NUMBER 4 ■ JUNE 2010

CRISE:

Centre for Research on Inequality, Human Security and Ethnicity

Monitoring and measuring horizontal inequalities

Frances Stewart, Graham Brown and Luca Mancini

UKaid from the Department for International Development

OVERVIEW

NUMBER 4 ■ JUNE 2010

Monitoring and measuring horizontal inequalities

Frances Stewart, Graham Brown
and Luca Mancini

OVERVIEW

Copyright

This document is an output from a project funded by UK aid from the UK Department for International Development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of DFID.

Design and layout: Rick Jones (rick@studioexile.com)

Contents

Executive summary

While the field of measuring vertical inequality (VI, i.e. inequality among individuals in a society) is well developed and sophisticated, much less attention has been paid to measuring horizontal inequalities (HI), or inequalities among groups in a society. This paper proposes ways of measuring and monitoring horizontal inequalities. It is important to monitor HIs because they affect the wellbeing of members of a group, and can be adverse to efficiency, growth and poverty reduction. Most important, they raise the risk of violent conflict. Yet for the most part, measurement and monitoring of HIs do not form a routine part of international or national data gathering and reporting.

Before proposing approaches to measurement the paper reviews certain challenges this poses, including the problem of defining group boundaries and data issues. Group delineation is difficult because identities are fluid, multiple and may be endogenous. Yet, felt differences are sufficiently clear and important in many societies to make it possible to measure group performance so long as one is sensitive to the possibility and implications of alterations in group boundaries. Historical and contemporary political analysis and surveys of people's own perceptions of significant identities contribute to defining salient group boundaries.

We emphasise the need to assess HIs across four main dimensions – social, economic, political and cultural status. Within each dimension, there are a number of elements – for example, within the economic dimension, employment, assets, and incomes are significant elements. A chronic problem is deficient data. There are some relevant international data sets, however, and in most multiethnic or religious countries, some useful data are available. However, there is a clear need to improve the data situation on HIs.

The aim in devising formulae to measure HIs is to identify appropriate measures which summarise HIs across a society, just as measures of VI do for distributions across individuals. Three prominent issues arise: first, how to combine information on multiple groups; secondly, how to take into account the whole distribution of a group and not only the average performance; and thirdly, whether to develop multidimensional indices of HIs.

A review of alternative aggregate measures of group inequality finds that the GCOV, the GGini and the GTheil, each with population weighting, all provide suitable measures that meet the main principles of a good measure of inequality – largely derived from principles developed for measures of vertical inequality. Data on South Africa show that including population weighting can be important, as the trend change in HI differs in this case between the measure that includes population weighting and the one that does not. Logic as well as intuitive interpretation suggests that the population-weighted measures are preferable. The population-weighted GCOV and the group Gini measures generally move in the same direction, with the GCOV showing the least change over time.

Evidence from Indonesia shows quite a high correlation across the three measures, so which is chosen may not matter much. There is low correlation, however, between HI and VI and none if homogeneous districts are included, indicating that VI is *not* an adequate measure of HI. Both VI and HI show persistence over time.

If one is particularly concerned with the position of one particular group, or of one group relative to another group, the clearest way to present the data is to take the ratio of the performance of the group to the mean, or the ratio of the performance of the two groups, rather than some aggregate societal group measure of inequality which also introduces other groups and may thus conceal the variable of interest.

We also explore alternative methods of assessing HIs allowing for the whole distribution within each group since the political and policy implications differ according to where the inequality lies within the distribution. We found that the α-means measure developed by James Foster provided the most instructive summary of this issue, since it indicates how HIs differ as one alters the weight given to different parts of the distribution. As data from South Africa shows, this approach also permits one to see how far any change in HIs over time is due to changing inequalities at the top, the bottom or throughout the distribution. Another approach is to compare the average achievements of different quantiles of the population (q-means) which can be easier to estimate and provides an intuitively comprehensible picture.

We discuss the issue of multidimensional indicators of HIs, and conclude that it is not desirable to provide a synthetic measure across dimensions because for policy purposes one needs to know the inequalities in each dimension. There might, however, be reasons to adopt a multidimensional indicator for elements *within* a dimension – e.g. for all types of economic or political inequality. One approach to providing a synthetic measure for a particular dimension is to generate subjective data on perceived inequalities in different dimensions, asking people whether they are better or worse off than other groups in economic terms, for example. Another approach is to develop an indicator based on the number of elements in which some degree of deprivation occurs. But for the most part, given scarce data it is usually most straightforward to present the data for each element separately, and most useful from a policy perspective.

The value of a measure of inequality depends on the purpose for which it is needed. One purpose is to make a general statement about whether a society is getting fairer or less fair over time from a group perspective. Another is to identify a variable which will enable us to test whether particular inequalities are correlated with other events, such as conflict, criminality, and unhappiness. For both purposes, group weighting by population would seem desirable, adopting an α-means approach. But it could be argued that the first objective might require measures which are widely understood – which could be a problem for these somewhat complex measures. From the perspective of intuitive understanding of how HIs vary across the distribution, a simple approach comparing outcomes in different quantiles of the population in each group could be useful.

A common question is how VI and HI are related. HI is a component of aggregate societal vertical inequality (VI), but there is no reason for any correlation between vertical inequality *within* a particular group and inter-group inequality. Empirical evidence shows that HIs generally form only a small component of VI.[1] On this reasoning, one would expect a small positive correlation between the two measures. This is shown in the Indonesian data, which also shows that VI is not at all an adequate proxy of VI. Consequently, one needs independent measures of HI to be monitor group inequalities. The common practice of using the contribution of between group inequality to total inequality as a measure of the extent of HIs is highly misleading from the perspective of assessing the severity of HIs. Hence independent measures of HIs are explored in this paper.

[1] See e.g. Papatheodorou (2000); Zhang and Kanbur (2003); Pradhan *et al.* (2001).

1 Introduction

Most discussion and measurement of inequality concerns vertical inequality, or inequality among individuals (VI), and is generally confined to a few economic variables, such as income, consumption, and sometimes assets. In this arena, the application of Lorenz curves and the Gini coefficient have been extensively and powerfully used. In doing so, inequalities between groups (defined here as horizontal inequality – HI) tend to be ignored. We argue that it is important to measure such inequality; and that its measurement should extend beyond the variables commonly used to measure vertical inequality, not only to other aspects of economic resources, but also to social and political entitlements. We consider some of the challenges raised by measuring HIs, which include defining group boundaries and getting data as well as the question of the most appropriate way of measuring HIs.

The next section of this paper briefly discusses why and when group inequalities matter; Section 3 considers problems in categorising groups, selecting dimensions of HIs and finding data. In measuring societal HIs for a society three issues need to be considered: first, how to aggregate inequalities across groups when there are more than two groups; secondly, how to take the whole distribution of variables of interest of each group into account, rather than compare the means alone; and thirdly, whether (and how) to arrive at a multidimensional index or indices of HIs. Section 4 considers problems in arriving at aggregate measures of HIs in a society where there are multiple groups and Section 5 discusses how the whole distribution might be taken into account in estimating HIs. Both propose specific measures and illustrate these empirically. Section 6 briefly discusses the question of multidimensional indices for HIs. Finally, Section 7 presents conclusions.

What horizontal inequalities are and why they matter

2

Horizontal inequalities are inequalities between groups. People can be grouped in many ways, and most people are members of many groups. There is a large range of types of group: national, racial, ethnic, religious, gender and age are some obvious important ways that people are categorised. There are also many rather transient types of groups – such as social clubs or producer networks. Some group affiliations are clearly more important and durable than others. In some cases the categorisation emerges largely from self-identification, in others, classification comes from legal factors (such as citizenship), as a result of categorisation by others, or some combination.

Group affiliation and inequalities matter because they affect economic and social objectives indirectly by affecting economic efficiency and social stability, and because they have a direct impact on individuals' well-being, particularly, when:

- group boundaries are relatively tight, so people cannot move easily (sometimes at all) from one group to another. An example is being of one gender; another is being a citizen of a particular country. If it is easy to change groups then the affiliation matters much less;
- being a group member leads to different treatment by others – e.g. via discrimination (in the case of gender, this may start at birth or even before. In the case of Africans in Apartheid South Africa, there was cradle to grave discrimination with political as well as economic dimensions). Privileges for particular groups also enhance the importance of group membership; and
- members of the group feel that being part of the group constitutes a significant aspect of their identity, and thereby group achievements contribute directly to members' well-being.

The direct impact of group inequality on members' well-being is important. People's well-being may be affected not only by their individual circumstances, but also by how well their group is doing relative to others.[2] This is partly because membership of the group is part of a person's own identity, and partly because relative impoverishment of the group increases the perceptions of members that they are likely to be trapped permanently in a poor position, or, if they have managed to do better than many in the group, there is a high risk that they will fall back into poverty.

Hence, it seems likely that the well-being of Muslims in Western Europe, Catholics in Northern Ireland, Hutus in Rwanda, Afro-Americans in the US, Africans in Apartheid South Africa, to take just a few of many examples, is (was) deeply affected by the relative impoverishment of the group – over and above the position of the individual themselves. Psychologists have shown, for example, that Afro-Americans suffer from many psychological ills due to the position of their group (Broman, 1997; Brown *et al.*, 1999). This is summarised by the title of one article: 'Being Black and Feeling Blue' (Brown, *et al.,* 1999). Hence the relative position of the group should enter into a person's welfare function (Akerlof and Kranton, 2000). The weight to be given to this element is an empirical matter on which more research is needed.

In addition, group inequality affects other objectives, in particular economic efficiency and social stability. Any economy in which there is significant discrimination against members of some group is likely to be less efficient than in the absence of such discrimination, since

[2] Akerlof and Kranton (2000) have adjusted individual utility to allow for this.

talented people in the group discriminated against will be held back, while too many resources, or too high a position, will go to less talented people in the favoured group. For example, Macours (2004) has argued that ethnic diversity in a context of weak property rights enforcement can result in market segmentation and less than optimal land allocation. In Guatemala, for example, informal land contracts are more likely to take place within the same ethnic group. Conversely, most studies show that affirmative action for Afro-Americans in the US has had a positive impact on efficiency (Badgett and Hartmann, 1995).

Moreover, if a group has experienced relative impoverishment historically, sometimes over many generations, this means that the present generation of a group may suffer relatively to others even without current overt discrimination. This is because family background, including nutrition and educational levels, influence a child's chances in life; and social networks operate disproportionately within a group and less between groups – indeed Blau (1977) regards having more in-group than out-group interactions as a defining characteristic of a group. Consequently, a member of a poor group has less advantageous social networks (or fewer good contacts from an economic perspective). The social networking point is less important where membership of the various groups to which an individual belongs is not coterminous, such as where particular social groups are multiracial or multiethnic.[3]

Perhaps the most important reason for concern with horizontal inequality is that it can be a source of violent conflict (Stewart, 2000; Stewart 2008). Group inequality provides powerful grievances which leaders can use to mobilise people for political action, by calling on cultural markers (a common history or language or religion) and pointing to group exploitation. This type of mobilisation seems especially likely to occur where there is political as well as economic inequality, so that the leaders are excluded from political power, while the mass of the population face unequal access to economic and social resources. Examples where group inequalities have been a factor in provoking conflict include: Côte d'Ivoire, Rwanda, Northern Ireland, Chiapas, and the Sudan to mention just a few (see Gurr, 1993; Stewart, 2002; Gurr and Moore, 1997; Langer, 2005; Murshed and Gates, 2005). Sharp horizontal inequalities within countries (and between them) are an important source of grievance and potentially of instability, independently of the extent of vertical inequality.

Finally, HIs can be an obstacle to the achievement of other targets, like poverty reduction or other MDGs. In some cases, it may not be possible to improve the position of individuals without tackling the position of the group. For example, programmes to advance credit to poor producers, or to promote universal education, can be impossible to achieve in the presence of severe group inequality. An example here is extending education to all girls which may be prevented not by a lack of schools or teachers, but by societal and parental attitudes to girls' education (Hafeez *et al.,* 1990; United States Information Service and United States Agency for International Development, 1992).

[3] Here it is assumed that every member of the group is treated equally. In practice, however, multiracial social groups often exercise segregation within the group.

Essential preconditions for measuring HIs

3

3.1 Determining the relevant groups

The first requirement in measuring HIs is to decide how people are to be categorised. The aim is to identify politically relevant and salient distinctions. These differ across societies – for example, religion is the important dividing line in some societies (like Northern Ireland or the Middle East), while it is ethnicity in many countries in Africa and Latin America; region can be important as indicated by many separatist struggles; and 'race' forms a dividing line in some societies, like Fiji or Zimbabwe. In many places, these identity distinctions overlap, and people are divided by religion and ethnicity, for example.

Given multiple identities and the social construction of identities, there are very few groups where boundaries are clear cut. For example, among Roma people in Eastern Europe, while 90.8% of people interviewed stated that they 'feel Roma', only 47.9% reported that they had declared themselves to be Roma in the census (UNDP, 2003). In Guatemala, 'under certain circumstances an individual can be born indigenous and become Ladino during the course of his or her life' (Caumartin, 2005: 8). While language is often used as a marker in Guatemala, the 1994 population census indicates that over a quarter of the self-declared indigenous people do not speak an indigenous language. Similar problems arise with most other groups that at first glance seem to be fairly easy to define. An example is Christians in the UK – does this group consist of the Church-going minority, or everyone who does not declare another religion? In Africa, there are many sub-ethnicities and much intermarriage which make boundaries fluid and ill-defined.

Moreover, an identity which seems important to the people themselves, or to others, at one point, may become quite trivial with time. For example, nowadays differentiation between descendants of the Angles and those of the Saxons or the Normans in England is insignificant, yet once these were key political distinctions. Moreover, many people declare themselves to have multiple identities, so that categorisation of each individual uniquely into one group becomes impossible. With political mobilisation, changes in how people see themselves and others may emerge: the *Orang Asli* (an active indigenous group in Malaysia), for example, were developed as a single group out of at least eighteen different identities in order to help people mobilise, first against the British and subsequently for their rights in independent Malaysia. Muslims in Europe have been mobilised as a group for political reasons as a result of the reactions to the 9/11 bombings and the Iraq war.[4]

To some extent, then, group boundaries become endogenous to group inequality. If people suffer discrimination (i.e. experience horizontal inequality) they may then feel cultural identity more strongly, particularly if others categorise them into groups for the express purpose of exercising discrimination (thereby creating or enforcing HIs). As Gurr (1993:3) has stated: 'The psychological bases of group identification are reinforced by cultural, economic and political differentials between the groups and others: treat a group differently by denial or privilege, and its members become more self-conscious about their common bonds and interests. Minimize differences and communal identification becomes less significant as a unifying principle'.

These types of argument could be used to discredit any attempt at measuring HIs. Yet this would be wrong: because, even though socially constructed and fluid, these differences do matter to people, as argued above. Moreover, even to test whether they matter or not, and which group distinctions matter, it is necessary to start with some categorisation. But the categorisation should, in so far as is possible, be sensitive to people's self-positioning. It is

[4] For example, in Denmark, Islamic leaders 'urged the country's 170,000 Muslims to vote against the ruling centre-right coalition' (*Financial Times,* 29/30 January 2005).

also desirable to explore whether adopting different categorisation criteria changes the results. To take the Guatemalan example, although there are many ways of classifying people into groups, the indigenous group, by whatever definition, still comes out at the bottom of the distribution with respect to virtually all variables. In practice, data deficiencies mean that only rather crude classifications are typically available. But once the importance of the issue is acknowledged, multiple classifications may emerge, as they have, for instance, in the UK census.

In measuring horizontal inequalities in a country, the first step is to classify the relevant identity groups: that is the group boundaries that people mind about, and the boundaries on the basis of which discrimination or favouritism occurs. Despite the fluidity of many group boundaries with cross-group marriages and multiple identities making classification of the politically salient groups difficult, in most multiethnic societies it is possible to identify the major politically relevant groups relatively easily. In order to identify relevant groups, the following procedures are recommended:

- An initial *in-depth investigation* of the history and political economy of the country in question, leading to indications of important group distinctions. In Sudan, for example, this would suggest differentiating Southern and Northern groups while in Rwanda, the Hutus and Tutsis would emerge as clearly salient groups.
- Drawing on *surveys of people's own perceptions* of identity distinctions, where available, or carrying out such a survey. Such surveys ask people which groups they feel are privileged or deprived, favoured or disfavoured by the government, and the importance of different aspects of identity to the people themselves.
- Adopting a multiple approach, taking a *variety of group classifications* (ethnic, regional, religious) and seeing where the main inequalities emerge.

3.2 Dimensions of horizontal inequality

Inequality has many dimensions. Economists tend to measure inequality in terms of incomes or consumption, although Sen has advocated adopting the space of capabilities (Sen, 1980). The appropriate space depends on what we are trying to assess. If it is well-being, income or consumption are very poor proxies since in general they do not include access to public goods, nor allow for distribution within the family, nor for the varying ways in which a given amount of income translates into capabilities (what people can be or do).[5] Hence the much more extensive multidimensional space of capabilities is preferable.

However, in practice capabilities (as potentials) are not easily measured, so inequality of functionings (what people actually are or are doing) is generally the only available measure. Capabilities, or indeed functionings, are the consequence of all sorts of circumstances – i.e. of possession or access to a variety of assets, to employment opportunities, the availability and quality of publicly provided goods and the capacity and constraints faced by the individual. While inequality of the outcome (functionings) is of concern, so is inequality of some of the more significant inputs – partly because this helps to diagnose the source of outcome inequality, and partly because such inequality contributes directly to people's well-being. For example, inequality in land ownership not only contributes to people's ability to be nourished, but it also contributes directly to their self-respect, status and well-being.

To the extent that we are concerned with the political impact of such inequalities, what is important are the elements that seem most significant to the people involved, i.e. what they believe to be central – indeed, in the extreme, the kinds of things which people will fight over. The evidence suggests that this differs across societies and groups: in Northern Ireland, for example, it appears that people are particularly concerned with their employment and housing inequality; in Zimbabwe their actions suggest they pay attention to land inequality; in Britain, young blacks object to being stopped by the police as they go about their business; in Sri

[5] Justino *et al.* (2004) show that there is no reason to assume that income inequality is a good proxy for other non-income inequalities.

Lanka, people's major concerns are employment prospects, access to higher education and to land. From both a well-being and a political perspective, then, these rather concrete variables may be of more importance than outcome variables, like life expectancy or nutrition levels, or incomes, which are less visible on a day-to-day basis. Moreover, political exclusion and HIs are highly relevant to well-being and to political mobilisation. Political variables include membership of the government, of parliamentary assemblies and local authorities, of the military and of the police. The perceived importance of these variables is illustrated by the large part they play in any post-conflict peace negotiations.

Hence, in principle, the measurement of HIs needs to extend to a large range of economic, social and political variables, the most relevant ones being specific to the context. In this spirit, Gurr has pointed to political, social, economic and demographic dimensions of relative deprivation, while Benabou has argued for the need to incorporate political rights in perspectives on inequality (Benabou, 2000). Gender inequality analyses almost invariably acknowledge this, exploring employment, education, nutrition, infant survival (UNDP, 1995; Croll, 2000), as well as a set of political variables as in the UNDP's Gender Empowerment Index (UNDP, 1995).

There are multiple elements of HIs within each of the four main dimensions:

- **Economic HIs** include inequalities in access to and ownership of assets—financial, human, natural resource-based and social. In addition, they comprise inequalities in income levels and employment opportunities, which depend on such assets and the general conditions of the economy.
- **Social HIs** include inequalities in access to a range of services, such as education, health care and housing, as well as in educational and health status.
- **Political HIs** include inequalities in the distribution of political opportunities and power among groups, including control over the the presidency, the cabinet, parliamentary assemblies, the bureaucracy, local and regional governments, , the army and the police. They also encompass inequalities in people's capabilities to participate politically and to express their needs.
- **Cultural status HIs** include disparities in the recognition and standing of different groups' languages, customs, norms and practices.

Evidence on the multiple dimensions (and elements) of HIs is essential because for policy one needs to know the source and nature of the HIs – for example, whether it is largely a problem of economic inequalities or also social and political. Moreover, for planning public expenditure one also needs to know which elements within the social and economic dimensions exhibit inequalities – for instance whether it is a matter of access to education, or poor economic opportunities or both; and whether access to land or employment shows the sharpest inequalities or are most felt by members of a group. Ideally, data should be collected on major inputs and outcomes in a wide range of areas, which, of course, involves extensive data requirements.

As noted earlier, some elements are more important in some societies, others elsewhere because of history, the structure of the economy and the stage of development. The aim should be to monitor the elements which are clearly key in the society in question. Moreover, because of strong correlations across the various elements and dimensions, getting information on one important element can be indicative of other inequalities. It is neither possible nor desirable to collect data on every aspect, but it is important to get an idea of HIs for some elements of every dimension.

Although data availability constrains which variables can be included in the short-term, certain elements that seem to be generally fundamental to well-being have priority for data collection. In the economic and social dimensions, these include: land and other assets; incomes; employment; educational access at various levels; health outcomes such as life expectancy;

and infant mortality. In political dimensions, members of the cabinet, the heads of departments and of the military and police are invariably of importance. For cultural status inequalities, language policy, the recognition given to the various religions and national holidays are generally all of relevance.

3.3 Data

In many cases, official statistics do not include data along group lines even for socio-economic variables, while there is almost never any readily available data for political or cultural status inequalities.

Socio-economic data: Such data are often sparse because ethno-cultural variables are not included in surveys, sometimes because of their political sensitivity – for example, in Nigeria ethnic data are not collected for this reason (Okolo 1999). Working with recipient countries to develop appropriate mapping tools forms an important part of the agenda in such countries, but for the quick assessment needed in fragile societies and post-conflict contexts, it may be necessary to take some other characteristic as a proxy for ethno-cultural difference. Two options are region and language.

Regional socio-economic data are often more readily available than data by ethnicity, from household surveys for example. Whether regional data are useful depends on how far identity groups are geographically segregated. As a rule of thumb, region is a useful proxy for measurement and policy if more than half the members of the deprived group are concentrated in the targeted region while less than half of the privileged group are in the targeted region. In many African countries, ethnic and religious groups *are* regionally concentrated so that regional inequality may be a suitable proxy for ethnic or religious inequality, and in some cases region itself defines group identities.

An alternative proxy for ethnicity is the use of a language variable, sometimes available where ethnic variables are not – as in Indonesian surveys in the New Order period. Comparison of language statistics from these surveys at the district level with ethnic data available at a later date shows close correlations, suggesting that language would be an effective proxy for ethnicity in this context (Mancini, 2005). But caution is needed here, as over time the language of the dominant groups tends to be adopted by subordinate groups: for example, in Sudan over several decades a number of marginalised groups have adopted Arabic as a lingua franca and even as a first language, due to a combination of deliberate policy by the regime and of an appreciation of the need to speak the dominant language in order to rise up the societal hierarchy. The same is true of indigenous language speaking in some Latin American countries.

In addition to these proxies, there are often particular surveys, and also censuses, that do include ethnic or religious variables. A priority, in any country, is to conduct an inventory of available data, and of urgent data needs. Imaginative use of existing data generally can contribute considerably to tracking HIs. Gaps may be filled by 'light' surveys including small surveys and use of focus groups.[6] Possible data sources include:

- Census data, where ethnic, language or religious data are often included, and sometimes religion. Appendix Table 1 shows all the developing countries in which at least one census from 1960 onwards asks questions on some group classifications – either ethnicity, or language or religion. Thirteen East and Southeast Asian countries are included, with questions about religion most commonly asked followed by language questions. In Sub-Saharan Africa as many as 34 countries ask about group membership, with questions about religion and ethnicity most common and language less so. In Latin America, censuses for 17 countries provide such data and here the language question is the most frequently asked, while in South Asia five countries have requested such information in at least one of their censuses, and the religious question is most common.

[6] Fuji (2006) discusses a variety of approaches to rapid poverty mapping.

- The Demographic and Health Surveys (DHS), many of which include ethnic identification, as well as information on access to social services and ownership of domestic assets. There are over 200 surveys carried out in 70 countries, they all contain regional data and about two-thirds specifically ask respondents about their ethnicity. Appendix Table 1 also shows the countries covered by DHS which include ethnic/religious/language questions. In some cases, the DHS questions are similar to those of the census, and in some they are much more extensive (e.g. the Philippines); while some countries have no such information in their censuses but do in the DHS surveys (e.g. the Cameroons).
- Living Standards Measurement Surveys (LSMS) which cover 31 countries and a range of socio-economic data, and sometimes include ethnic variables.
- Regional data from household surveys, the census and sometimes public expenditure accounts.
- Specific sectoral data (e.g. from schools and hospitals) which often contain ethnic and regional information.
- National Human Development Reports which generally have regional data and sometimes ethnic and religious too.

Political data: Information needed to assess political HIs includes the group distribution of positions in the cabinet, parliament, and bureaucracy, the army and police, and so forth. This requires knowledge of the background of the relevant politicians or officials. In a few cases this may be publicly available (for example, in Nepal, see Brown and Stewart, 2006; for Kenya see Kanyinga, 2007). But more generally it is not. In some contexts, political data can be collected through 'name recognition' techniques to attribute group background (see, for example, Langer, 2005). Moreover, where group distinctions are important, most informed political observers in a society can readily provide broad trends on important political positions. A key concept here is that of 'relative representation' (Langer, 2005), defined as each group's share of the positions available divided by its share of the population, which can be used as a measure of political HIs, though we should note that relative representation may be fair, yet certain groups remain disempowered as minorities where majority decision-making occurs. So for a true appreication it is necessary to go beyond the numbers to an understanding of the political system.

Political HIs are particularly relevant to societal stability, especially where they are in the same direction as socio-economic inequalities. So despite the difficulties in getting data, it is very important to get an understanding of the existence and dynamics of political HIs.

Information on cultural status inequalities: this is a matter of recording the cultural recognition given to the various groups; how far their holidays, buildings, languages and practices are respected, and changes over time in these aspects. Judgment enters here as much as numbers. Sources for information include the media and local academics.

Perceptions: The prime aim of this paper is to measure and monitor 'objective' HIs, but people's perceptions of inequalities ('subjective' HIs) are crucial for determining their socio-political impacts. Thus investigating such perceptions is a critical element of any study of group behaviour and mobilization (violent or otherwise). Moreover, the extent to which people's perceptions of HIs accurately reflect a country's 'objective' political and economic realities is an important question since distortions of realities can be brought about by political manipulation or the media and might be corrected by efforts to publicize the objective situation.

Surveys provide an appropriate way of obtaining data on perceptions of HIs. Such surveys should aim to obtain people's perceptions on the following issues (among others):

- Their own and other groups' socio-economic situation, progress and access to economic resources and networks;
- Their own and other groups' representation and positions in the most important political and administrative institutions;

- The treatment by the state, parastatal institutions and other groups of their own and other groups' cultural norms, practices and customs;
- The role of 'culturally' defined identities (such as ethnicity or religion) in gaining access to state employment, services and resources, and similarly with respect to the private sector;
- The extent to which the state (and other prominent actors) are seen as favouring or discriminating against their own and/or other groups in a variety of ways.

Developing aggregate measures of HIs

4

In developing an aggregate measure of HIs the aim is to arrive at a single figure which is a good representation of the entire distribution. While the question of how to aggregate HIs across society is similar to that of identifying a suitable measure of vertical inequality, there are some important differences. The most obvious is that, by definition, there are fewer groups in a society than individuals – in fact for many countries there may be only a small number of salient ethnic or religious groups. Indeed, in some contexts there are just two groups, and then a straightforward comparison of means is possible. Even with three or four groups, it is possible to look at the dyadic differences separately and aggregation into a single measure is not only less needed but may actually conceal important information. In contrast, however, when we are dealing with a large number of groups, aggregate measures of group inequalities can be helpful. For example, in Nigeria, there are more than 300 ethnic groups, though not all of them are political salient. We then need to move beyond simple comparisons of means.

A further complication in measuring inequality among groups is that the size of the group (the number of individuals in each group) generally differs. Hence it is necessary to decide whether to weight the measure by the size of the group. With an unweighted measure, the position of small groups would get the same weight as those of large groups. Yet from a well-being and a political perspective, this represents a different situation in that very different numbers of individuals are affected. Therefore, a population-weighted index is generally desirable.

A third difference between group and individual inequality is that each group is made up of a number of individuals – so the *intra-group* distribution may be of interest as well as the *inter-group.* It is possible to include a measure of within-group inequality in the measure of each group's performance. However, this would in effect conflate two separate issues – inter-group and intra-group disparities. Hence, in this section we consider how to measure HIs without consideration of the distribution within each group. In the next section we turn to issues of intra-group distribution.

Finally, because of the essentially multidimensional nature of HIs, the question arises of whether and how to amalgamate each dimension into a single index. This question also arises for measures of vertical inequality, and a number of methods for developing multidimensional indices have been developed (Bourguignon and Chakravarty, 2003; Deutsch and Silber, 2005). We do not develop a multidimensional index here, but we discuss the issue and propose some possible approaches in Section 6.

4.1 Some principles underlying measurement of inequality

Three well established general principles of a good measure of inequality have been developed for VI. These are:

1. independence of the distribution from the mean;
2. the principle of transfers (Pigou-Dalton): that is transfers from a richer person to a poorer person reduce inequality; and
3. the transfer of an equal amount from rich to poor counts for more in terms of reducing inequality than one from rich to less rich.

While the first two principles apply equally to a measure of group inequality, there does not seem to be good reason to accept the third principle.[7] However, for arriving at a useful measure of HI we would add two further principles:

4. Insofar as possible, the aim should be to find a measure which is descriptive, not evaluative so that the inequalities can be observed as such and not masked by the evaluative element. This is not perfectly achievable since any measure involves some implicit valuation, but we aim to minimize this and hence discard measures which have explicit normative elements, notably inequality aversion, built in; and
5. The measure should quantify group inequality as such, not the contribution of group inequality to either social welfare as a whole (like the gender-weighted Human Development Index), nor to (vertical) income distribution as a whole. The reason is that the aim is to assess the extent of group inequality and not the contribution of horizontal inequality to overall inequality; the latter will vary not only according to the extent of group inequalities but also according to other sources of overall vertical inequality.

Ratios of average achievements of pertinent groups are the most straightforward and intuitively appealing measures of group inequality (for example, the ratio of black to white per capita incomes in South Africa). However, such ratios only show inequality between two groups, and other measures are needed where there are a larger number. Østby (2008; 2003) deals with this problem by choosing the two largest groups and calculating the ratios for these groups. Another possibility would be to choose the two groups that seem to be politically competitive (not necessarily the two largest in population size) in the particular context.[8] However, this would impart a large element of political judgement into the choice. In general, both to assess how fair a society is and to test how far group inequality affects various objectives, there is a need for a synthetic measure which incorporates *all* group inequalities in any one variable into a single measure of HI. Nonetheless, we should bear in mind the possibility that the synthetic aggregate measure may be influenced by 'irrelevant' alternatives in the sense that it incorporates information about groups which may not be considered relevant comparators by other groups. Hence for some purposes, especially when the number of relevant groups is small, it is helpful to look at simple ratios of each group to the mean, and/or ratios of major groups to each other, as well as the synthetic measures to be discussed below.

4.2 Possible aggregate measures of horizontal inequalities

Three common measures of inequality, devised to measure VI, seem to be particularly appropriate for measurements of group inequality:

1. *The coefficient of variation* (COV), that is, the variance divided by the mean (see Box 1), which is a common measure of regional disparities (for example, Williamson, 1965; Quah, 1996). This measure attaches equal weight to redistributions at different income levels (violating the third principle above), but there is no reason to think this matters in relation to measuring HI. The COV gives more weight to the extremes, in a somewhat arbitrary fashion, because it involves squaring the deviations from the mean (Sen 1997: 28). As a measure of group inequality (GCOV), it can be weighted by the population size of each group as shown in Box 1. Without population weighting, changes in the position of a very small group would have the same effect as one involving a large group. Consequently, we recommend a population-weighted measure.
2. *Group Gini* (GGINI: Box 1). This compares every group with every other group, in contrast to the GCOV which measures differences from the mean, for each group. Generally, in using Gini as a measure of inequality, observations are grouped according to achievement

[7] Subramanian (2007) also argues that the principles adopted for measuring VI do not necessarily apply in relation to HI.

[8] This follows Posner's recommendation of considering only politically relevant groups (Posner, 2004).

Box 1 Measures of HIs (population weighted)

$$\text{GCOV}=\frac{1}{\bar{y}}\left(\sum_{r}^{R} p_r(\bar{y}_r-\bar{y})^2\right)^{\frac{1}{2}}$$

$$\text{GGINI}=\frac{1}{2\bar{y}}\sum_{r}^{R}\sum_{s}^{S} p_r p_s\left|\bar{y}_r-\bar{y}_s\right|$$

$$\text{GTHEIL}=\sum_{r}^{R} p_r \frac{\bar{y}_r}{\bar{y}}\log\left(\frac{\bar{y}_r}{\bar{y}}\right),$$

where $\bar{y}_r=\frac{1}{n_r}\sum_{i}^{n_r} y_{ir}$ is group r mean value; R is group r's population size; p_r is group r's population share; y_{ir} is the quantity of the variable of interest (e.g. income or years of education) of the i^{th} member of group r; Y_r is the value of y for group r; and Y is the grand total of variable y in the sample.

on the variable of interest. However, for HIs, we group people by 'non-economic' characteristics (religion, ethnicity, race, and so on) and not by the variable for which the inequality is being calculated.

3. Finally, the *Theil index* (GTHEIL: Box 1), which is especially sensitive to the lower end of the distribution, can be used. It also compares each group with the mean (like the GCOV) in contrast to the group Gini. Unlike the Gini coefficient, the Theil is always precisely decomposable and is often used to divide overall VI into inequality due to within-group inequality (WG) and that due to between-group (BG) (Anand, 1983; Heshmati, 2004). Some researchers have questioned the intuitive appeal of the Theil inequality index (Sen, 1992). However, as far as group inequality is concerned, no intuitive understanding has built up of the group Gini either, because of the lack of experience in interpreting it. The group Theil represents a reasonable alternative to the two previous measures for measuring group inequality.

Other measures of inequality are utility-based with normative implications – including some element of inequality aversion in the measure – for example, Dalton's and Atkinson's index (Atkinson 1970). They are not suitable for measuring HIs because of the strong element of evaluation, thereby violating our fourth principle above.

There are also some measures which have been specifically designed to explore aspects of group inequality:

4. The Esteban/Ray (ER) polarization index (ER(1.5): Box 2) (Esteban and Ray, 1994; Duclos *et al.,* 2004). The ER index is similar to a group Gini[9], weighted by population, but includes

Box 2 Other proposed measures of group inequality

Estaban-Ray polarisation index:

$$\text{ER}(k,\alpha)=k\sum_{r}^{R}\sum_{s}^{S} p_r^{1+\alpha} p_s\left|\bar{y}_r-\bar{y}_s\right|,\ k=\frac{1}{\bar{y}},\ \alpha=1.5$$

Zhang and Kanbur index: ZK=BG/WG, $\text{WG}=\sum_{r}^{R}\frac{\bar{Y}_r}{Y}\left[\frac{1}{n_r}\sum_{i}^{n_r}\frac{y_{ir}}{\bar{y}_r}\log\left(\frac{y_{ir}}{\bar{y}_r}\right)\right]$,

[BG=THEIL–WG].

where BG = between group inequality, WG = within group inequality, and α is a scalar, and the other variables are the same as Box 1.

[9] Where $\alpha = 0$, this is equivalent to the group Gini.

an index ($1+\alpha$, where $1 \leq \alpha < 1.6$) which is higher the greater the weight attached to the share of group population. A common value for α is 1.5 (Kanbur and Zhang, 1999; Zhang and Kanbur, 2003). This index is not decomposable. The point of α is to increase the weight given to large groups, so that the index rises as the population is distributed among fewer and more equally sized groups. Consequently, two populations might have the same value of the index, despite one having less variance in resource access or incomes between the groups than the other, so long as the one with less variance had a smaller number of larger groups than the other. Esteban and Ray do not intend the index to be a measure of group inequality but of societal polarization. The disadvantage of the measure from our perspective is first, that it includes two elements that we wish to explore separately – the demographic polarization of the groups and the extent of inequality among them – and second, that there is an arbitrariness about the choice of α. The ER measure also violates the Pigou-Dalton condition because increased demographic polarization can offset a given income transfer from a richer to a poorer group;

5. Zhang and Kanbur have proposed an index (ZK: Box 2)) which is the ratio of BG/WG, using Theil to measure BG and WG (Zhang and Kanbur 2003). This is higher the more HIs contribute to overall inequality. This violates the fifth principle we put forward – i.e. the size of the measure will vary according to within-group variance. Thus the same between-group variance will lead to different measures according to WG. Consequently, this measure is not suitable for quantifying HIs.

6. Chakrabarty (2001) has proposed an odds ratio as a way of measuring group inequality. This measure calculates the odds of individuals in a particular group falling into a particular category (for example, rich or poor), and then expresses group inequality as the ratio of these odds. This basically resembles the method of using simple ratios of performance, except that the performance of each group is not the average but the odds of being poor, calculated as a proportion of the poor to the total population in each group. Hence it requires somewhat more data than the ratios of averages. Since it is designed to measure differences between two groups, some other method would need to be introduced to generalize to many groups (for example, by adopting a Gini of the odds).

Much work in the area of quantifying group inequality has tended to be less interested in measuring group inequality as such, and more interested in devising a measure of general welfare that accounts for group inequality – for example, Anand and Sen (1995) devised a gender-weighted Human Development Index (the Gender Development Index or GDI). This is a social welfare evaluation for society as a whole, which weights trade-offs between higher average achievement and more inequality between genders. Similarly, Majumdar and Subramanian (2001) adjust a measure of deprivation, or capability failure, by a group deprivation index. The capability index is a weighted index of several deprivations. The index is adjusted by a formula for the deprivation of the particular group. Again the aim is to adjust aggregate welfare by some valuation of group disparities rather than to describe group inequality. These approaches are not suitable for measuring HIs because they violate principle 5 above, i.e. they produce measures of societal welfare adjusted for group inequality but not direct measures of HIs.

4.3 Conclusions on approaches to measuring group inequality where there are more than two groups

First, given that we aim – as far as possible – to arrive at a descriptive measure of His, we reject those measures which involve a strong explicitly evaluative element, that is, the Dalton/Atkinson measures. This is also a problem with the ER polarization index. Nonetheless, the other measures also, unavoidably, contain some element of valuation – for example, the GCOV by squaring the observations gives more weight to observations further from the mean and the Gini gives relatively more weight to the middle of the distribution.

Secondly, we aim to arrive at a measure of HI which is distinct from other influences such as VI and population distribution. This means that neither the ER nor the ZK index are suitable. The ER index combines two elements: inequality and population polarization. Hence the same distance between groups would get greater weight the more the population is demographically polarized. The ZK index incorporates both BG and WG inequality into the index. Thus the same HIs would get different values according to the extent of WG. From our perspective this is not desirable, as our aim is to find a measure that will help to monitor group inequalities as such; and in empirical work to test what impact this has on various outcomes. While the impact of HIs may vary according to the extent of heterogeneity within the group or the extent of demographic polarization, we wish to test both these elements separately, which is not possible if they are incorporated in a single index. Similarly, for policy work it is the extent of actual HIs which is relevant, and this is disguised by these measures. Thirdly, in addition to simple ratios between two groups, it is useful to have a measure which captures horizontal inequality in a society with several groups. Hence the odds ratio by itself is insufficient, though it could provide the inputs into another measure, such as a group Gini, which aggregates across a number of groups.

This leaves us with the first three measures discussed above, each of which can be weighted according to the size of the groups' population shares:

- The coefficient of variation among groups (GCOV);
- The Group Gini (GGINI) coefficient; and
- The Group Theil (GTHEIL).

4.4 Applications of the aggregate measures: South Africa and Indonesia

This section explores how the three chosen measures of inequality behave in practice on the basis of data from South Africa and Indonesia for both VI and HI. This gives insight, first, into how far it matters which measure we use; and secondly whether one or other measure seems to fit an intuitive understanding of the situation.

Longitudinal income data from South Africa (Figure 1) is interesting because it shows the importance of weighting by group population share. The unweighted GCOV among racial groups shows a rise in HI since 1975, while the population weighted GCOV, and the GGINI each show a trend fall. The surprising rising trend in the unweighted GCOV can be attributed to a particular feature of the trends in these data. Over the period, the mean income of all four population groups increased at a faster rate than the mean of the population as a whole, due to the concomitant increase in population share of the poorest group, black South Africans, which effectively dragged down the rate of increase in the overall mean. While the various weighted measures compensate for this unusual trend precisely because they incorporate a degree of population weighting, the unweighted GCOV treats every group as static over time in terms of size. The increase in the unweighted GCOV is counterintuitive: the ratio of black/white income *per capita* fell significantly.

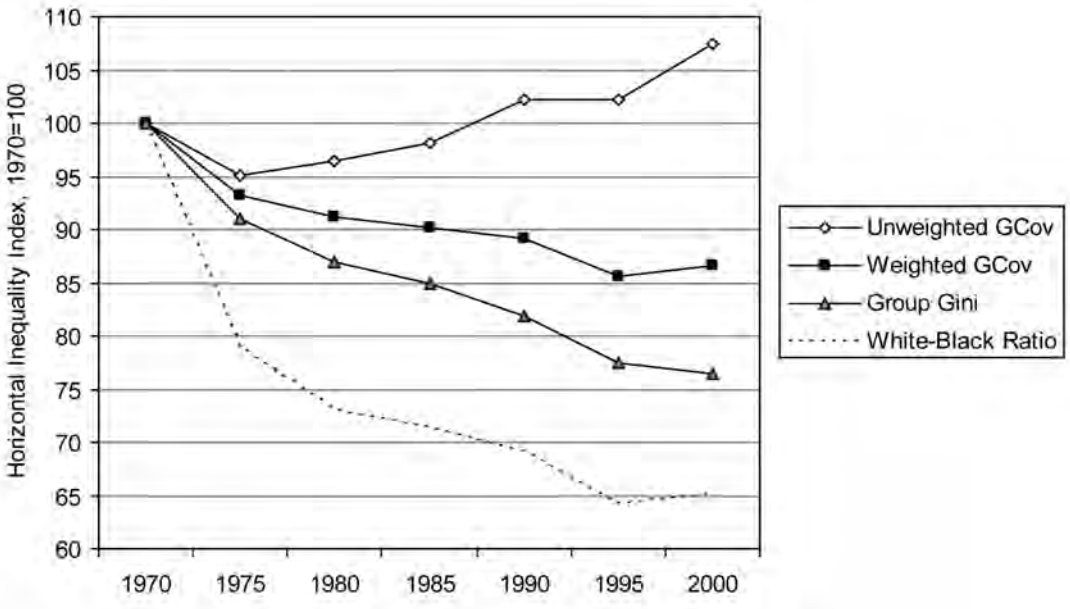
Data for educational inequalities from religiously diverse districts in Indonesia for 1990 and 1995 illustrates the behaviour of different measures of VI and HI.[10] Table 1 shows the correlation between measures of VI and HI based on years of education for 88 *religiously diverse* districts in Indonesia.[11] The numbers on the diagonal (in bold) show the correlation of each indicator with its own value in the two years. The following results emerge:

- The indices of vertical inequality – Gini, Theil and the coefficient of variation (COV) – are nearly perfectly correlated in both years, indicating that for this dataset it is not important which measure is selected;

[10] The 1995 dataset contains 304 rather than 282 districts because the province of East Timor was excluded from the 1990 dataset due to undersampling.

[11] Groups are defined by religion, which in both samples is categorized as: Islam, Catholic, Protestant/Other Christian, Hindu, Buddhist and Other.

Figure 1 South Africa – Aggregate measures of horizontal inequality in per capita income by ethnic group, 1970–2000



Source: Calculated from data in van den Berg and Louw (2004).

Note: Louw and van der Berg give two separate estimates for the year 2000 – one 'optimistic', the other 'pessimistic'. Here, the mid-way point between the two estimates was used.

- HI measures are also highly correlated with each other in both years, with coefficients ranging between 0.76 and 0.96 – again suggesting that the choice of group inequality indicator may not be important;
- The correlations between measures of VI and HI are much smaller than the correlations within either the VI or HI measures. In Table 1, there is generally a positive correlation between the VI and HI measures. Table 1, however, excludes ethnically/religiously homogeneous districts. When these are included (not shown) the correlation between each measure of HI and VI is negative, and generally statistically non-significant at standard levels. It is relevant to include homogeneous districts if one wants to explore how far a measure of VI can proxy for HI in society as a whole. However, given that our main purpose is to contrast measures of HIs, it seems more relevant to confine the data to multi-group districts.
- In this data set HI measures tend to be less correlated over time than VI indicators, although all show a quite high correlation over time. The difference in correlations over time between VI and HI measures implies that HI and VI dynamics differed across Indonesian districts over the five-year period. From one perspective this is surprising – if group inequality stemmed from discrimination or asset inequality one would not expect this to change quickly while individuals might move more readily up or down the income scale. However, the very large numbers involved in VI mean that upward and downward changes are more likely to offset each other than in the case of HI which is based on small numbers of groups. The data on South African white/black income ratios shown above (Figure 1) similarly shows much more change than VI in South Africa over the same period.

Carrying out a similar exercise with income data for 154 religiously diverse districts (Table 2) in 1995 broadly confirms the results of Table 1. Again, there are quite high correlations among alternative measures of HI, but much lower correlations between HI and VI.

Table 1 Correlation matrix between inequality measures in years of education in Indonesia 1990 and 1995 in religiously diverse districts

		VERTICAL INEQUALITY						HORIZONTAL INEQUALITY					
	YEAR	GINI		THEIL		COV		GGINI		GTHEIL		GCOV	
VI													
GINI		0.87	***										
THEIL	1995	0.99	***	0.83	***								
	1990	0.98	***										
COV	1995	0.99	***	0.99	***	0.86	***						
	1990	0.99	***	0.99	***								
HI													
GGINI	1995	0.40	***	0.42	***	0.40	***	0.81	***				
	1990	0.41	***	0.36	***	0.39	***						
GTHEIL	1995	0.52	***	0.54	***	0.52	***	0.91	***	0.77	***		
	1990	0.43	***	0.39	***	0.41	***	0.91	***				
GCOV	1995	0.46	***	0.47	***	0.45	***	0.93	***	0.91	***	0.74	***
	1990	0.47	***	0.42	***	0.44	***	0.95	***	0.91	***		

Notes:
88 religiously diverse districts (ERF>=0.1).
***, **, * indicate significance at 1%, 5%, and 10% level, respectively.
The diagonal elements shown in bold for each index are the correlation coefficients of each indicator over time, i.e. between 1990 and 1995.

Source: Indonesian Census 1990; 1995.

Table 2 Correlation matrix between inequality measures in per capita income in Indonesia 1995 in religiously diverse districts

	VERTICAL INEQUALITY						HORIZONTAL INEQUALITY				
	GINI		THEIL		COV		GGINI		GTHEIL		GCOV
VI											
GINI	1.00										
THEIL	0.94	***	1.00								
COV	0.81	***	0.94	***	1.00						
HI											
GGINI	0.25	***	0.34	***	0.40	***	1.00				
GTHEIL	0.36	***	0.49	***	0.56	***	0.80	***	1.00		
GCOV	0.37	***	0.46	***	0.52	***	0.86	***	0.84	***	1.00

Notes:
154 religiously diverse districts are considered (ERF>=0.1).
***, **, * indicate significance at 1%, 5%, and 10% level, respectively.

Source: Indonesian Census 1995.

5

Measuring HIs in different segments of the income distribution

The performance of a group includes the *whole* distribution of the group (i.e. the performance of every member). When this is represented by the average per capita performance, distributional differences within groups are concealed. Yet from political and policy design perspectives, how groups compare at different points in the distribution is relevant. The following cases need to be distinguished:

- Case 1, where one group outperforms another at every income level;
- Case 2 where a group's elite (e.g. top 5%) has incomes (or other measure of welfare or resource access) far higher than the elite of the other group, but the remaining 95% of the distribution have similar income levels;
- Case 3, where both groups are broadly equal at the top, but the bottom 40% of one group is far poorer than the bottom 40% of the other;
- Case 4 where one group has higher income levels at the top of the distribution, but lower at the bottom.

These differences have both political and policy implications. From a political perspective, in the first case, both the elite and the masses have grievances and so rebellion may be likely. Malaysia in the 1960s and Apartheid South Africa are examples. In the second case, the elite has a grievance but not the rest of the population which makes rebellion less likely.[12] In the third and fourth cases, the biggest grievances will be among the lower-income groups – but lacking leadership they may not rebel, unless the elite are politically disenfranchised or threatened, in which case they will find it easy to mobilise support among the lower-income groups – the Rwandan situation in 1994 is an example.

The differences are relevant from a policy perspective too. In the first case, policies are needed both to increase opportunities at the top (in capital ownership, entrepreneurial activities, and senior civil service positions) and to provide basic services and economic opportunities to the masses; in the second case, focus should be on elite opportunities, while in the third and fourth cases it is a matter of improving the position of the worst off.

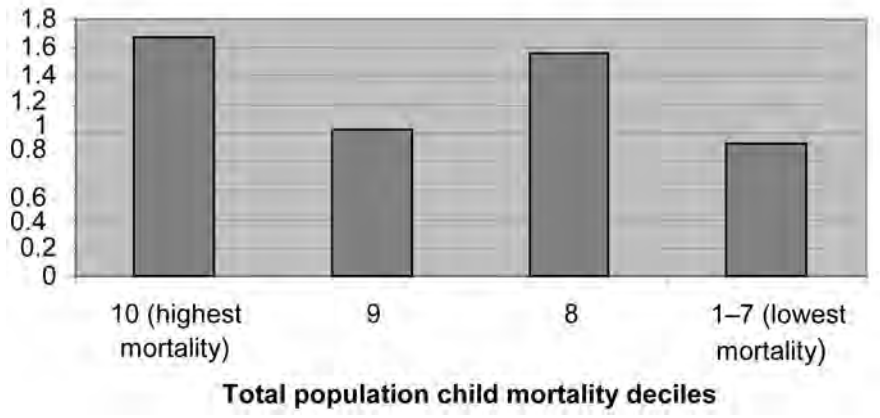
Comparisons of the whole distribution of each group are therefore needed. A simple approach to making such a comparison consists in dividing the income distribution of each group into *n* quantiles and comparing group means for each quantile (to be described as *q*-means henceforth). This approach allows us to see how the richest quantile of one group compares with that of other, and similarly for the other quantiles. Figure 3 provides an example for child mortality for a region in Indonesia. In this case, Muslims have higher mortality than Christians among the worst three deciles, but slightly better mortality than Christians among the better off 70%.

However, a more comprehensive way of comparing HIs across distributions is offered by a general means approach, derived from the work of James Foster.[13] The formula for this is presented in Box 3. Basically, it calculates group means for each group at different points of the income distribution, using parametric means. The value of the parameter, α, determines how much weight is given to different sections of the distribution. Hence the estimate of HI varies according to the chosen value of the parameter. Where α =1, the measure is equiva-

[12] Calabar in Nigeria is an example where the elites of the Quas and the Efuts perceived themselves disadvantaged compared with the Efiks in political appointments and business opportunities, but the mass of the population felt they were equally treated with respect to employment and amenities. No conflict occurred. In neighbouring Warri, where there were both elite and popular perception of HIs, there have been recurrent violent conflicts (Ukiwo, 2006).

[13] See Foster, Lopez-Calva, and Szekely (2003) for details. Foster *et al.* define these parametric means as 'general means'.

Figure 2 Relative representation of Muslims vis-à-vis Christians in Central Sulawesi, Indonesia



lent to a normal mean.[14] For values of, $\alpha < 1$, α-means are more sensitive to population performance at *low values* of achievement, and increasingly so as α becomes smaller. For values of $\alpha > 1$, more weight is given to *higher achievements*, again increasing the higher the value of α, and where $\alpha=1$ (the arithmetic mean) the value is most sensitive to median incomes. Comparing α-means for different values of α thus indicates how HIs differ in different parts of the income distributions of the two groups.

We illustrate this approach using income data from the 1996 and 2001 South African censuses (Figure 3).[15]

Box 3 The α-means approach to measuring HIs across different segments of the groups' distributions

α-means:

$$\begin{cases} \overline{y}_{r\alpha} = \left(\sum_{i}^{n} y_{ir}^{\alpha} \Big/ n \right)^{\frac{1}{\alpha}} & \alpha \neq 0, \alpha \in \Re \\ \overline{y}_{r\alpha} = \exp\left(\dfrac{\sum_{i}^{n} \ln\left(y_{ir}\right)}{n} \right) & \alpha = 0 \end{cases}$$

where $\overline{y}_{r} = \frac{1}{n_r} \sum_{i}^{n_r} y_{ir}$ is the mean value of *y*, which is a measure of the element of interest (e.g. income) of group *r*, p_r is the population share of group *r*, y_{ir} is the value of *y* for the i^{th} member of group *r*, Y_r is the total value of *y* for group *r* as a whole, *Y* is the grand total value of *y* in the sample/population, and α is a scalar.

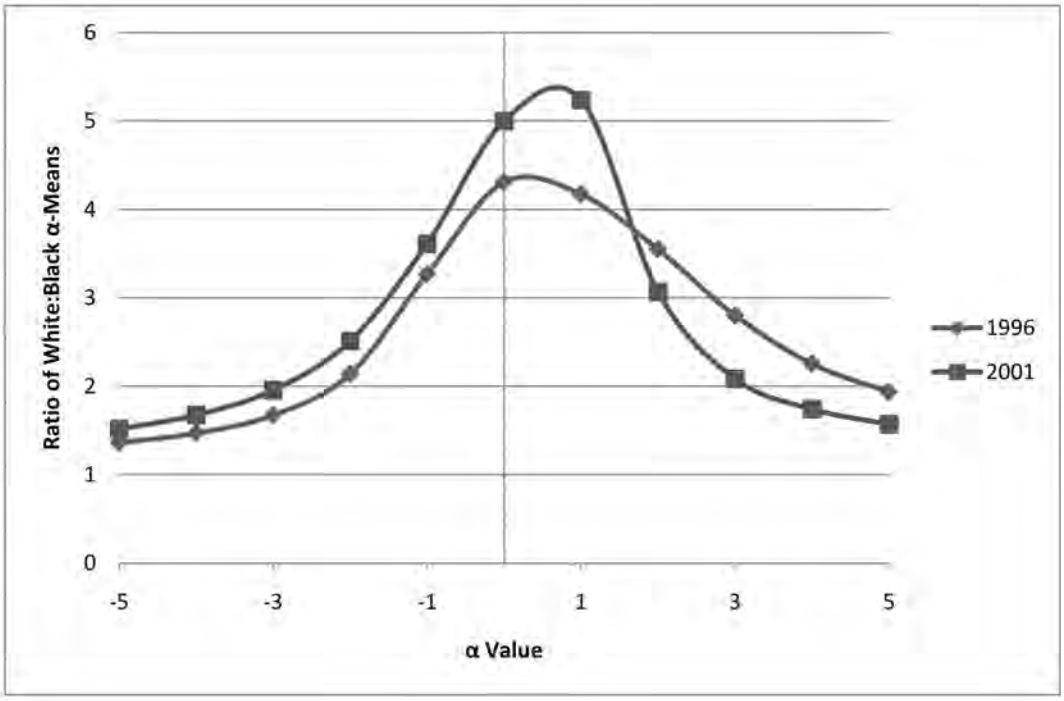
What Figure 3 shows is that differences between blacks and whites are highest in the middle of the distribution, while they are less for both the rich and the poor. This is true both in 1996 and 2001. Comparing 2001 and 1996 shows that the gap between rich blacks and rich whites has lessened, while that between poor blacks and poor whites has widened. Taking the normal mean (α =1), the gap has narrowed as shown in Figure 1.

The main difference between the *q*-means and the α-means approaches is that whereas *q*-means are calculated on sections of the group's income distribution, α-means are calculated

[14] (α = 1) is the geometric mean; and (α = -1) is the harmonic mean.

[15] Data obtained from the Integrated Public Use Microdata Series (IPUMS) at the Minnesota Population Center, University of Minnesota.

Figure 3 Comparison of White:Black α-means ratios, South Africa 1996 and 2001



over the entire range of the group's distribution but with different portions weighted differently according to the value of α. It is in principle possible to set α such that the difference between α-means and *q*-means becomes negligible.

Measuring multidimensionality

6

The multidimensionality of HIs naturally leads to the question of whether it would be desirable to develop a multidimensional index of HIs. There are several reasons why we do *not* think this is desirable in general:

- First, because in general, people perceive inequality in terms of specific elements (such as housing or employment) and not in terms of an aggregate measure;
- Secondly, policy also needs to be directed towards specific elements and for this it is necessary to have information on the individual HIs. For example, if educational HIs are small, but employment ones are large, then policy needs to be designed to correct the employment HIs not the educational ones, but this would be concealed in a multidimensional index.
- Thirdly, a general research finding is that HIs in different dimensions (e.g. economic or political) have different implications for social stability, and it is relevant also whether the various inequalities go in consistent directions.[16] Again this would be concealed by a multidimensional indicator.
- Fourthly, each of the elements within any dimension is measured in different ways and with different metrics and these differences are even more marked across dimensions. When one is looking at one indicator on its own, then the different metrics can be interpreted intuitively. Of course, data can be standardised, but this conceals differences in the measurements and in the interpretation of their significance, it does not eliminate them. Even different ways of measuring the *same* indicator can lead to differences in HIs. For example, HIs in literacy are generally much lower than HIs in illiteracy. Yet in order to arrive at multidimensional indices, it is necessary to impart unique values to the HI on each element/dimension.
- Fifthly, inequality in different dimensions might be of different normative importance; hence different inequality measures would be appropriate for each dimension. For example in health we may be concerned with relative inequality;[17] in income space we may wish to incorporate inequality aversion in the measure, as in the Atkinson measure; and in literacy we may be concerned with absolute inequality. While it is possible to construct such inequality measures, it is difficult to interpret them or use such multidimensional measures as a guide to policy.

Thus there is not a case for developing a multidimensional indicator for inequality across the four main dimensions, which are fundamentally incommensurate. Yet each dimension contains a number of elements. Given scarce data, there is often only information for one or two elements within any dimension, so the issue of developing a synthetic measure may not arise. Nonetheless, where there is more extensive information, it may be helpful sometimes to have a single indicator for the various elements *within each dimension,* so as to be able to see whether one group is broadly better off in the dimension, e.g. in economic terms, or in political or cultural terms.

Two approaches to developing indicators to summarise performance across the various elements within each dimension are worth exploring:

1. First, where perceptions' surveys are available or where they can be conducted, a question on whether people perceive their group to be better off, equal to or worse off than

[16] One finding is that the risk of violent conflict is greater if political and economic HIs are consistent, with one group being favoured in both economic and political terms and the other(s) disfavoured (Stewart, 2008).

[17] For which a Kolm measure would be appropriate (Vranken *et al.,* 2001; Atkinson, 2009).

other groups in general in the economic realm provides a general measure of subjective HIs in this area. Similar questions can help make assessments in the social, political and cultural dimensions.

2. Secondly, Alkire and Foster's (2008) proposed method for aggregating across dimensions in poverty measures has potential for HIs. Their method essentially involves two steps. First defining the poverty line for each dimension; and then defining the multidimensional incidence of poverty as being the percentage of people who fall below the specified poverty lines on some selected number of dimensions. For example, suppose there are three poverty dimensions – health, income and education – and the selected cutoff rate for multidimensional poverty is to be poor in at least two dimensions. Then a person who is poor in any two dimensions would count as poor and the percentage of the population who meet this criteria would constitute the headcount poverty rate. A similar approach to HIs could be adopted. First, it would be necessary to specify a level of HIs, for any element, which would define a group as being deprived/rich – for example, 25% below or 25% above the average. Thus any group which was 25% below the average in any element would be relatively deprived. It would then be possible to show whether a group faced *consistent* deprivation within a dimension, or was deprived in some elements but not in others. Following Alkire and Foster, a group could be defined as experiencing HIs only if it was deprived in a certain proportion (say more than half) the elements measured. However, while this would indicate the extensiveness of the HIs across elements, it would not allow for the severity in any of the elements. Severity could be explored by increasing the cut-off line for defining HIs in any element – say to 50% below the average.[18] Of course, it would also be possible to develop an index which combined both a measure of the severity of HIs and their extensiveness, but again the weighting would be arbitrary.

   In general, in the many cases where there are rather few salient groups and limited data, presenting the information on the range of available data is more informative than presenting a synthetic measure of it.

[18] Alkire and Foster propose a measure, M1, which combines 'breadth' of poverty with 'depth' of poverty. This basically multiplies the headcount poverty, H by the average weighted sum of dimensions in which people are deprived A, times the average shortfall from the poverty line for all deprivations (normalized),G. (Alkire and Foster, 2008, plus personal communication from Sabina Alkire)

Some conclusions

7

Measurement of HIs is important because HIs affect the wellbeing of members of a group, and can be adverse to efficiency, growth and poverty-reduction as well as raising the risk of violent conflict. Yet for the most part, measurement of HIs does not form a routine part of international or national data gathering and reporting. This paper has considered alternative approaches to measurement of HIs.

We acknowledge the severe problem of defining group boundaries, since identities are fluid, multiple and may even be endogenous. Nonetheless, felt differences seem sufficiently clear in many societies to make it possible to measure group performance so long as one is sensitive to the possibility and implications of alterations in group boundaries. A further chronic problem is deficient data but in most multiethnic or religious countries, some useful data are available as well as some international data sources. However, there is a clear need to improve the data situation on HIs.

A review of alternative aggregate measures of group inequality found that the GCOV, the GGini and the GTheil, each with population weighting, all provide suitable measures. Data on South Africa showed that including population weighting can be important, as the trend change in HI differed in this case between the measure that included population weighting and the one that did not. Logic as well as intuitive interpretation suggests that the population-weighted measures are preferable. The population-weighted GCOV and the group Gini measures generally move in the same direction, with the GCOV showing the least change.

Evidence from Indonesia showed quite a high correlation across the three measures among religiously diverse districts, so which is chosen may not matter much. There is low correlation, however, between measures of HI and measures of VI and none if homogeneous districts are included, indicating that VI is *not* an adequate measure of HI. Both VI and HI show persistence over time. But for education, within districts, HI changed more over time than VI, which could be important from the perspective of political stability since durable HI is likely to be particularly provocative.

If one is particularly concerned with the position of one particular group, or of that one group relative to another group, the clearest way to present the data is to take the ratio of the performance of the group to the mean, or the ratio of the performance of the two groups, rather than some aggregate societal group measure of inequality which also introduces other groups and may thus conceal the inequalities of interest.

We also explored alternative methods of assessing HIs allowing for the whole distribution within each group since the political and policy implications differ according to where the inequality lies within the distribution. We found that the α-means measure developed by James Foster provided the most instructive summary of the way HIs alter as one alters the weights given to different parts of the distribution. However, a comparison of HIs in different quantiles of the distribution is intuitively appealing and straightforward to present.

The value of a measure of inequality depends on the purpose for which it is needed. One purpose is to make a general statement about whether a society is getting fairer or less fair over time from a group perspective. Another objective is to identify a variable which will enable us to test whether particular inequalities are correlated with other events, such as conflict, criminality, and unhappiness. For both purposes, group weighting by population would seem desirable, adopting an α-means approach. However, it could be argued that the first objective might require measures which are widely understood – which could be a problem for

these somewhat complex measures. From the perspective of assessing how fair a society is, the measure of inequality could include some element of evaluation (as in some measures of inequality that have been proposed – e.g. the Estaban-Ray polarisation measure). However, unless this valuation is widely understood and shared, it is better to present the data without such a valuation element.

For the second objective – to identify a variable we can use to explore consequences of HIs – it is preferable to exclude elements of evaluation as far as possible, and to have a separate measure of HI and VI. The population-weighted GCOV or the group Gini appear to be the preferred measures for this objective. These measures can be used in combination with the α-means measure.

A common question is how VI and HI are related. HI is a component of aggregate societal vertical inequality, although it generally forms only a small component. However, there is no reason for any correlation between vertical inequality within a particular group and inter-group inequality. From this perspective one would expect a small positive correlation between the measures of HI and measures of VI. This is shown in the Indonesian data, but it also shows that measures of VI are not at all adequate proxies of VI, indicating the importance of arriving at independent measures of HI. The common practice of measuring between group inequality in terms of the contribution group inequality makes to total societal vertical inequality is highly misleading from the perspective of assessing the severity of HIs. Hence the need to measure HIs independently as explored in this paper.

References

Akerlof, G. A., and Kranton, R. E. 2000. 'Economics and Identity'. *The Quarterly Journal of Economics* cxv(3): 715-53.

Alkire, S., and Foster, J. 2008. 'Counting and multidimensional poverty measurement'. Oxford: OPHI.

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

Anand, S., and Sen, A. 1995. 'Gender inequality in human development: theories and Measurement'. *Human Development Report Office Occasional Paper.* New York: UNDP.

Atkinson, A. B. 1970. 'On the measurement of inequality'. *Journal of Economic Theory* 2(3).

Atkinson, A.B. 2009 'Income, health and Multidimensionality' in de Boyser, K., Dewilde, C., Dierck, D. and Friedrichs, J. (eds) *Between the Social and the Spatial: Exploring the Multiple Dimensions of Poverty and Social Exclusion.* Farnham, Surrey: Ashgate.

Badgett, M. V. L., and Hartmann H. L. 1995. 'The effectiveness of equal employment opportunity policies' in Simms, M. C. (ed.) *Economic Perspectives on Affirmative Action.* University Press of America.

Benabou, R. 2000. 'Unequal societies: income distribution and the social contract'. *American Economic Review* 90(1): 96-129.

Blau, P.M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure.* New York: Free Press.

Bourguignon, F., and Chakravarty, S.R. 2003. The measurement of multidimensional poverty. *Journal of Economic Inequality* 1(1): 25-49.

Broman, C. 1997. 'Race-related factors and life satisfaction among African Americans.' *Journal of Black Psychology* 23(1): 36-49.

Brown, T.N., Williams, D.R., Jackson, J.S., Neighbours, H., Sellers, S., Myriam, T., and Brown, K. 1999. 'Being black and feeling blue: mental health consequences of racial discrimination'. *Race and Society* 2(2): 117-131.

Brown, G., and Stewart, F. 2006. 'The implications of Horizontal Inequality for aid'. *CRISE Working Paper No. 26.* Oxford: Centre for Research on Inequality, Human Secrurity and Ethnicity, University of Oxford.

Caumartin, C. 2005. 'Racism, violence and inequality: An overview of the Guatemalan case'. *CRISE Working Paper No. 11.* Oxford: Centre for Research on Inequality, Human Security and Ethnicity, University of Oxford.

Chakrabarty, A. 2001. 'The concept and measurement of group inequality'. *Centre for Development Studies Working Paper No. 315.* Thiruvananthapuram: Centre for Development Studies.

Croll, E. 2000. *Endangered Daughters: Discrimination and Development in Asia.* London: Routledge.

Dalton, H. 1920. The measurement of the inequality of incomes. *Economic Journal* 30(119): 348-61.

Deutsch, J., and Silber, J. 2005. 'Measuring multidimensional poverty: An empirical comparison of various approaches'. *Review of Income and Wealth* 51(1): 145-174.

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

Estaban, J.-M., and Ray, D. 1994. 'On the measurement of polarisation'. *Econometrica* 62(4): 819-51.

Esteban, J.-M., and Ray, D. 1998. 'Conflict and distribution'. *Journal of Economic Theory* 87(2): 379-415.

Foster, J.E., and Szekely, M. 2006. 'Is economic growth good for the poor? Tracking low income using general means'. *Mimeo.* Vanderbilt University.

Foster, J. E., Lopez-Calva, L. F., and Szekely, M. 2003. 'Measuring the distribution of Human Development: Methodology and an application to Mexico'. *Estudios Sobre Desarrollo Humano,* PNUD Mexico, No. 2003-4.

Fuji, T. 2006. 'How Well Can We Target Resources with "Quick-and-Dirty" data? Empirical Results from Cambodia'. *SMY Economics and Statistics Working Paper* 03-2006. Singapore Management University.

Gurr, T. R. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts.* Washington, DC: United States Institute of Peace Press.

Gurr, T.R., and Moore, W. H. 1997. 'Ethnopolitical rebellion: a cross-sectional analysis of the 1980s with risk assessments for the 1990s'. *American Journal of Political Science* 41(4): 1079-103.

Hafeez, S., UNICEF, & South Asian Association for Regional Cooperation. 1990. *The girl child in Pakistan: priority concerns.* Islamabad.

Heshmati, A. 2004. 'Regional inequality in selected large countries'. *Discussion Paper Series* No. 1307. Berlin: Institute for the Study of Labor.

Justino, P., Litchfiled, J. and Niimi, Y. 2004. 'Multidimensional inequality: An empirical application to Brazil'. *PRUS Working paper No. 24.* Brighton: Poverty Research Unit at Sussex, University of Sussex

Kanyinga, K. 2007. 'Governance institutions and inequality' in Society for International Development (SID) *Readings on inequality in Kenya.* Nairobi: SID.

Kanbur, R., and Zhang, X. 1999. 'Which regional inequality? The evolution of rural-urban and inland-coastal inequality in China from 1983 to 1995'. *Journal of Comparative Economics* 27(4): 686-701.

Langer, A. 2005. 'Horizontal inequalities and violent group mobilization in Côte d'Ivoire'. *Oxford Development Studies* 33(1): 25–45.

Macours, K. 2004. 'Ethnic Divisions, Contract Choice and Search Costs in the Guatemalan Land Rental Market'. Washington DC: Paul H. Nitze School of Advanced International Studies (SAIS) Johns Hopkins University

Majumdar, M., and Subramanian, S. 2001. 'Capability failure and group disparities: some evidence from India for the 1980s'. *Journal of Development Studies* 37(5): 104-40.

Mancini, L. 2005. 'Horizontal inequalities and communal violence: Evidence from Indonesian Districts'. *CRISE Working Paper* No. 22. Oxford: Centre for Research on Inequality, Human Secrurity and Ethnicity, University of Oxford.

Murshed, S., and Gates, S. 2005. 'Spatial-horizontal inequality and the Maoist insurgency in Nepal'. *Review of Development Economics* 9(1):121-34.

Okolo, A. 1999. 'The Nigerian census: Problems and prospects'. *The American Statistician* 53(4): 312-25.

Østby, G. 2008. 'Inequalities, the Political Environment and Civil Conflict: Evidence from 55 Developing Countries' in Stewart, F. (ed.) *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* Basingstoke: Palgrave Macmillan.

Østby, G. 2003. 'Horizontal inequalities and civil war'. *MA Thesis,* Norwegian University of Science and Technology, Oslo.

Papatheodorou, C. (2000) 'Decomposing Inequality in Greece: Results and Policy Implications'. *Distributional Analysis Research Programme Papers* No. 49. London: STICERD.

Posner, D. N. 2004 'Measuring ethnic fractionalization in Africa'. *American Journal of Political Science* 48(4): 849-63.

Pradhan, M., Sahn, D., and Younger, S. 2001. 'Decomposing World Health Inequality'. *Tinbergen Institute Discussion Papers* No. 01-091/2. Tinbergen Institute.

Quah, D. 1996. 'Regional convergences clusters across Europe'. *Centre for Economic Performance Discussion Paper No. 0274.* London: Centre for Economic Performance.

Sen, A. 1980. 'Equality of what?' in *Tanner Lectures on Human Values.* Cambridge: Cambridge University Press.

Sen, A. 1992. *Inequality Reexamined.* Oxford: Clarendon Press.

Sen, A. 1997. *On Economic Inequality.* Oxford: Clarendon Press.

Stewart, F. 2000. 'The root causes of humanitarian emergencies' in Nafziger, W.E., Stewart, F., and Väyrynen, R. (eds) *War, Hunger and Displacement: The Origins of Humanitarian Emergencies, Vol. 1.* Oxford: Oxford University Press.

Stewart, F. 2002. 'Horizontal inequalities: A neglected dimension of development'. *QEH Working Paper* No. 81. Oxford: Queen Elizabeth House, University of Oxford.

Stewart, F. 2008 (ed.) *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies.* Basingstoke: Palgrave Macmillan.

Subramanian, S. 2007. 'Indicators of Inequality and Poverty' in McGillivray M. (ed.) *Human Well-being: Concept and Measurement.* Basingstoke: Palgrave Macmillan.

Ukiwo, U. 2006. 'Horizontal Inequalities and Violent Ethnic Conflicts: A Comparative Analysis of Ethnic Relations in Calabar and Warri, Southern Nigeria'. *DPhil Thesis.* University of Oxford.

UNDP. 1995. *Human Development Report 1995.* New York: Oxford University Press.

UNDP. 2003. *Avoiding the Dependency Trap.* Bratislava: UNDP

United States Information Service & United States Agency for International Development. 1992. *Educating the Girl Child: a seminar,* Islamabad: United States Information Service.

Van den Berg, S., and Louw, M. 2004 'Changing patterns of South African income distribution: Towards time series estimates of distribution and poverty'. *South African Journal of Economics* 72(3) 546-72.

Vranken, J., De Keulenaer, F., Estivill, J., Aiguabella, J., Breuer, W., Engels, D. & Sellin, C. 2001. 'Towards an Integrated Approach of European Policies on Social Exclusion and Inclusion'. Antwerpen: UFSIA - OASeS.

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

Zhang, X., and Kanbur R. 2003. 'Spatial Inequality and Health Care in China'. Washington: IFPRI.

Appendix Table 1. Availability of information on ethnicity/ language/religion in censuses and DHS

Key to read the chart:

✓ Information available

- Information not available

na Questionnaire was not found

The main sources of information are Population and Household Censuses. DHS surveys were consulted whenever the Censuses did not provide enough information.

East Asia, Central Asia and Pacific

Cambodia	Census 1962	Census 1998	Census 2008	DHS 2000	DHS 2005
Ethnic/racial origin	na	-	✓	-	-
Religion	na	✓	✓	✓	✓
Language	na	✓	✓	-	-

Kazakhstan	Census 1979	Census 1989	Census 1999	DHS 1995	DHS 1999
Ethnic/racial origin	na	na	-	-	-
Religion	na	na	-	✓	✓
Language	na	na	✓	✓	✓

Indonesia	Census 1961	Census 1971	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	-	-	-	-	✓
Religion	✓	✓	✓	✓	✓
Language	-	-	✓	✓	-

Indonesia	DHS 1991	DHS 1994	DHS 1997	DHS 2002-03	DHS 2007
Ethnic/racial origin	-	-	-	-	-
Religion	✓	✓	✓	✓	✓
Language	✓	✓	✓	✓	-

Azerbaijan	Census 1970	Census 1979	Census 1989	Census 1999	DHS 2006
Ethnic/racial origin	na	na	na	-	✓
Religion	na	na	na	-	✓
Language	na	na	na	✓	-

Lao PDR	Census 1973	Census 1985	Census 1995	Census 2005
Ethnic/racial origin	✓	✓	✓	✓
Religion	-	-	✓	✓
Language	-	-	-	-

Malaysia	Census 1970	Census 1980	Census 1991	Census 2000
Ethnic/racial origin	-	✓	✓	✓
Religion	✓	✓	✓	✓
Language	✓	✓	-	-

Papua New Guinea	Census 1966	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	na	-	na	-
Religion	na	-	na	✓
Language	na	-	na	-

Myanmar	Census 1973	1983
Ethnic/racial origin	na	✓
Religion	na	✓
Language	na	-

Philippines	Census 1960 & 1970	Census 1990	Census 1995	Census 2000	Census 2007
Ethnic/racial origin	na	-	✓	-	-
Religion	na	✓	-	✓	✓
Language	na	✓	✓	✓	-

Philippines	DHS 1993	DHS 1998	DHS 2003	DHS 2008
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	✓	✓	✓
Language	✓	✓	✓	✓

Singapore	Census 1970	Census 1980	Census 1990	Census 2000	Census 2010
Ethnic/racial origin	✓	✓	✓	✓	✓
Religion	-	✓	✓	✓	✓
Language	-	-	✓	✓	✓

Thailand	Census 1960	Census 1970	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	-	-	-	-	-
Religion	-	✓	✓	-	✓
Language	-	-	-	✓	✓

Timor-Leste	Census 2004
Ethnic/racial origin	-
Religion	✓
Language	✓

Vietnam	Census 1960	Census 1970's	Census 1989	Census 1999
Ethnic/racial origin	-	na	✓	✓
Religion	✓	na	-	✓
Language	✓	na	-	-

Sub-Saharan Africa

Burundi	Census 1970	Census 1979	Census 1990	DHS 1987
Ethnic/racial origin	na	✓	na	-
Religion	na	✓	na	-
Language	na	-	na	-

Benin	Census 1978	Census 1992	DHS 1996	DHS 2001	DHS 2006
Ethnic/racial origin	✓	✓	✓	✓	✓
Religion	-	✓	✓	✓	✓
Language	-	-	✓	✓	✓

Burkina Faso	Census 1962	Census 1976	Census 1985	Census 1996	Census 2006
Ethnic/racial origin	na	✓	-	na	-
Religion	na	✓	-	na	✓
Language	na		-	na	-

Burkina Faso	DHS 1993	DHS 1998-99	DHS 2003
Ethnic/racial origin	✓	✓	✓
Religion	✓	✓	✓
Language	✓	✓	✓

Botswana	Census 1971	Census 1981	Census 1991	Census 2001
Ethnic/racial origin	na	na	-	-
Religion	na	na	-	✓
Language	na	na	-	✓

Cameroon	Census 1976	Census 1987	DHS 1991	DHS 1998	DHS 2004
Ethnic/racial origin	-	-	-	✓	✓
Religion	-	-	✓	✓	✓
Language	-	-	✓	✓	✓

Central African Republic	Census 1961	Census 1975	Census 1988	Census 1998	DHS 1994-95
Ethnic/racial origin	na	-	-	na	✓
Religion	na	-	-	na	✓
Language	na	-	✓	na	✓

Cape Verde	Census 1970	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	✓	-	-	-
Religion	✓	-	-	-
Language	-	-	-	-

Chad	Census 1962	Census 1973	Census 1989	DHS 1996-97	DHS 2004
Ethnic/racial origin	✓	na	✓	✓	✓
Religion	-	na	✓	✓	✓
Language	-	na	-	✓	✓

Comoros	Census 1966	Census 1978	Census 1991	DHS 1996
Ethnic/racial origin	na	-	na	-
Religion	na	-	na	-
Language	na	✓	na	-

Congo, Rep.	Census 1974	Census 1996	DHS 2005
Ethnic/racial origin	na	na	✓
Religion	na	na	✓
Language	na	na	✓

Congo, Dem. Rep.	Census 1984	DHS 2007
Ethnic/racial origin	✓	✓
Religion	-	✓
Language	-	✓

Cote d'Ivore	Census 1975	Census 1988	Census 1998	DHS 1994	DHS 2005
Ethnic/racial origin	✓	na	✓	✓	✓
Religion	✓	na	✓	✓	✓
Language	-	na	-	✓	✓

Ethiopia	Census 1968	Census 1984	Census 1994	DHS 2000	DHS 2005
Ethnic/racial origin	na	✓	✓	✓	✓
Religion	na	✓	✓	✓	✓
Language	na	✓	✓	✓	✓

Eritrea	Census 1984	DHS 1995	DHS 2002
Ethnic/racial origin	na	✓	✓
Religion	na	✓	✓
Language	na	✓	✓

Gabon	Census 1960	Census 1969	Census 1980	Census 1993	DHS 2000
Ethnic/racial origin	na	✓	✓	na	✓
Religion	na	-	-	na	✓
Language	na	-	-	na	-

Gambia	Census 1963	Census 1973	Census 1983	Census 1993
Ethnic/racial origin	na	✓	✓	✓
Religion	na		✓	✓
Language	na		-	-

Ghana	Census 1960	Census 1970	Census 1984	Census 2000
Ethnic/racial origin	na	✓	-	✓
Religion	na	✓	-	✓
Language	na	✓	-	-

Ghana	DHS 1988	DHS 1998	DHS 2008
Ethnic/racial origin	✓	✓	✓
Religion	✓	✓	✓
Language	-	✓	✓

Guinea	Census 1972	Census 1983	Census 1996	DHS 1999	DHS 2005
Ethnic/racial origin	na	-	-	✓	✓
Religion	na	✓	✓	✓	✓
Language	na	-	-	✓	✓

Kenya	Census 1969	Census 1979	Census 1989	Census 1999
Ethnic/racial origin	✓	✓	✓	✓
Religion	-	-	-	✓
Language	-	-	-	-

Kenya	DHS 1993	DHS 1998	DHS 2003
Ethnic/racial origin	✓	✓	✓
Religion	✓	✓	✓
Language	✓	✓	✓

Lesotho	Census 1966	Census 1976	Census 1986	Census 1996	DHS 2004
Ethnic/racial origin	✓	✓	-	-	-
Religion	✓	✓	-	-	✓
Language	-	-	-	-	✓

Liberia	Census 1962	Census 1974	Census 1984	DHS 1986	DHS 2007
Ethnic/racial origin	na	✓	✓	✓	-
Religion	na	-	✓	✓	✓
Language	na	-	-	-	✓

Malawi	Census 1977	Census 1987	Census 1998	DHS 2000	DHS 2004
Ethnic/racial origin	✓	✓	-	✓	✓
Religion	-	-	✓	✓	✓
Language	-	-	✓	✓	✓

Mauritius	Census 1972	Census 1983	Census 1990	Census 2000
Ethnic/racial origin	na	na	-	-
Religion	na	na	✓	✓
Language	na	na	✓	✓

Mauritania	Census 1977	Census 1988	Census 1999
Ethnic/racial origin	na	✓	✓
Religion	na	-	-
Language	na	-	-

Mozambique	Census 1997	DHS 1997	DHS 2003
Ethnic/racial origin	✓	-	-
Religion	✓	✓	✓
Language	✓	✓	✓

Nigeria	Census 1973	Census 1991	DHS 1990	DHS 1999	DHS 2008
Ethnic/racial origin	✓	-	-	✓	✓
Religion	✓	-	✓	✓	✓
Language	✓	-	✓	-	✓

Niger	Census 1977	Census 1988	Census 1999	DHS 1992	DHS 2008
Ethnic/racial origin	✓	✓	na	✓	✓
Religion	-	✓	na	✓	✓
Language	✓	✓	na	✓	✓

Rwanda	Census 1978	Census 1991	Census 2002	DHS 1992	DHS 2005
Ethnic/racial origin	✓	✓	✓	✓	-
Religion	✓	✓	✓	✓	✓
Language	-	-	✓	-	✓

Senegal	Census 1955	Census 1976	Census 1988	Census 2002
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	-	✓	✓
Language	-	-	✓	✓

Senegal	DHS 1992	DHS 1999	DHS 2005
Ethnic/racial origin	✓	✓	✓
Religion	-	-	✓
Language	✓	✓	✓

Seychelles	Census 1987	Census 1997	Census 2002
Ethnic/racial origin	-	-	na
Religion	✓	-	na
Language	✓	-	na

Sierra Leone	Census 1963	Census 1974	Census 2004
Ethnic/racial origin	na	✓	✓
Religion	na	-	✓
Language	na	-	✓

South Africa	Census 1970	Census 1980	Census 1996	Census 2001
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	✓	✓	✓
Language	✓	✓	✓	✓

Sudan	Census 1964	Census 1973	Census 1983	Census 1993	DHS 1989-90
Ethnic/racial origin	na	na	-	✓	-
Religion	na	na	-	✓	✓
Language	na	na	-	✓	-

Swaziland	Census 1966	Census 1976	Census 1986	Census 1997	DHS 2006
Ethnic/racial origin	✓	na	-	✓	-
Religion	-	na	-	-	✓
Language	-	na	-	-	✓

Tanzania	Census 1978	Census 1988	Census 2002	DHS 2003	DHS 2004
Ethnic/racial origin	na	-	-	-	-
Religion	na	-	-	✓	✓
Language	na	-	-	-	✓

Togo	Census 1970	Census 1981	Census 1993	DHS 1988	DHS 1998
Ethnic/racial origin	na	✓	✓	✓	✓
Religion	na	✓	✓	✓	✓
Language	na	-	-	-	✓

Uganda	Census 1969	Census 1980	Census 1991	Census 2002
Ethnic/racial origin	na	-	✓	✓
Religion	na	-	✓	✓
Language	na	-	-	-

Uganda	DHS 1988-89	DHS 1995	DHS 2000-01	DHS 2006
Ethnic/racial origin	✓	✓	-	-
Religion	✓	✓	✓	✓
Language	-	✓	✓	✓

Zambia	Census 1969	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	na	✓	✓	✓
Religion	na	-	-	✓
Language	na	✓	✓	✓

Zambia	DHS 1992	DHS 1996	DHS 2001-02	DHS 2007
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	✓	✓	✓
Language	✓	✓	✓	✓

Zimbabwe	Census 1969	Census 1982	Census 1992	Census 2002
Ethnic/racial origin	✓	✓	✓	✓
Religion	-	✓	-	-
Language	-	✓	-	-

Zimbabwe	DHS 1988	DHS 1994	DHS 1999	DHS 2005-06
Ethnic/racial origin	✓	✓	-	-
Religion	✓	✓	✓	✓
Language	✓	✓	✓	✓

North Africa

Algeria	Census 1960	Census 1966	Census 1977	Census 1987	Census 1998
Ethnic/racial origin	-	-	-	na	na
Religion	-	-	-	na	na
Language	✓	✓	-	na	na

Egypt	Census 1966	Census 1976	Census 1986	Census 1996	Census 2006
Ethnic/racial origin	na	na	na	-	-
Religion	na	na	na	✓	✓
Language	na	na	na	-	-

Egypt	DHS 1988	DHS 1995	DHS 2000	DHS 2008
Ethnic/racial origin	-	-	-	-
Religion	✓	✓	-	✓
Language	-	-	-	-

Morocco	Census 1960	Census 1971	Census 1982	Census 1994
Ethnic/racial origin	na	-	-	na
Religion	na	-	-	na
Language	na	-	-	na

Morocco	DHS 1987	DHS 1992	DHS 1995	DHS 2003-04
Ethnic/racial origin	-	-	-	-
Religion	-	-	-	-
Language	-	-	-	-

Tunisia	Census 1966	Census 1975	Census 1984	Census 1994
Ethnic/racial origin	na	-	-	na
Religion	na	✓	-	na
Language	na	✓	-	na

Latin America and the Caribbean

Argentina	Census 1960	Census 1970	Census 1980	Census 1991	Census 2001
Ethnic/racial origin	-	-	-	-	✓
Religion	✓	-	-	-	-
Language	-	-	-	-	-

Bolivia	Census 1976	Census 1992	Census 2001	DHS 1994	DHS 1998
Ethnic/racial origin	-	-	✓	-	-
Religion	-	-	-	-	-
Language	✓	✓	✓	✓	✓

Brazil	Census 1960	Census 1970	Census 1980	Census 1991	Census 2001
Ethnic/racial origin	✓	-	✓	✓	✓
Religion	✓	✓	✓	✓	✓
Language	-	-	-	-	-

Chile	Census 1960	Census 1970	Census 1982	Census 1992	Census 2002
Ethnic/racial origin	-	-	-	✓	✓
Religion	✓	✓	-	✓	✓
Language	-	-	-	-	-

Colombia	Census 1964	Census 1973	Census 1985	Census 1993	Census 2005
Ethnic/racial origin	-	-	-	✓	✓
Religion	-	-	-	-	-
Language	-	-	-	-	✓

Costa Rica	Census 1963	Census 1973	Census 1984	Census 2000
Ethnic/racial origin	-	-	-	✓
Religion	-	-	-	-
Language	-	-	-	-

Dominican Republic	Census 1960	Census 1970	Census 1981	Census 1993	Census 2002
Ethnic/racial origin	✓	-	-	-	-
Religion	✓	-	-	-	-
Language	-	-	-	-	-

Ecuador	Census 1962	Census 1974	Census 1982	Census 1992	Census 2001
Ethnic/racial origin	-	-	-	-	✓
Religion	-	-	-	-	-
Language	-	-	-	-	✓

El Salvador	Census 1961	Census 1971	Census 1979	Census 1992	Census 2001
Ethnic/racial origin	-	-	-	-	na
Religion	-	-	-	-	na
Language	-	-	-	-	na

Guatemala	Census 1964	Census 1973	Census 1981	Census 1994	Census 2002
Ethnic/racial origin	✓	✓	✓	✓	✓
Religion	✓	-	-	-	-
Language	✓	-	✓	✓	✓

Guatemala	DHS 1987	DHS 1995	DHS 1998-99
Ethnic/racial origin	✓	✓	✓
Religion	-	-	✓
Language	✓	✓	✓

Haiti	Census 1961	Census 1971	Census 1982	Census 1992	Census 2001
Ethnic/racial origin	na	-	-	na	-
Religion	na	✓	✓	na	✓
Language	na	-	✓	na	-

Honduras	Census 1961	Census 1974	Census 1988	Census 1992	Census 2001
Ethnic/racial origin	-	-	na	na	✓
Religion	-	-	na	na	-
Language	-	-	na	na	-

Mexico	Census 1960	Census 1970	Census 1980	Census 1990	Census 2000	Census 2005
Ethnic/racial origin	-	-	-	-	✓	-
Religion	✓	✓	✓	✓	✓	-
Language	✓	✓	✓	✓	✓	✓

Nicaragua	Census 1963	Census 1971	Census 1995	Census 2005
Ethnic/racial origin	-	-	-	na
Religion	✓	-	✓	na
Language	✓	-	✓	na

Panama	Census 1960	Census 1970	Census 1980	Census 1990	Census 2000
Ethnic/racial origin	-	-	-	✓	✓
Religion	-	-	-	-	-
Language	-	-	-	-	-

Paraguay	Census 1962	Census 1972	Census 1982	Census 1992	Census 2002
Ethnic/racial origin	-	-	-	-	-
Religion	✓	-	-	✓	✓
Language	✓	-	✓	✓	✓

Peru	Census 1961	Census 1972	Census 1981	Census 1993	Census 2005
Ethnic/racial origin	-	-	-	-	-
Religion	✓	✓	✓	✓	✓
Language	✓	✓	✓	✓	✓

Uruguay	Census 1962	Census 1975	Census 1996
Ethnic/racial origin	-	-	-
Religion	-	-	-
Language	-	-	-

Venezuela	Census 1961	Census 1971	Census 1981	Census 1990	Census 2001
Ethnic/racial origin	-	-	-	-	✓
Religion	-	-	-	-	-
Language	-	-	-	-	✓

South Asia

Bangladesh	Census 1981	Census 1991	Census 2001	DHS 2004	DHS 2007
Ethnic/racial origin	✓	-	-	-	-
Religion	✓	✓	✓	✓	✓
Language	-	-	-	-	-

India	Census 1971	Census 1981	Census 1991	Census 2001
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	✓	✓	✓
Language	✓	✓	✓	✓

Nepal	Census 1971	Census 1981	Census 2001	DHS 1996	DHS 2006
Ethnic/racial origin	na	-	✓	✓	✓
Religion	na	✓	✓	✓	✓
Language	na	✓	✓	✓	✓

Pakistan	Census 1971	Census 1981	Census 1998	DHS 1990-91	DHS 2006-07
Ethnic/racial origin	-	-	-	-	-
Religion	✓	✓	✓	-	-
Language	✓	-	✓	✓	✓

Sri Lanka	Census 1971	Census 1981	Census 2001	DHS 1987
Ethnic/racial origin	✓	✓	✓	✓
Religion	✓	✓	✓	✓
Language	-	-	-	-

Online resources:

https://international.ipums.org

http://www.surveynetwork.org

http://www.acap.upenn.edu

http://www.measuredhs.com

'. . . simple inequality between rich and poor is not enough to cause violent conflict. What is highly explosive is . . . horizontal inequality: when power and resources are unequally distributed between groups that are also differentiated in other ways – for instance by race, religion or language.' Kofi Annan

CRISE is a development research centre within the University of Oxford supported by the United Kingdom Department for International Development (DFID).

Set up in 2003 under the direction of Professor Frances Stewart, CRISE aims to investigate the relationships between ethnicity, inequality and conflict in order to identify policies that will promote peace.

UNIVERSITY OF OXFORD

Centre for Research on Inequality, Human Security and Ethnicity (CRISE)
Oxford Department of International Development (Queen Elizabeth House)
University of Oxford, 3 Mansfield Road, Oxford OX1 3TB, UK

T +44 1865 281810 F +44 1865 281801 W www.crise.ox.ac.uk