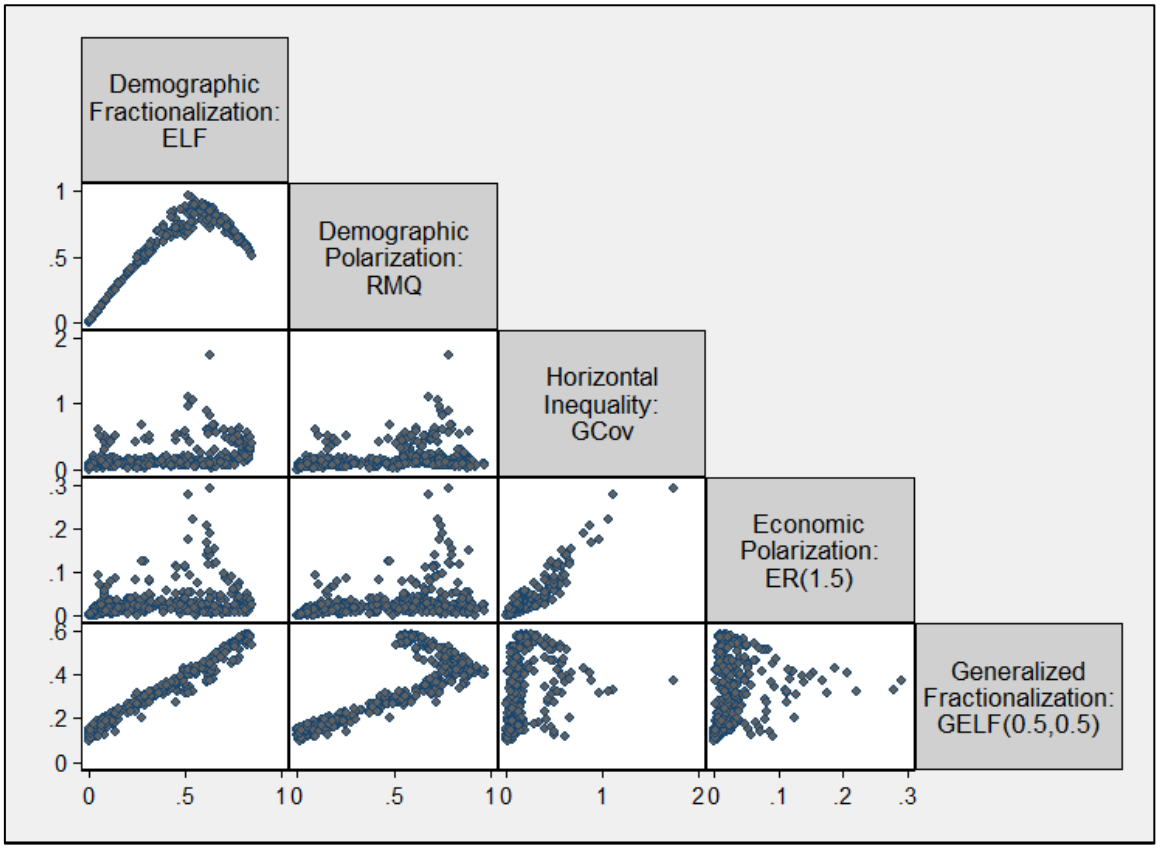


Figure 8 produces a scatter matrix to compare the performance of these measures at the subnational level, along with the 'pure' demographic fractionalization and polarization measures; Table 2 reports the equivalent Pearson's correlation matrix. Three particular observations are worth making of these data:

- As expected, the relationship between the 'pure' demographic measures of fractionalization and polarization traces a broad quadratic form, with greater variance at the 'peak' of the inverted-U curve;
- The relationship between the GCov measure of horizontal inequality and the ER index of economic polarization is very strong, with a correlation coefficient of 0.920. This confirms the findings of the stylized comparison that, while the two measures purport to identify different distributional relationships between groups, they appear to perform more or less identically when applied;
- The relationship between the demographic polarization index (RMQ) and its socio-economic extension (ER) is relatively weak (corr. 0.358), but the relationship between the demographic fractionalization index (ELF) and its socio-economic extension is much stronger (0.990).

Thus far, however, we have only considered the relationship between the different measures based on the specifications used in the stylized comparison. Both the ER and GELF measures, however, can be specified differently. With the ER measure, we can vary the degree of 'polarization sensitivity' by varying the value of α. As noted above, the GELF index is open to potentially limitless specifications of how the calculation of pairwise similarities, although the collapsing of the similarity matrix into a single result is fixed. Here we simply vary the 'weight' given to each respective dimension – religious similarity and educational similarity – calculating 11 different measures ranging from GELF(0.0, 1.0) in which no weight is given to religious similarity (and which is hence entirely a measure of 'vertical' inequality between individuals) to GELF(1.0, 0.0) which, conversely, gives no weight to educational similarities and which hence collapses into the simple demographic fractionalization index (ELF). The equally weighted measure used above is equivalent to GELF(0.5, 0.5). Other possible ways of modify the similarity matrix would be, for instance,

to include a dimension of 'inequality aversion' by squaring the reciprocal of the difference in educational achievement.

Rather than produce more correlation matrices for all these variables, we graph the changing correlation coefficient as we change the free variable of interest; Figure 9 does this for correlations between the ER measure and the ELF (equivalent to GELF(1.0,0.0)), RMQ, GCOV, and two other specifications of GELF as α changes; Figure 10 tracks similar changes in GELF as we increase the weight apportioned to religious similarities. (In Figure 10, the dotted sections of the curve represent correlations that are not significant at the 1% level; all other correlations in both figures are significant at this level). These results are, of course, based on our real world data, not on stylized comparisons as before.

Figure 9. Impact of changing α in ER(α) on correlation with other measures of ethnicity

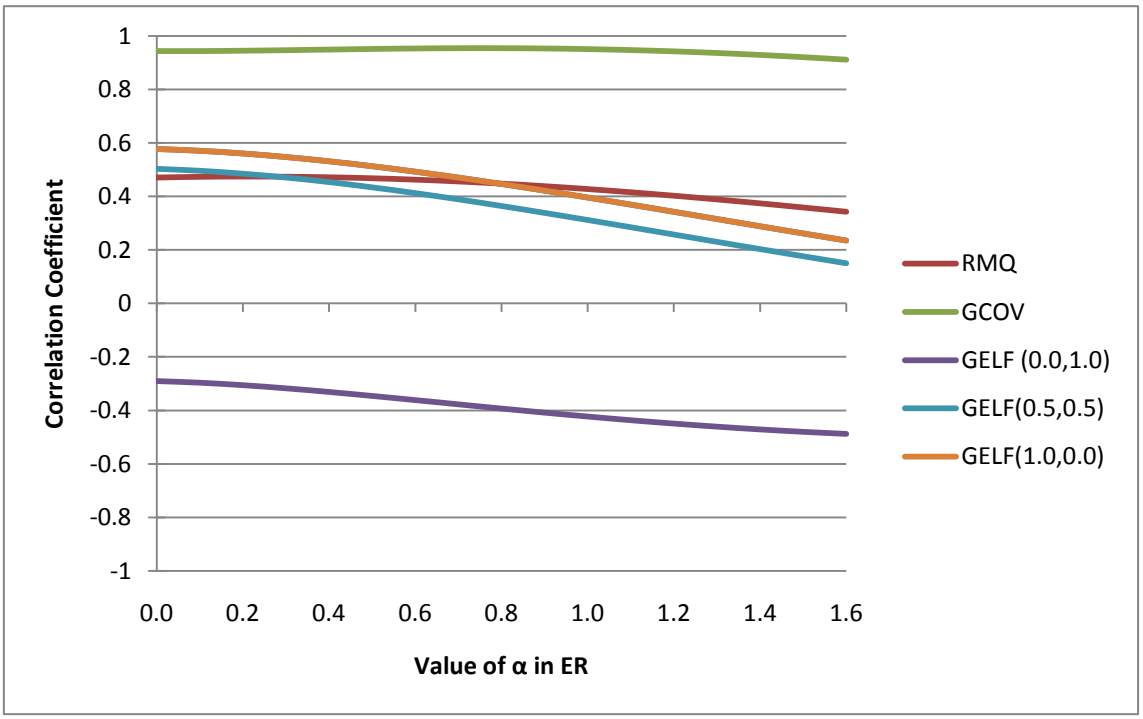
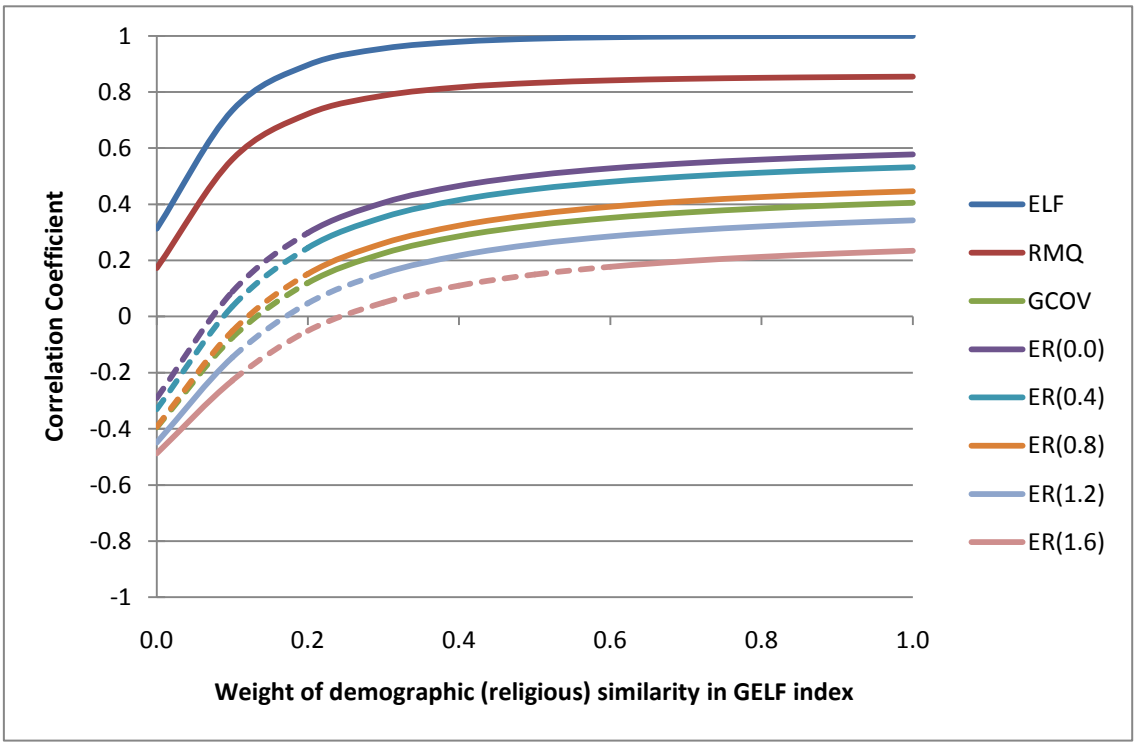


Figure 10. Impact of changing demographic weight in GELF on correlation with other measures



What can we make of these results? Increasing the 'polarization sensitivity' of the ER measure by increasing α results in a decrease in its correlation with all the other measures, except GELF(0.0, 1.0), which is negatively correlated and which increases in significance as α increases. GELF(0.0, 1.0), we should remind ourselves, is a purely 'vertical' measure, with no information about horizontal distributions. Across all values of α, however, the correlation with GCOV remains problematically high. As far as GELF is concerned, the correlation is unsurprisingly low – and, for the various measures of ER, insignificant – where the weight to religious similarity is likewise low, and at all level of weighting the correlation between GELF and the 'pure' measures of demographic diversity is noticeably higher than that for the other synthetic socio-economic measure. What is perhaps more surprising the speed with which this correlation 'picks up', such that with a weight of 0.2 to religious diversity and 0.8 to educational disparity – both between group and within group (which is not picked up by any

other measure) – the GELF is positively a significantly correlated with GCOV and all but the most 'polarization sensitive' specifications of ER.

Conclusions

In this paper, we have reviewed the ontological debates over the nature of ethnicity and the different ways in which it is operationalized and 'measured' for quantitative research. In the first two sections, we have defended the use of such measures by arguing that while the apparently 'primordial' implications of the ethnicity data that form the basis of such measurements are incommensurate with sophisticated theories of ethnicity, their employment in quantitative measures can be reconciled in so far as we treat 'ethnicity' as an indicator of concepts of diversity and social distance, rather than as the concept being measured in itself. In the latter two sections, however, we identified a more practical problem with these measures: while they purport to capture very different distributional relationship and dimensions of 'social distance' they are, empirically, very closely correlated, at least on the data we present here. This high level of correlation combined with a high likelihood of measurement error make the interpretation of results based on these indices extremely difficult. Put another way, if we were to find an strong statistical correlation between any one of these measures and a particular economic or political outcome – economic growth, provision of public goods, incidence of violent conflict – we would assert that we could reasonably take this as evidence of some kind of relationship between the top-level concept of social distance and the particular outcome we test, but that we could be less certain about the precise nature of this relationship.

Does this, then, save the methodological rigour of quantitative inquiry into ethnicity, but doom it again on practical grounds? We think not, and in conclusion we wish to suggest three possible ways forward. Firstly, while these different measures appear to be highly correlated against each other, quantitative enquiries that have tested different measures against

a particular outcome *have* produced noticeable differences. Montalvo and Reynal-Querol (2002), for instance, directly test their measure of demographic polarization against the fractionalization index, using the same data, and find that polarization is a significant correlate of ethnic conflict where fractionalization isn't. Similarly, Posner (2004) is able to use the differences between the significance of the ELF and his PREG index to explore the precise mechanisms through which ethnic diversity affects economic growth in Africa. The high empirical correlation between these various measures demonstrated here certainly caution us to be extremely careful in interpreting such results, but as data improve and the possibility of meta-analyses emerge, we may gain more confidence in such results. Brown and Langer (2011 [forthcoming]), for instance, replicate Sambanis' (2005) meta-study of civil war incidence across different conflict datasets with critical attention to the passage of time, and find – in qualified support of Montalvo and Reynal-Querol – that while ethnic fractionalization does not provide a strong correlate of conflict in any time period across any definition of conflict, ethnic polarization does appear to provide a consistent correlate in a particular time period.

Secondly, we suggest that the results here encourage us towards more theoretically sophisticated measure such as the N-Star that not only 'map' social distance but also incorporate a specific theorized *mechanism* linking social distance to a specific outcome. While data limitations prevented us from comparing the N-Star measure with the other measure of political disparity review, the PREG, Cederman and Girardin themselves compare the N-Star with the basic ELF, finding that N-Star is a significant correlate of ethnic conflict, where ELF is not. Crucially, the direct correlation between their measure and N-Star is very low (correlation coefficient 0.42). In this case, then, we have strong evidence of a particular mechanism translating social distance in a particular political outcome, and the low correlation between this theoretically-rich 'process' measure and the less sophisticated 'mapping' measure should comfort us that we have, indeed, identified something specific. This is

encouraging, but also clearly demands more attention to developing such sophisticated measures and less 'off-the-shelf' (Kalyvas 2008) style analyses that combine existing datasets with different outcomes. As noted above, the advantage of some of the measures discussed, such as the GELF and the segregation indices, is their flexibility. The advantage of the N-Star measure, however, is its *inflexibility* – it purports to capture a precise relationship between a particular form of social distance and a particular outcome and can be tested as such.

The final conclusion that we can draw from this review of quantitative approaches to ethnicity is the need for more engagement with qualitative analyses. Where through data or theoretical limitations we can only use indices that are highly correlated empirical but purport to pick up different relationships, a properly iterative combination of quantitative and qualitative enquiry should provide better insights into the particular mechanisms at work.

References

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2003. Fractionalization. *Journal of Economic Growth* 8: 155-194.

Anand, S. 1983. *Inequality and Poverty in Malaysia: Measurement and Decomposition*. Oxford: Oxford University Press.

Anderson, B. 1983. *Imagined communities: reflections on the origin and spread of nationalism*. London: Verso.

Bangura, Y. 2001. *Ethnic structure and governance: reforming Africa in the 21st century*. Geneva: UNRISD.

Bates, R. 2004. Ethnicity. *The Elgar Companion to Development Studies*. Harvard University.

Brass, P. 1985. Ethnic groups and the state. In *Ethnic groups and the state*, ed. P. Brass. London: Croom Helm.

Brown, G. K., and A. Langer. Forthcoming (2011). Riding the ever-rolling stream: Time and the ontology of violent conflict. World Development.

Brubaker, R., and F. Cooper. 2000. Beyond "identity." *Theory and Society* 29(1): 1-47.

Bruk, S. I. i., and V. S. Apenchenko. 1964. *Atlas Narodov Mira*. Moscow: Glavnoye Upravleniye Geodezii i Kartografii Gosudarstvennogo Geologicheskogo Komiteta SSSR and Institut Etnografii im N. N. Miklukho-Maklaya Akademii Nauk SSSR.

Cederman, L.-E., and L. Girardin. 2007. Beyond fractionalization: Mapping ethnicity onto nationalist insurgencies. *American Political Science Review* 101(1): 173-185.

Chandra, K. 2001. Cumulative Findings in the Study of Ethnic Politics: Constructivist Findings and Their Non-Interpretation. *APSA-CP* (Winter) 12(1): 7-11.

Cohen, R. 1999. The making of ethnicity: a modest defence of primordialism. In *People, nation and state*, ed. E. Mortimer and R. Fine. London: I.B. Tauris.

Coleman, J., T. Hoffer, and S. Kilgore. 1982. Achievement and segregation in secondary schools: A further look at public and private school differences. *Sociology of Education* 55: 162-182.

Collier, P., and A. Hoeffler. 1998. On economic causes of civil war. *Oxford Economic Papers* 50(4): 563-573.

——. 2002. Greed and gievance in civil war. World Bank, DECRG.

http://econ.worldbank.org/programs/conflict.

Doornbos, M. 1991. Linking the future to the past - Ethnicity and pluralism. In *Ethnicity and the state in Eastern Africa*, ed. M. M. A. Salih and J. Markasis. Stockholm: Elanders Gotab.

Duclos, J.-Y., J.-M. Esteban, and D. Ray. 2004. Polarization: Concepts, measurement, estimation. *Econometrica* 72 (6): 1737-1772.

Easterly, W. 1997. Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112: 1203-1250.

Easterly, W., and R. Levine. 1997. Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112: 1203-1250.

Esteban, J.-M., and D. Ray. 1994. On the measurement of polarization. *Econometrica* 62 (4): 819-851.

Ezcurra, R. 2009. Does income polarization affect economic growth? The case of the European regions. *Regional Studies* 43 (2): 267-285.

Fearon, J., and D. D. Laitin. 1996. Explaining interethnic cooperation. *American Political Science Review* 90 (4): 715-735.

Fearon, J. D. 2003. Ethnic and Cultural Diversity by Country. *Journal of Economic Growth* 8: 195-222.

Geertz, C. 1963. The integrative revolution - Primordial sentiments and civil politics in the new states. In *Old societies and new states - The quest for modernity in Asia and Africa*, ed. C. Geertz. London: Collier-Macmillan.

Glazer, N., and D. P. Moynihan. 1976. *Ethnicity: theory and experience*. Cambridge, MA: Harvard University Press.

Goertz, G. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.

Green, E. D. 2004. The (Mis)use of Ethnicity in the Current Political Economy Literature: Conceptual and Data Issues. Paper prepared for WIDER Conference on Making Peace Work. Helsinki, Finland, 4-5 June 2004.

Hardin, R. 1995. *One for all-The logic of group conflict*. Princeton, NJ: Princeton University Press.

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

Kalyvas, S. N. 2008. Promises and pitfalls of an emerging research program: The microdynamics of civil war. In *Order, Conflict, and Violence*, ed. S. N. Kalyvas, I. Shapiro, and T. Masoud. Cambridge: Cambridge University Press.

Kanbur, R., and X. Zhang. 1999. Which regional inequality? The evolution of rural-urban and inland-coastal inequality in China from 1983 to 1995. *Journal of Comparative Economics* 27: 686-701.

Kertzer, D. I., and D. Arel. 2002. Censuses, identity formation, and the struggle for political power. In *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*, ed. D. I. Kertzer and D. Arel. Cambridge: Cambridge University Press.

Laitin, D., and D. Posner. 2001. The Implications of Constructivism for Constructing Ethnic Fractionalization Indices. *APSA-CP* (Winter) 12(1): 13-17.

Maasoumi, E. 1986. The measurement and decomposition of multi-dimensional inequality. *Econometrica* 54(4): 991-997.

Nafziger, W. E., and J. Auvinen. 2002. Economic development, inequality, war, and state violence. *World Development* 30(2): 153-163.

Nagel, J. 1994. Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture. *Social Problems* 41(1): 152-176

Posner, D. 2004. Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science* 48(4): 849-863.

——. 2005. The Implications of Constructivism for Studying the Relationship between Ethnic Diversity and Economic Growth. Department of Political Science, University of California, Los Angeles.

Premdas, R. 1995. *Ethnic conflict and Development: The case of Guyana*. Avebury: Aldershot.

Rabushka, A., and K. A. Shepsle. 1972. *Politics in plural societies: a theory of democratic stability*. Columbus, Ohio: Charles E. Merrill Publishing Company.

Ranger, T. 1983. The invention of tradition in colonial Africa. In *The Invention of Tradition*, ed. E. Hobsbawm and T. Ranger. Cambridge: University of Cambridge Press.

Reardon, S. F., and G. Firebaugh. 2002. Measures of multigroup segregation. *Sociological Methodology* 32 (1): 33-67.

Reardon, S. F., and J. T. Yun. 2001. Suburban racial change and suburban school segregation, 1987-1995. *Sociology of Education* 74: 79-101.

Rodrik, D. 1999. Where Did All the Growth Go? External Shocks, Social Conflict and Growth Collapses. *Journal of Economic Growth* 4: 385-412.

Sambanis, N. 1999. Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature. *World Politics* 52(4): 437-483.

——. 2005. What is civil war? Conceptual and empirical complexities of an operational definition. *Journal of Conflict Resolution* 48(6): 814-858.

Shils, E. 1957. Primordial, Person, Sacred, and Civil Ties. *The British Journal of Sociology* 8: 130-145.

Stewart, F., C. Ruggeri Laderchi, and R. Saith. 2003. Does it matter how we define poverty? *Oxford Development Studies*.

Ukiwo, U. 2005. The study of ethnicity in Nigeria. *Oxford Development Studies* 33(1): 7-24.

——. 2006. Horizontal inequalities and violent ethnic conflicts: A comparative study of ethnic relations in Calabar and warri, Southern Nigeria. PhD thesis, Oxford Department of International Development, University of Oxford.

Van den Berg, S., and M. Louw. 2004. Changing patterns of South African income distribution: Towards time series estimates of distribution and poverty. *South African Journal of Economics* 72(3).

Van den Berghe, P. L. 1978. Race and ethnicity: a sociological perspective. *Ethnic and Racial Studies* 1(4): 401-411.

Varshney, A. 2001. Ethnic conflicts and ancient hatreds: cultural concerns. *International Encyclopaedia of the Social & Behavioural Sciences*: 4810-4813.

Williamson, J. G. 1965. Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change* 13(4): 1-84.

Young, C. 1985. Ethnicity and the colonial and post-colonial state in Africa. In *Ethnic groups and the state*, ed. P. Brass. London: Croom Helm.

——. 2003. Explaining the conflict potential of ethnicity. In *Contemporary peacemaking: conflict, violence and peace process*, ed. J. Darby and R. MacGinty. Basingstoke: Palgrave Macmillan.

Abstract (in Japanese)

要約

本稿では、「民族（ethnicity）」の概念がどのように定量的研究に運用されてきたかを検討する。民族に関する現在の理論と統計的処理における民族の用法の間に概念的なミスマッチが存在する。このミスマッチは、「民族」をより幅広い概念である「社会的距離」の指標として用いるものとして計量経済学的方法を再構築することによって、克服できる。本稿では、「社会的距離」を更に「多様性」と「格差」に細分化し、先行研究で用いられている既存の測定方法の強みと弱みを検討する。これらを踏まえ、各測定方法ごとの定型的な測定分布、更にはガーナ、ウガンダ、インドネシアの県レベルの実際のデータを用い、民族に関する異なった測定方法がどの程度相関関係を持っているかを検討する。その結果、かなり異なった基礎分布を取り上げたとしている測定方法でさえ、実際には強い相関関係を持っていることがわかった。この結果は、計量経済的結果を解釈する際には一層注意を要することを示唆している。特定の政治的結果と特定の分布を結びつけるような理論的に更に洗練された民族の測定方法を用いた定量的な研究は、因果関係を綿密に検討する定性的な研究と組み合わせることによって、民族に関する実証研究の適切な将来の方向性を指し示すものとなるのかもしれない。

JICA Research Institute

Working Papers from the same research project

"Ethnic Diversity and Economic Instability in Africa: Policies for Harmonious Development"

JICA-RI Working Paper No. 6

Diversity of Communities and Economic Development: An Overview

Gustav Ranis

JICA-RI Working Paper No. 7

Market, Democracy, and Diversity of Individual Preferences and Values

Satish Jain

JICA-RI Working Paper No. 8

Stability of a Market Economy with Diverse Economic Agents

Anjan Mukherji