

Working Paper No. 2012-4

Methodology of the Indices of Social Development

*Roberto Foa*¹

*Jeffery C. Tanner*²

August 2012

¹ Harvard University, Department of Government. 1737 Cambridge St, Cambridge, MA 02138. foa@fas.harvard.edu.

² Pardee RAND Graduate School. 1776 Main Street, Santa Monica, CA 90407. jtanner@rand.org.

We are grateful to the World Bank for funding and hosting this research. The views expressed in this paper are solely those of the authors and should not be ascribed to their associated institutions.

ISSN 2213-6614

The Institute of Social Studies is Europe's longest-established centre of higher education and research in development studies. On 1 July 2009, it became a University Institute of the Erasmus University Rotterdam (EUR). Post-graduate teaching programmes range from six-week diploma courses to the PhD programme.

Research at ISS is fundamental in the sense of laying a scientific basis for the formulation of appropriate development policies. The academic work of ISS is disseminated in the form of books, journal articles, teaching texts, monographs and working papers. The Working Paper series provides a forum for work in progress which seeks to elicit comments and generate discussion. The series includes academic research by staff, PhD participants and visiting fellows, and award-winning research papers by graduate students.

Working Papers are available in electronic format at www.iss.nl

Please address comments and/or queries for information to:

Institute of Social Studies

Irene van Staveren

P.O. Box 29776

2502 LT The Hague

The Netherlands

or

E-mail: isd@iss.nl

Table of Contents

Abstract	4
I. Introduction	5
II. Methods of Generating Composite Indices	7
2.1 Selection.....	7
2.2 Weights	8
III. Constructs in the Indices of Social Development	12
IV. The Matching Percentiles Methodology in the Indices of Social Development	16
V. Why Matching Percentiles?	19
VI. Results.....	22
VII. Diagnostics	25
7.1 Discussion of the Clustering of the Indicators	26
7.2 Perceptions-based Versus Actionable Indicators.....	28
7.3 Selection of Indicators -- What's in?	33
7.3.1 Outlier Observations	33
7.3.2 Results.....	42
7.4 What's Where?.....	43
7.4.1 Hierarchical Cluster Analysis	43
7.4.2 Factor Analysis	46
7.4.3 Testing the Robustness of the Aggregation Methodology – Sensitivity Analysis	52
VIII. Conclusion	59
References.....	62
Appendix.....	67

Abstract

In recent years, international organizations, think-tanks, and the social sciences have contributed to a dramatic expansion in the range of composite indices measuring concepts such as human development, governance, or social capital. This paper discusses challenges faced in the design of composite indices, and suggests the method known as matching percentiles as a solution to the problem of how to aggregate social institutional data. We produce a set of indices measuring the health of five dimensions of social development: civic activism, clubs and associations, inter-group cohesion, interpersonal safety and trust, and gender equity. We then conduct a series of diagnostic tests which demonstrate the robustness of the new measures, and assess their degree of construct validity.

I. Introduction

In recent years, international organizations, think-tanks, and academics in the quantitative social sciences have overseen a proliferation of composite indices designed to assess broad social science concepts such as governance or human development. International organizations such as the United Nations and the World Bank have produced aggregate development indicators including the Human Development Index (HDI), the Gender Empowerment Measure (GEM), the Doing Business (DB) indicators, and the Worldwide Governance Indicators (WGI), while think-tanks and consultancies such as Freedom House, the Economist Intelligence Unit, and Transparency International have produced indices such as the Political Rights and Civil Liberties indices, the Quality of Life index, and the Corruptions Perceptions Index (United Nations 2007, Doing Business 2005, Kaufman et al. 1999, Lambsdorff 2006). A review of the phenomenal growth of composite indices conducted by Bandura and Martin del Campo (2006), found that of the 160 composite cross-country indices now in existence, 83% had been generated since 1991 and 50% in the previous 5 years alone, while, before 1991 there were less than 20% of the composite indices found available today.

Why have composite indices found such favor among development organizations and researchers in recent years? First, a composite measure has the ability to summarize complex or multi-dimensional issues in a simple manner, making it possible for policymakers to get a tractable and representative sense of the situation in a given country and in comparison with others. Second, because they provide a single estimate, composite indices have substantial ease of interpretation over the use of multiple benchmarks, while quantification of a concept makes it possible to assess progress over time and to highlight cases where intervention may be needed. Third, the commitment to regularly produce and update quantitative ratings facilitates communication with ordinary citizens, including stakeholders in developing countries, showing both the commitment of an organization to a particular set of development challenges. Finally, composite indices are an important starting point for debate. Until the publication of the Worldwide Governance Indicators (1999) or the Ease of Doing Business Index (2001), ‘good governance’ was largely a catchphrase. By defining and

measuring it, a process of dialogue has begun over what is and ought to be understood by quality of governance.

However, as critics have frequently alleged, when poorly designed, composite indices also carry attendant risks. Nardo et al. (2005), for example, note how ill-constituted composite measures can send misleading policy messages or invite simplistic policy conclusions. Saltelli and Tarantola (2007), for example, give the case of a sustainability index, cited in a major newspaper, that rewards oil and gas exporting countries higher due to the large budget surplus engendered by a temporary boom in commodities prices. As the authors note, in this case readers are better informed by examining the separate indicators individually, rather than referring to the aggregate score. For these reasons, key decisions have to be made when designing composite indices regarding the selection of indicators, weighting schemes, and how to deal with missing data so as to ensure that final index scores reflect a meaningful distillation of the available information.

Specifically, the purpose of this paper is to discuss the methodological decisions and outcomes taken in the construction of the indices of social development, a set of indices that combine over 200 indicators from 25 sources to provide a comprehensive picture of social institutions across the world. In this paper we describe some of the methodological challenges, provide a brief outline of the matching percentiles method used in the aggregation process, and conduct a series of brief diagnostic tests in order to demonstrate the reliability and validity of the new measures. Section II summarizes the methodological decisions required when constructing aggregate measures. Section III provides an outline of the constructs used in the design of the indices of social development. Section IV advances our own solution to the problem of aggregating social data, based on the matching percentiles method, and section V discusses some of its advantages over other methods. Section VI briefly discusses the results of the matching percentiles method. Section VII addresses potential concerns with this aggregation method, via a series of brief diagnostic tests, and finally section VI concludes.

II. Methods of Generating Composite Indices

2.1 Selection

The first major area that the designer of a composite index must decide is with regard to the selection of variables for use in constructing the measure. Specifically, the designer must decide whether to concentrate on just one or two 'key variables', or to adopt a more comprehensive approach using data from a wide range of indicators of varying data quality. An example of an index using a smaller number of select variables is the Human Development Index, which in its most recent iteration uses just six items: life expectancy, the literacy rate, income per capita, and the primary, secondary and tertiary enrolment rates (Human Development Report 2007). At the other extreme, the Worldwide Governance Indicators project constitutes among the most comprehensive exercises in data aggregation, with over 300 indicators from 33 separate data sources (Kaufmann et al. 2007).

The judgment as to whether the index designer should adopt a broad or a narrow selection of indicators depends largely upon the latent variable that the measure is intended to capture. An index of, say, cardiovascular health may reasonably rely on just one or perhaps several indicators, such as active and resting heart rate, blood pressure, and history of myocardial infarction. Here a single indicator may be at once reliable, valid, and representative of a large number of cases. In the case of measuring dimension of governance, for example a measure of the rule of law, no such indicator exists. Researchers should therefore examine a broader array of measures. Some items, such as a rating of contract enforceability by a consultancy organization, may be valid and cover a large number of countries, yet be considered unreliable due to perceptions bias and correlated error with other ratings. Meanwhile another indicator, such as a public opinion survey regarding the incidence of consumer fraud, may be reliable and fairly valid, but cover only a small portion of countries. Using a larger pool of indicators, therefore, is likely to be the only means of accurately measuring that concept.

Another dilemma in index design is the choice between reliability and representation. If the indicator selection is narrow, it is simpler to understand how the index is derived, and the index may gain credibility from the reliability of its sources, assuming these are well selected.

However, the risk of a narrow selection is that the indicators chosen may not be relevant to what the index purports to measure: for example, due to its reliance on demographic data, the Human Development Index is sometimes misunderstood as a measure of human capital (that is, physical and mental productive capacity) rather than a measure of human capabilities in general (Fukuda-Parr and Kumar 2004). At the other extreme, a more encompassing selection of indicators improves the ability to validly measure every aspect of a phenomenon, but it can then prove more difficult for readers to understand what the index scores represent. There is no clear set of criteria to determine which approach is more advisable, but in general the ‘fuzzier’ the concept and the weaker the available data, the more likely only a large pool of indicators can accurately capture the construct in question.

2.2 Weights

The second consideration is the assignment of weights to indicators in order to produce the final index. Four basic types of solution to this problem can be found among the existing range of composite indices: the use of equal weights among items (Basic Capabilities Index, E-Government Index, Failed States Index); theoretically categorized weights (HDI, Gender Empowerment Measure, Doing Business Index, Environmental Sustainability Index, Economic Freedom Index); schematic weights (EIU Quality of Life Index); and variable weights (Worldwide Governance Indicators, Corruptions Perceptions Index). Equal weighting simply means that each item of data used by an index is averaged in order to produce a final score. For example, the Basic Capabilities Index (formerly the Quality of Life Index) takes a simple average of the rescaled values of the primary completion rate, the child mortality rate, and the percentage of births attended by skilled personnel (Social Watch 2007). Strict use of equal weighting is comparatively rare, and far more common is the categorization of indicators into theoretically derived subcomponents. For example, the Human Development Index assigns its 6 indicators into 3 component areas: a long and healthy life (measured by life expectancy); a good income (measured by income per capita); and skills and knowledge (combining adult literacy, primary school enrolment, secondary school enrolment, and university enrolment). While each of the components has equal weight in producing the final index score, each indicator within them does not: life

expectancy and income per capita each account for 1/3 of the variation in final scores, and each of the 4 education variables account for 1/12.

Clustering by thematic area is very common in composite index projects that use a large number of items. Other indices which make use of subcomponents aggregated before the final indexing process include the Doing Business Indicators, which cluster items into 10 different areas relevant to starting, managing, and closing an enterprise, and the Environmental Sustainability Index, which categorizes items from 76 datasets into 21 areas. However, while use of equal weights among theoretically specified subcomponents has the advantage of clarity, it courts potential difficulties. For example, if a measure of 'human capital' clusters subcomponents into three areas - 'knowledge', 'health' and 'information' – then it is quite likely that the 'knowledge' and 'information' clusters will overlap substantially, perhaps reflected in a high statistical correlation between the two measures.

A range of statistical procedures have been developed in order to ascertain an appropriate weighting scheme. The first of these is principle components analysis (PCA), which assigns factor loadings based upon whether a subsequent indicator shares a common factor with another variable in the dataset. For example, PCA is likely to weigh down 'knowledge' and 'information' in the above example, as both depend upon a single underlying factor (the skills and education of the population). In practice, however, very few composite indices use PCA weights, in part because it is difficult to explain the process to non-statisticians, in part because the weights themselves change as the data changes over time, but mainly because the results using equal weights and PCA weights tend not to differ substantially (the Doing Business Indicators 2005 report conducts such a comparison, and shows minimal discrepancies). A second method of deriving weights for use in index aggregation is through regression processes. This can be done when a highly valid and reliable measure of the latent variable exists, but only for a restricted subset of countries. In that case the reliable measure can be used as the dependent variable in a regression framework, with the index indicators or subcomponents used as independent variables, and the resulting coefficients used as weights. In essence, such a design tells us what the scores 'should have' been for the reliable indicator, were a larger country sample available. An example of this is the Quality of Life Index produced by the Economist Intelligence Unit in 2005 (EIU 2005). Veenhoven (2005)

posits that responses to public opinion surveys asking respondents how happy or satisfied they are at the present time is a reliable, valid measure of human wellbeing. However, this data exists only for those countries in which public opinion surveys have been fielded. Therefore, the designers of the Quality of Life Index designed a regression model using nine indicators as independent variables, such as income per capita, political stability, or gender equality, and use these nine variables to project quality of life estimates for a total of 111 countries.

A clear limitation of the regression approach, of course, is that very rarely does a direct measure of a latent variable exist in the same way that survey data on subjective wellbeing provides a direct measure of people's quality of life. Another limitation with the regression approach to weighting, the PCA approach, and the use of theoretically derived weights, is that in themselves they offer no solution in cases where data may be missing for particular countries.

Composite indices can deal with the problem of missing data in one of three ways. The first and simplest solution is casewise deletion: to drop any country for which complete data does not exist. This naturally avoids a great number of methodological tangles, and is the approach used in such indices as the Doing Business Indicators. The Doing Business Indicators utilize a team of over 30 researchers working constantly to extract the required information from their rated governments. Analogous to casewise deleting is indicator deletion: dropping variables which are incomplete for the full set of countries or, as with the Human Development Index, selecting variables for which complete data across the domain of countries is relatively easy to obtain. The Human Development Index, as we have discussed, restricts itself to a small set of fairly narrow topic indicators (such as literacy or income per capita) for which complete data exists. In the absence of either a very narrow indicator set or the resources to collect primary data, casewise or indicator deletion is simply not feasible. The consequence of casewise deletion for an indicator of civic engagement, for example, would be to reduce the sample to a very small number of countries, or even none at all.

The second solution to the problem of missing data is to impute missing values. Use of imputation is rare in indices produced by international organizations, but relatively more common in academic indices and datasets. The Environmental Sustainability Index (ESI), for example, uses Markov Chain Monte Carlo simulation to impute values for the missing variables in the dataset. Imputation solutions, however, have two potential drawbacks. The first is that imputation is unreliable in cases where appropriate estimation models cannot be determined from available variables. Furthermore, it can lead to highly erroneous results when data is missing for a very large number of countries on a given variable. As Dempster and Rubin (1983) remark, imputation “is seductive because it can lull the user into the pleasurable state of believing that the data are complete after all, and it is dangerous because it lumps together situations where the problem is sufficiently minor that it can legitimately be handled in this way and situations where standard estimators applied to real and imputed data have substantial bias.” The second problem is that there is a serious problem of legitimacy where nations are rated on a given dimension of country performance based on data that is merely estimated, rather than actual. In such cases, it is very difficult to guard against challenges by critics from countries which are rated poorly in such an exercise that the scores are inaccurate; and it is precisely for this reason that use of imputation is more common among academicians than among international organizations such as the United Nations, the European Commission, or the World Bank.

A third alternative approach to dealing with missing data is to use the existing data entirely (no case-wise deletion) and exclusively (no imputation) in the estimation of the index, and to supplement this with an estimated margin of error, based *inter alia* on the number of missing items. This is the approach of a number of more recent indices such as the Corruptions Perceptions Index (CPI) produced by Transparency International and the World Bank’s Worldwide Governance Indicators (WGI). Such approaches carry a dual advantage, in that they allow scores to be estimated for a maximal number of countries, and can use a broader range of indicators to triangulate indices for nebulous constructs.

III. Constructs in the Indices of Social Development

In recent years, there has been a growing range of cross-country sources providing new data on social institutions and practices. The World Values Surveys have since 1981 steadily expanded their coverage to include 92 societies around the globe, and are now complemented by a broad array of regional survey projects: Latinobarometer, founded in 1996, which covers 18 Latin American societies; Afrobarometer, founded in 1999, which covers 18 sub-Saharan African societies; and Asia Barometer, founded in 2003, which covers 18 Asian societies. Meanwhile, the Eurobarometer surveys, which were founded in 1973, presently cover the 27 countries of the European Union plus Turkey and Russia³, and the International Social Survey Programme, started in 1984, covers a total of 45 countries. The result is a vast wealth of comparative survey data for comparing social practices and outcomes: Donsbach and Traugott (2007), for example, list 66 comparative survey projects alone in the postwar era, 44 of which were initiated in the past two decades. In addition to comparative survey projects, there is also a growing array of numerical ratings based on qualitative assessment, including the Civicus *Civil Society Index* (CSI), a rating of the state of civil society based upon expert reports, and the *Minorities at Risk* project, which provides comparative assessments of discrimination and exclusion of minority groups in 118 societies across the world, from which to base judgments of country social institutions.

As yet, however, there are few comprehensive cross-country indices that combine these data to form a judgment of the state of social institutions and practices, such as civil society development, intergroup cohesion, or gender discrimination. The Indices of Social Development fill this gap by aggregating the vast range of available social institutional data into a set of cross-country surveys and ratings. To create these estimates, the indices draw from over 200 indicators from 25 sources, including international organizations, comparative survey projects, rating agencies, and academic assessments⁴. Five basic ‘types’ of social institution are highlighted according to the means by which they achieve this end: **civic**

³ The Eurobarometer project encompasses several surveys, of which some are regionally specific within the continent.

⁴ A summary is provided in Appendix I.

activism, clubs and associations, intergroup cohesion, interpersonal safety and trust, and gender equity.

The first of the measures, *civic activism*, measures the extent to which social practices encourage more active and critical engagement with political authorities. It is rooted in the work of political scientists such as Almond and Verba (1963) or Putnam et al.'s (1993) work on civic engagement. At its core is the existence of what Norris (1999) calls 'critical citizens', that is, political subjects able and willing to articulate and represent their interests before government. Such practices include but are not limited to engagement in debate and reflection over public policy, through the media for example, and the practice of actively representing citizen interests through contacting public officials and protesting unpopular policies. A similar concept of civic engagement has been deployed in studies by Paxton (2002) and Inglehart and Welzel (2005). In terms of benefits for social welfare, greater civic engagement is able to foster collective action for the delivery of better governance.

The second of the measures, *clubs and associations*, measures the level of participation in voluntary activities conducted amongst individuals in the same locale, such as a village or neighborhood. The existence of such an associative life is central to what Woolcock and Narayan (2000) term the 'communitarian' understanding of social capital, and its core elements include local organizations such as clubs, associations, and civic groups. Social capital is widely associated with the works of Robert Putnam (2000) and Amitai Etzioni (1997), whose studies have examined the health of associative life in western democracies. In terms of their benefits for social welfare, community networks are essential worldwide for risk mitigation, as well as playing important roles in the oversight and disbursement of local level funds (Guggenheim 2006).

The third of the measures is *inter-group cohesion*. This index specifically measures the extent or absence of routinized conflict between ethnic, religious, or other social identity groups. Though a specific literature on identity-based social conflict is less well-defined than for, say, social trust, debates in social psychology have made similar use of the term 'social conflict', although this term may also include non-violent exchanges (Seeman et al, 1994; Schuster et al, 1990). The term 'collective violence' has also been used to designate a range of events in

which group identities are mobilized to produce conflict, including among non-state and quasi-state actors (Tilly 2003). Essential to the definition of this concept deployed by the indices of social development are acts of violence that are i) conducted by non-state actors and ii) perpetrated by and directed against individuals of specific identity groups. In terms of its benefits for social welfare, inter-group cohesion reduces transaction costs, as well as reducing the negative externalities associated with conflictive behavior, that is, through the spillover of violence to third-parties.

The fourth measure, *interpersonal safety and trust*, is supported by a longstanding theoretical and empirical literature on the concept of social trust (e.g. Fukuyama 1995, Arrow 1974, Knack and Keefer 1997). A central aspect of this definition is that it reflects a norm of reciprocity binding individuals who may not have met or have a deep knowledge of one another. Furthermore, generalized trust includes the willingness of individuals to cooperate with other members of society with whom they may be poorly acquainted, as well as the extent to which others exhibit trustworthiness by refraining from activities that violate norms of cooperation between individuals. The importance of both trust and trustworthiness for the concept has been highlighted by Glaeser et al. (2000), who show that survey questions on social trust are as reflective of the likelihood of trust violations by others as they are the willingness to make trusting choices. Examples of safety and trust violation may range from minor infringements such as cheating or fraud, to more serious violations such as theft, assault, or other interpersonal violence. Social trust primarily contributes to improving human welfare by reducing transaction costs, with secondary benefits in the form of enhanced collective action and is argued to play a key role in ensuring the security of property rights and contract (Knack and Keefer 1997).

Finally, the fifth measure, *gender equity*, specifically estimates the level of discrimination occurring against women. Gender discrimination is the subject of a vast literature in the fields of labor economics, sociology, and demography, and in terms of its welfare effects, gender equity implies improved allocative efficiency benefitting individual and aggregate economic outcomes (Schultz, 2002, Esteve-Volart, 2004, Morrison et al. 2007). Empirical evidence also shows that increases in female education yield larger impacts than similar improvements in male education on human development outcomes such as child survival,

health and schooling (World Bank 2001, Schultz 2002, Thomas et al. 1991, King and Hill, 1993).

The five categories outlined above do not constitute the only means of dividing the broad construct of social institutions, but one that is conceptually coherent and in accordance with previous work in this field (Stone 2001, Paldman, 2000, Narayan and Cassidy 2001). An overview of the behaviors which define each of these dimensions of social institutions is found in Table 1 below, while a more detailed comparison of how this taxonomy compares with those proposed for previous attempts to provide cross-country measures of social institutions or ‘social capital’, based on a review of the appropriate literature, is outlined in the Diagnostics section in the second half of this document.

Table 1 **Indices of Social Development, Benefits, Defining Practices**

<u>Construct</u>	<u>Benefits</u>	<u>Defining Practices</u>
Civic Activism	Collective action	Signing petitions, writing to representatives, electoral turnout, keeping informed of local politics, peaceful demonstration
Clubs and Associations	Collective action	Voluntary social work, participation in social networks, giving to local community organizations
Inter-group Cohesion	Reduced transaction cost	Ethnic and religious riots, attacks on members of another community
Interpersonal safety and trust	Transaction costs, collective action	Petty criminality (theft, burglary, fraud) to serious crime (assault, murder)
Gender equity	Allocative efficiency	Refusal of jobs and other benefits to women based on non-merit criteria

IV. The Matching Percentiles Methodology in the Indices of Social Development

The aggregation method chosen to combine our set of 200 indicators into these five indices is a variant of the matching percentiles methodology used by Lambsdorff (Lambsdorff 1999, 2006). In this approach, scores are assigned to countries based on ordinal rankings. The ranks of countries for variables included in the index are used to assign equivalent values to countries with equivalent ranks. In this section, we outline the exact stages used in progressing from a set of individual indicators to a final compositive index score.

For the Indices of Social Development, we take three steps in preparation for the indexing process itself. First, each component variable is standardized to have a mean zero and standard deviation one and oriented such that more desirable outcomes receive a higher (positive) value and less desirable outcomes receive a lower (negative) value. Next, we take an observation-level (row) average of the variables for the five years around each of our anchor years for the indices: 1990, 1995, 2000, and 2005. So, for example, the 1995 value will be a row (country) average of the 1993, 1994, 1995, 1996, and 1997 values. Finally, in an effort to overcome issues of correlated error - in instances where we have multiple variables from the same source covering the same type of item, for example - we take a row average of the error-correlated variables to create a series of “subindices” used as the input variables

in the matching percentiles process.⁵ This averaging also nets out a portion of any classical measurement error which may exist in these variables.

As with any indexing methodology, the basic premise is that for each of the five dimensions of social development there exists some latent construct (L_i) representing the true level of that construct in country i . Each of the k available indicators x_{ik} is a manifestation, on a different functional transformation $h_k()$ and with varying degrees of measurement error e_{ik} , of the level of L_i such that:

$$x_{ik} = h_k(L_i) + e_{ik} \quad (\text{eq. 1})$$

and the latent variable can be recovered through some function $g()$:

$$L_i = g_i(x_{ik}) + u_i \quad (\text{eq. 2})$$

In other indexing applications, the functional form of h is usually assumed to be known (and usually assumed linear and isomorphic across k). Here we recognize that we are unable to estimate the functional form of h_k and so employ a nonparametric aggregation methodology, with no assumptions regarding the functional form of the relationship between L and x_k . We merely assume that the relative position of countries on x for variable k reflects a better or worse underlying condition with respect to L . The ranks of successive indicators in the index are then used to assign values to countries.

Our method uses a recursive process of matching observational ranks over pairs of variables: a master and an input variable. The initial master variable is a random variable $\sim N(0.5, 0.15)$ and the input variables are the “sub-indices” described earlier. Taking each of the input variables in turn, the algorithm first determines which observations appear in both the master and input variables. Observations for this conjoint set are then ranked separately for

⁵ A Note on Data coverage:

For subindices of the Local Community or Safety & Trust indices which use variables not available for the current round such that the entire subindex is missing, we use data from the previous round for that subindex. If the previous round is not available we leave the subindex as missing. This does not occur much in our data, but where it does it is almost always for the 2005 round (as data from 2006 and 2007 are often not yet available) and only ever for the Local Community and Safety & Trust indices.

the master and the input variables. Having obtained master and input variable ranks for each observation, we next create a “match” variable which rescales the input variable by assigning the cardinal value of the country in the master variable to the country with the same ordinal rank in the input variable. For example, if Albania, Burundi, Cameroon, and Denmark were to have master variable scores of 0.45, 0.61, 0.65, and 0.89, and input variable scores of 0.82, 0.94, 0.31, and 0.46, then they would receive “match” scores of 0.65, 0.89, 0.45, and 0.61.

Each observation which has a value in the input variable will receive a matched value, matched against the master variable for each of the K input variables (sub-indices) used in creating the index. Once the match values are assigned for each of the input variables, the K match variables are averaged to create the index score for each country⁶.

As the indexing process is obviously influenced by the draw of the random normal master variable, the newly created index score is fed back through the indexing process as a new master variable. This process iterates recursively until the index reaches convergence. Our convergence parameter is 10^{-4} for the sum of the squared differences between the master variable and resulting index within a particular iteration. As a further check, this convergence process is run with 1,000 Monte Carlo runs for each index. A country’s final score is its average score across the 1,000 runs. The reported standard error for each country is the average of the standard error across 1000 runs where the standard error in each individual run is the standard error across the K matched variable scores in the final (converged) indexing iteration. Because the law of large numbers assures that the means of the index means over the 1,000 Monte Carlo runs is distributed asymptotically normal, assigning the initial master variable a normal distribution at the outset does not seem egregious.

⁶ The denominator of this mean for country i is the number of matched input variables for which country i has a non-missing value.

V. Why Matching Percentiles?

Using the matching percentiles methodology for the Indices of Social Development has several advantages over the alternative methods described earlier. Among the most important of these is the ability to handle variables which have many missing values.

The ability to use incompletely populated variables is tremendously important in developing a global index for a latent construct that is difficult to measure. Such an index benefits from being able to use as much good data as possible, given that data items for these constructs frequently do not often cover every country. Discarding a variable because it is incompletely populated would waste the useful information collected for countries which the variable does treat.

By contrast, most other methods—regression principle components analysis, factor analysis—require completeness of information. Methods based on a regression framework, for example, require that any observation used in estimation have data for the dependent variable and each of the regressors, and observations which are missing data in any one of these variables will be dropped. As the data used in indexing social development is likely not missing at random (much less missing completely at random), dropping observations with partially missing data is likely to bias estimates for the remaining countries—and obviously leads to fewer scored observations and consequently less robust indices.

As mentioned previously, one strategy for dealing with missing data is imputation, which estimates observations with partially missing data based on those observations that are available.

However, unless these observations are missing at random, these imputations are susceptible to omitted variables bias, and so can bias other index estimates in the sample. Imputation is also invoked in aggregation procedures that use a correlation matrix, including principal

components, factor analysis, and correspondence analysis, and require a completely populated observation-by-variable matrix for most indexing calculations.

Yet data imputation is unattractive in this context for both technical and political reasons. As discussed, data for a globally comparable component of a worldwide indicator is likely to have large swathes of the data of any particular variable missing if only because sources most often focus on a particular set of countries with a common geography, polity, or economy. Imputing missing data would yield a low ratio of imputing to imputed data. Furthermore, any imputation method using a maximum likelihood process assumes independence in observations. As countries may have an incentive to not report particular outcomes or to not allow some types of surveys to be done, they are likely not independent, nor would they be conditionally independent as we are not able to observe factors which may condition data availability. Consequently, it may be better to abandon maximum likelihood completely rather than violate its assumptions.

A final limitation of imputation is that regardless of why a country's data may be missing, a government may take umbrage with a score relying heavily on imputed data, as the country can claim that an imputation does not accurately reflect their true score. Furthermore, imputation methods may impute data to be far outside of the sample or have strange outcomes. The likelihood of imputations being incorrect increases as the share of missing data increases under an imputing mechanism.

The matching percentiles method described here does not assign scores based on imputed missing data for component variables and so avoids the problems of imputation. When the matching percentiles method as described here encounters a country with a missing value for a particular variable, that observation is skipped and no "matched score" is calculated for that country-variable. Because the index score of each Monte Carlo run is calculated by taking a country mean of the matched input variables for which the country does have a score, we can score countries which have missing data over some variables, even if the standard errors for that score cover a larger area of the density. Furthermore, in the Indices of Social Development, each observation's score is independent of other observations, and is hence largely immune to the effects of an omitted variables bias.

Because our matching ranks method does not rely on imputation for missing data, it is sufficiently robust to handle variables which cover few countries. Furthermore, because the process relies on convergence and scores only those countries in the conjoint set of master and input variables, the process is not biased when encountering variables which report only “good” or only “bad” countries as can happen in other standardization processes whereby the worst of the good countries can be unduly punished and the best of the bad countries can be unduly rewarded.

A further benefit of the matching percentiles method outlined here is that it dispenses with the linearity assumption, an assumption that is often unrealistic when dealing with indicators that may have important diminishing returns or threshold effects. Where other indexing methods often rely on assumptions of distribution of data or errors or on a specific functional form for the relationship between the latent construct and the observed variables (typically, a linear relationship), matching percentiles relies on ordinal rather than cardinal information and the asymptotic properties invoked by the law of large numbers. Recalling equations 1 and 2, our functional form assumptions are nonparametric in b_k in that we rely only on rankings rather than cardinal values over the input variables, and linear in g_i as we average across the input variables to estimate the latent variable. In our view this linear treatment is at least as robust and conceptually more straight-forward than any competing functional form.

The intuition behind the aggregation process of the matching ranks method is arguably easier for a lay person to grasp than procedures that put the data through a range of statistical and mathematical machinery to pull out the portion of the variation of the variables explained by the unobserved component. Rather, we are essentially giving countries the same score who have the same rank, and then averaging all of those matched scores together.

Finally, our variant of the matching percentiles method employed for the Indices of Social Development allows two simple advantages over other matching percentiles exercises: first, our method allows us to estimate all anchor years simultaneously in addition to indexing

social development within each anchor year. This not only allows us to examine global distributions, but also gives us the great advantage of being able to compare countries with themselves and each other across time. Second, where we do index social development using only data within an anchor year, our method also allows us to do these estimations independent of previous years' results so as to avoid the methodological entrenchment which may inhibit showing any time variation which may have occurred in these already slow-moving latent variables.

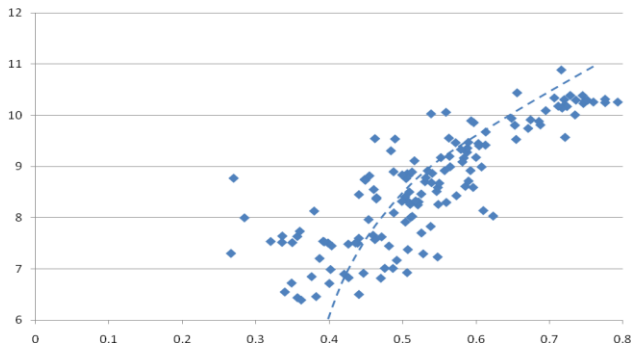
VI. Results

The matching percentiles method is applied to each of a subgroup of the 200 indicators of social development to combine these into our five Indices of Social Development: *civic activism, clubs and associations, intergroup cohesion, interpersonal safety and trust, and gender equity*. A detailed breakdown by index of indicators used is provided in the Appendix. How do these indices of development compare with more conventional development measures? Figure 1 shows the relationship between economic development, as measured by log per capita GDP at PPP, and the quality of social institutions, measured by each of the five social development clusters. Plotted on the charts are the estimated relationship between log per capita GDP and each of the indices, based on the coefficients for log GDP and log GDP squared that are reported in Table 2. It can be seen that for all of the clusters except for membership of clubs and associations, there is a positive relationship between national income and the quality of a country's social institutions⁷. Countries with higher levels of gender equity, civic activism, intergroup cohesion, and interpersonal trust tend to have to have higher national income per capita, and vice versa.

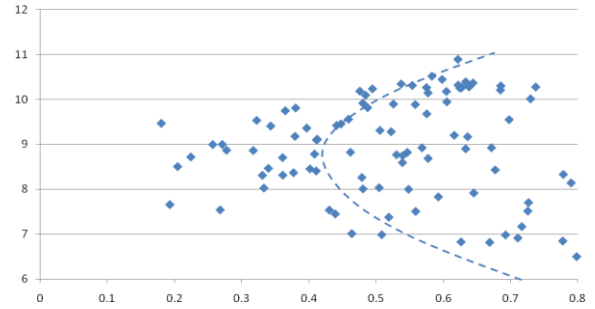
Figure 1 Social Institutions and Log GDP per Capita

⁷ The relationship between voluntary association and economic development is roughly curvilinear, with levels of engagement in community activities high in low-income, agrarian societies, falling among urbanizing, medium-income states, and high again among advanced, postindustrial societies. This result is consistent with an established sociological literature according to which initial development leads to the breakdown of traditional communal bonds ('communities of necessity'), based around village or parish networks, which are then reconstructed around friendship and voluntary organizations ('communities of choice'), once the cost of social organization falls with improved transport, telecommunications, and leisure hours (Tonnies [1887] 2001, Castells 2000, Inglehart 1997).

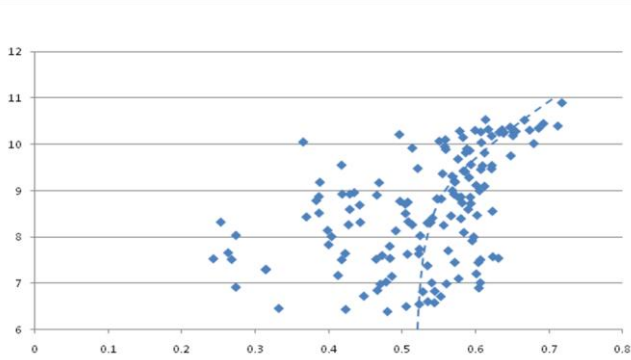
Civic Activism



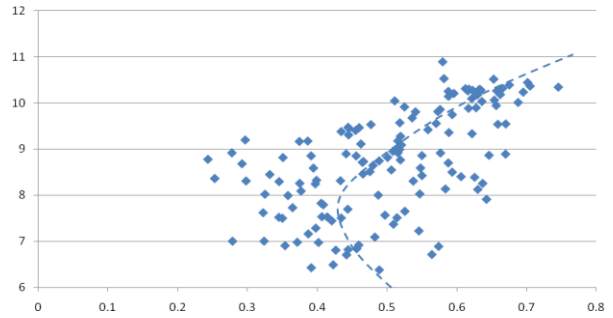
Clubs and Associations



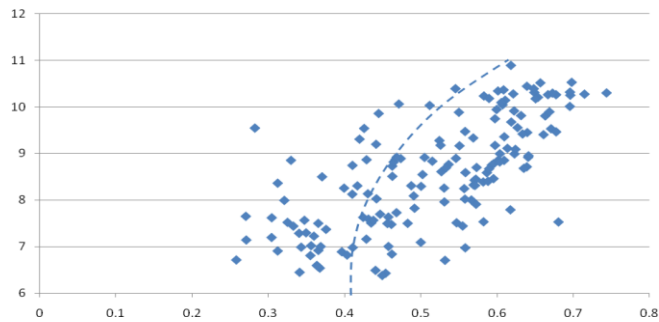
Intergroup Cohesion



Interpersonal Safety and Trust



Gender Equity



We can explore this association further by examining the results of a simple multivariate regression with each of the five social institutional clusters as our dependent variable, and a range of controls plus income per capita as independent variables. Results are shown in Table 2.

Table 2 **Regression Models, Social Institutions**

	Civic Activism	Clubs and Associations	Intergroup Cohesion	Safety and Trust	Gender Equity
Log GDP per capita	-0.157 (0.075)*	-0.719 (0.262)**	-0.081 (0.108)	-0.395 (0.116)**	-0.083 (0.07)
Log GDP per capita, squared	0.013 (0.004)**	0.041 (0.015)**	0.006 (0.006)	0.026 (0.007)***	0.006 (0.004)
Voice and Accountability, 2005	0.014 (0.01)	0.00 (0.034)	0.062 (0.016)***	0.01 (0.016)	0.023 (0.01)*
Ethnic fractionalization	0.008 (0.03)	-0.013 (0.102)	0.027 (0.046)	-0.093 (0.046)*	-0.008 (0.028)
Religious Fractionalization	0.07 (0.027)*	0.078 (0.092)	-0.049 (0.038)	0.033 (0.04)	-0.028 (0.026)
Linguistic Fractionalization	-0.071 (0.03)*	0.122 (0.098)	0.016 (0.044)	-0.037 (0.044)	-0.007 (0.028)
Fraction Protestant	0.118 (0.031)***	0.058 (0.102)	-0.059 (0.046)	0.025 (0.046)	0.136 (0.03)***
Fraction Muslim	-0.026 (0.028)	0.009 (0.098)	-0.015 (0.043)	0.124 (0.042)**	-0.012 (0.026)
Fraction Hindu	0.079 (0.042)	-0.032 (0.167)	-0.349 (0.08)***	0.12 (0.061)	0.008 (0.041)
Fraction Catholic	0.034 (0.028)	0.015 (0.081)	-0.091 (0.04)*	-0.034 (0.04)	0.118 (0.027)***
Fraction Orthodox	0.009 (0.047)	0.041 (0.122)	-0.056 (0.066)	0.033 (0.067)	0.018 (0.045)
Fraction Buddhist	0.124 (0.051)*	-0.182 (0.175)	-0.096 (0.062)	0.08 (0.061)	0.179 (0.041)***
Constant	0.864 (0.321)**	3.555 (1.133)**	0.795 (0.455)	1.945 (0.495)***	0.698 (0.297)*
N	111	75	108	102	118
Adj. r ²	0.82	0.20	0.49	0.62	0.81

The results confirm the positive relationship between economic development and the quality of social institutions. Moving from a medium to a high income status, as shown by the coefficient for the square of log GDP and the coefficient effects plotted on the graphs of Figure 1, is positively associated with all social institutional measures, and significantly so in the case of civic activism, safety and trust, and membership of clubs and associations. The only qualification is that during early stages of economic development, while societies are undergoing major social transformations such as urbanization, industrialization, and the demographic transition, significantly poorer outcomes develop in the cases of membership of clubs and associations, crime and trust, and civic activism. This curvilinear relationship,

while initially surprising, is consistent however with a long and established literature in both sociology and political science (Durkheim [1897] 1951, Huntington 1968, Inglehart 1997, Inglehart and Baker 2000).

Beyond a certain threshold, GDP per capita and social institutional quality are strongly positively related. This may be due to one of two reasons. First, social institutions may be the outcome of processes of sustained economic development. As societies become more affluent, individuals are empowered materially, educationally, and socially, and this leads to predictable changes in social norms and values in response to a new pattern of incentives. For example, citizens gain the human and financial resources needed to participate in civic activities; states can afford to finance formal institutions, such as the police and judiciary, to enforce laws and prevent inter-group violence; and women are able to receive an education and enter the labor market (Inglehart and Baker 2000). The converse perspective is that sustainable long-run economic growth is the product of a certain set of social institutions - the norms, precedents, and cultural expectations that accumulate over the course of a country's history. These may either distort or protect the pattern of incentives to engage in economically productive activity, depending upon whether they protect property rights, encourage work, and reduce transaction costs. Such a view is the cornerstone of the new institutional economics (NIE), according to which formal institutions (courts that protect property rights and enforce the rule of law) and informal institutions (social trust, cohesion, and voluntary activity) constitute long-run determinants of sustained capital accumulation, that is, economic growth (North 1991, Hall and Jones 1999). A more detailed discussion of the interdependent relationship between social institutions and economic growth is not the place of this paper, but will form an important basis of future research surrounding the indices of social development.

VII. Diagnostics

The long history of the use of composite indicators in the social sciences, in particular in the disciplines of psychology and demography, has led researchers to develop a range of diagnostic tests designed to assess construct validity and indicator reliability. These include

tests designed to help identify outlier indicators, assess the degree to which indicators reflect a single underlying dimension, and identify redundancies or assign weights among the indicator set. Notably, techniques such as confirmatory factor analysis, cluster analysis, and calculations of statistical leverage and influence are standard practices in the process of index design and analysis.

This section of the paper therefore provides an assessment of the construct validity of the social development indices, by firstly considering in greater detail the theoretical motivation for indices in reference to other studies, and then presenting the results of a range of tests designed to assess i) the presence of outlier variables or values; ii) the convergence of indicators together into the dimensions assigned to them by the indices of social development (convergent validity); iii) the appropriate weighting and removal, if required, among the indicators used.

7.1 Discussion of the Clustering of the Indicators

As the definition of almost all social science concepts are contested among practitioners, a key requirement of construct validity in the social sciences is that concepts be clearly defined and justified (Carmines and Zeller 1976). The taxonomy deployed in the Indices of Social Development is broadly consistent with previous attempts to classify the realm of social institutions, as well as the broader literature discussing the multidimensional nature of social capital (Stone 2001, Paldman, 2000, Narayan and Cassidy 2001). A recent and comprehensive attempt to provide a multidimensional metric of social institutions is the Social Capital Integrated Questionnaire (SCIQ), developed by Grootaert, Narayan, Jones and Woolcock (2004), which has been fielded in a number of the Living Standards Measurement Surveys (LSMS). The SCIQ offers six dimensions of social capital, listed as: i) groups and networks, ii) trust and solidarity, iii) collective action and cooperation, iv) information and communication, v) social cohesion and inclusion, and vi) empowerment and political action. Of these categories, the second (trust and solidarity) clearly parallels the social development construct for interpersonal safety and trust; the third (collective action and cooperation) parallels the construct for clubs and associations; the third (information and communication)

and sixth (empowerment and political action) together broadly constitute civic activism; while the fifth dimension (social cohesion and inclusion) corresponds with intergroup cohesion. Thus while the terminology may vary, basic taxonomies of social institutions often delineate similar norms and practices.

The breakdown of social institutions into the categories developed by Grootaert et al. (2004) and into those used in the social development indices reflects a lengthy process of debate and refinement within the literature on social capital. Paldam (2000), for example, summarizes an early distinction made between three basic aspects, those of social trust, cooperation, and networks. While social trust refers to abstract interpersonal norms of reciprocity, cooperation refers to specific instances of working together in voluntary associations and groups, and networks are simply one's web of personal connections that can be drawn upon for information, support, or favors. This distinction is ultimately reflected in the design of the social development indices, which separate trust into one dimension (interpersonal safety and trust), while placing cooperation in another (clubs and associations). The 'networks' component of social capital, meanwhile, is not included in the social development indices. This is due to theoretical concerns regarding the issue of whether social networks *per se* are beneficial for aggregate social welfare, or whether they only benefit the individuals within them at the expense of others (cf. Rubio 1997), and also simply due to the difficulty of measuring 'network density' on a national and cross-country level. In a more extensive overview of the social capital concept, Stone (2001) also separates 'norms' and 'networks', and within each of these two categories, distinguishes a range of sub-categories that are relevant to the distinctions made in the social development indices, such as whether networks are homogenous or heterogeneous, or whether norms of trust are generalized or only exist among familiars. These distinctions are also reflected in the five dimensions used in the social development indices; with interpersonal safety and trust, for example, referring only to generalized norms of cohesion.

7.2 Perceptions-based Versus Actionable Indicators

In recent years there has been a lively debate on the use of *perceptions-based* indicators, in particular in the study of quality of governance and above all the measurement of political corruption (Urrea 2007, De 2010). There have also been debates around the appropriate choice of indicators in development research, with practitioners highlighting the need to focus on so-called *actionable* indicators which are capable of being quickly improved through appropriate policy intervention. This section therefore discusses the nature of the indicators used in the indices of social development.

The sources used for the Indices of Social Development include indicators for a wide range of concepts relating to social institutions, using different techniques, and covering different groups of countries. In general, however, we can divide the data sources into two categories, reflecting the different methods via which the measures were generated: actionable indicators, which are based on direct measurement of social institutions and their outcomes, and perception-based indicators, which are based on assessments by public opinion surveys, private agencies and non-governmental organizations, of the nature of social institutions in that country. In general, actionable indicators may be preferable to perception-based indicators, in that they are more responsive to changes in underlying social conditions, and cannot be influenced by changes in perception independent of substantive social change. However, because some norms and practices are difficult to measure directly, perceptions measures are sometimes needed to supplement these data.

We can further subdivide two groups within these two categories depending on how the data are generated. Actionable indicators are either proxy variables based upon the measurable outcome of social institutions, or information on reported social behavior taken from nationally representative surveys, while perceptions-based measures can likewise be divided into two categories, public opinion and expert assessment, depending upon the nature of the underlying data source. A breakdown of the indicators used in each of the Indices of Social Development, by indicator type following this schema, is provided in Table 3.

Proxy variables in our dataset are those typically used in studies of social capital, such as per capita newspaper circulation, the density of international non-governmental organizations, or the reported number of ethnic or other violent street riots⁸. The use of proxy variables such as newspaper circulation depend upon reasonable inferences regarding the causes and consequences of social action, such as the assumption that greater newspaper readership reflects greater citizen propensity to engage in civic activism, or that violent riots reflect the breakdown of cohesion among social groups. The validity of these inferences is confirmed by the often high degree of correlation between each of these measures and other indicators of social institutional structure. Behavioral items taken from comparative, nationally representative survey projects include responses to questions such as the signing of petitions, domestic violence, or membership of voluntary associations.

Table 3. Types of Social Institutional Indicators

	Actionable		Perception-based	
	Proxy variables	Behavioral	Public opinion	Expert assessment
Civic Activism	Membership of international NGOs, per capita access to radios	"Here is a list of actions that people sometimes take as citizens. For each of these, please tell me whether you, personally, have done any of these things during the past year." - Attended a demonstration or protest march	"How much confidence do you have in civil society organizations in your country?"	Civics ratings on effectiveness of civil society organizations
Clubs and Associations	-	"Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an active member, an inactive member, or not a member?" - Development associations	"Would you say that, in your neighbourhood, people generally help one another?"	-

⁸ Key sources include the World Values Survey, founded in 1981, which currently provides comparative measures such as social trust, tolerance of minorities, and voluntary associational membership for almost 90 societies around the world, as well as regional survey projects, such as Latinobarometer, founded in 1996, Afrobarometer, founded in 1999, and Asian Barometer, founded in 2003, which cover a joint total of 49 societies.

Intergroup Cohesion	Number of newspaper-reported ethnic riots	"On this list are various groups of people. Could you please sort out any that you would not have as neighbors?" - those of another ethnicity, those of another religion	"Is your ethnic group ever treated unfairly in society?"	International Country Risk Guide rating on levels of ethnic or religious tensions
Interpersonal Safety and Trust	Deaths from homicide	"Over the past five years has anyone taken something from you, by using force, or threatening you?"	"Generally speaking, would you say that most people can be trusted or that your can't be too careful in dealing with people?"	Economist Intelligence Unit rating on 'level of social mistrust'
Gender Equity	Ratio of female to male mortality rates, ratio of female to male school enrolment rates	Proportion of wives reporting having suffered an act of domestic violence in the past year	Proportion of women who feel that "women have the chance to earn the same salary as men in their country"	Cingranelli-Richard rating on women's social rights

The first form of perception-based data come from the nationally representative public opinion surveys just mentioned, and include responses to those questions which ask the respondent to give their opinion on some issue, such as their level of confidence citizens feel in their civil society organizations, the level of discrimination women or minorities feel they encounter in their daily lives, or the trust people have in their fellow citizens. While 'softer' than survey questions on actual behavior, these items allow researchers to tap into a range of additional issue areas that harder data may be lacking. The second category consists of numerical ratings produced from expert assessments, in which academics, non-governmental organizations, and private rating agencies assess the nature of social institutions across countries. Such assessments have become more widespread in recent years, as researchers have sought to make social facts visible to quantitative analysis⁹. These efforts to code descriptive assessments of the quality or otherwise of social institutions into numerical form open up a further rich source of information for researchers wishing to comparatively assess the social environment of different countries.

⁹ The Minorities at Risk project, for example, was started in 1986 and has been updated over three successive waves, providing comparative measures of discrimination and exclusion of minority groups in 118 societies across the world. The International Country Risk Guide has since 1980 provided assessments of a range of social variables, in addition to purely political and economic factors, such as the level of ethnic or religious tensions. Meanwhile, since 2003 the Civicus civil society network has been developing a range of indices for the health of civil society.

Table 4 below summarizes the relative prevalence of each of these data categories in the database of social development indicators. Shown are the proportion of country-year data points in each category. Overall, actionable proxy variables account for half of the data used, with survey data on social actions and behavior forming an additional quarter. The remaining quarter of data points are split between perceptions-based indicators derived from public opinion surveys and expert assessments.

Table 4. Proportion of Data Points in Each of the Indices, by Type

	Actionable		Perception-based	
	Proxy variables	Behavioral	Public opinion	Expert assessment
Civic Activism	0.71	0.24	0.00	0.05
Clubs and Association	0.00	0.79	0.21	0.00
Intergroup Cohesion	0.34	0.13	0.02	0.51
Safety and Trust	0.37	0.12	0.21	0.22
Gender Equity	0.75	0.07	0.04	0.15
AVERAGE	0.48	0.27	0.10	0.15

Not all sources provide observations for every indicator in each country, but together, these data sources allow for comprehensive estimates of the nature of social behavior and norms of interaction across a broad range of societies. Thus, for the *civic activism* measure, the score for Bolivia in 2005 is estimated using 9 pieces of information from 4 different sources: the LSE Civil Society Yearbook figures for the number of International NGO secretariats based in that country (9) and connections to international NGOs (1216); Latinobarometer survey items for participation in petitions (48%), lawful demonstrations (45%), and the average number of days in the week that a respondent follows current affairs via television (4.78/7), newspapers (2.27/7), and radio (4.73/7); UNESCO data on the number of daily newspaper titles, per capita (2.13); and a *Civicus* Civil Society assessment rating the quality of civil institutions with regard to their structure, environment, values, and effectiveness (1.65/4). These data are aggregated by source and combined using the matching percentiles method, to generate a composite score for Bolivia which is roughly in the middle of the world distribution, and within Latin America and the Caribbean, falls roughly halfway between Peru and Ecuador¹⁰.

¹⁰ For further reference, the high and the low scores within Latin America and the Caribbean are held by the Bahamas, and Haiti, respectively.

7.3 Selection of Indicators -- What's in?

With the foregoing theoretical basis underpinning the division of indicators into five constructs, the second step in building the indices was to assign the circa. 200 indicators into these five categories. However, doing so on an *ad hoc* basis risks potentially assigning indicators into the wrong category, as well as including indicators which, due to outlier observations, weak cohesion with the latent construct, or unreliable measurement, may have the effect of excessively leveraging the final index score of one or more countries, or which simply may not fit into any dimension. In the following sections we deal with the problem of testing the reliability of indicators and the accuracy of observations.

In section 7.3.1 below, we employ diagnostic tests designed to identify outlier observations from within any one of the approximately 200 indicators. We then move on in section 7.4 to examine the reliability of the individual indicators themselves.

7.3.1 Outlier Observations

Several diagnostic tests enable us to identify outlier observations. First, we are able to use the standard deviations of the variables to identify cases that are clearly outside of the range of a normal distribution. This approach, while satisfactory for an initial summary of the data, fails to highlight whether the presence of such outliers has undue effect in the generation of final index scores. Second, therefore, we are able to supplement such an analysis with statistical tests of both leverage and influence. This enables us to better determine which values for each of the indicators are outliers, due to either misreported data or measurement error, and therefore constitute candidates for deletion.

7.3.1.1 Standard Deviations

The first such test is to identify outlier observations using the standard deviation. A random variable which is normally distributed will have 99.7% of its data within 3 standard

deviations from the mean. We explore each variable individually to determine which observations have values greater or less than 3 standard deviations. To test for this we first convert each variable into a standard normal.

Because many of the variables are clearly not normally distributed (most frequently binary or other non-continuous variables) upon inspection with a graphical analysis of a quantile-quantile plot, we also employ a nonparametric standard deviation test for outliers. According to Chebyshev's inequality, any randomly distributed variable of (almost) *any* distribution will have 95% of its data distributed within 4.5 standard deviations of the mean. We test for observations falling outside of these bounds as well.

Results are reported in Appendix I. In all we found that 31 of our variables had at least one outlier according to our standard deviation tests. The existence of outlier cases does not necessarily mean that a particular datum is inaccurate, as a variable with 100 cases which follows a distribution with 95% of observations within 4.5 standard deviations of the mean ought naturally to have five observations fall outside these bounds. However, as extreme distributions are unusual, observations outside these bounds are typically indicative of inaccurate or unreliable data.

7.3.1.2 Leverage and Influence

The second and more rigorous approach to identify outliers is to use calculations of statistical leverage and influence. Leverage and influence constitute a useful diagnostic tool for identifying outliers. The method is straightforward—regress an index on each of its corresponding indicators in a series of bivariate regression and determine which observations (countries) carry undue weight in that regression.

Where in a simple average each point contributes the same weight, such is not the case in a regression framework. Though each index is a nonlinear function of the collection of its corresponding indicators, using a bivariate regression in this way is still useful for exploring outliers.

7.3.1.2.1 Leverage

Points which have high leverage are those with unusual x -values. Though points with high leverage alone do not affect the estimates of regression coefficients, these may affect the values of model summary statistics such as goodness-of-fit and standard errors of coefficients. Clearly we would prefer that the effect an indicator has on its index be representative of the entire sample of observations for that index.

Leverage for a point i is the distance of the i th observation from the center of the x -space, and obtained using the principal diagonal of the “hat” matrix as $h_{ii} = \mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$. The average of distance from the center of the x -space is given as $\bar{h} = p/n$ where p is the number of parameters in the regression including the intercept (here p is 2 for our bivariate regressions) and n is the number of observations in the regression. A point with distance from the x center greater than a cutoff $h_{ii} > 2p/n$ is considered a leverage point. High leverage points mark a potential for influencing regression coefficients but may not affect the coefficient at all (e.g. if it lies directly on the regression line, even though isolated in x -space from the rest of the data).

7.3.1.2.2 Influence—Cook’s Distance

Where leverage points indicate a potential for changing regression coefficients, influence points are those which actually do affect coefficients. Influence considers both x - and y -space by examining how the regression would behave differently if the i th point were deleted—denoted as (i) . In this way, the three main modes for measuring influence are all deletion diagnostics. One influence measure, Cook’s Distance¹¹, uses a measure of the squared distance between the least-squares coefficient based on all n points and the regression coefficient calculated without point i . Each point thus has its own measure of Cook’s Distance, written as:

¹¹ Common alternate deletion diagnostic measures of Influence include ΔFBETAS and DFFITS , which measure difference in beta in standard deviation units and the difference in fitted y -values, respectively.

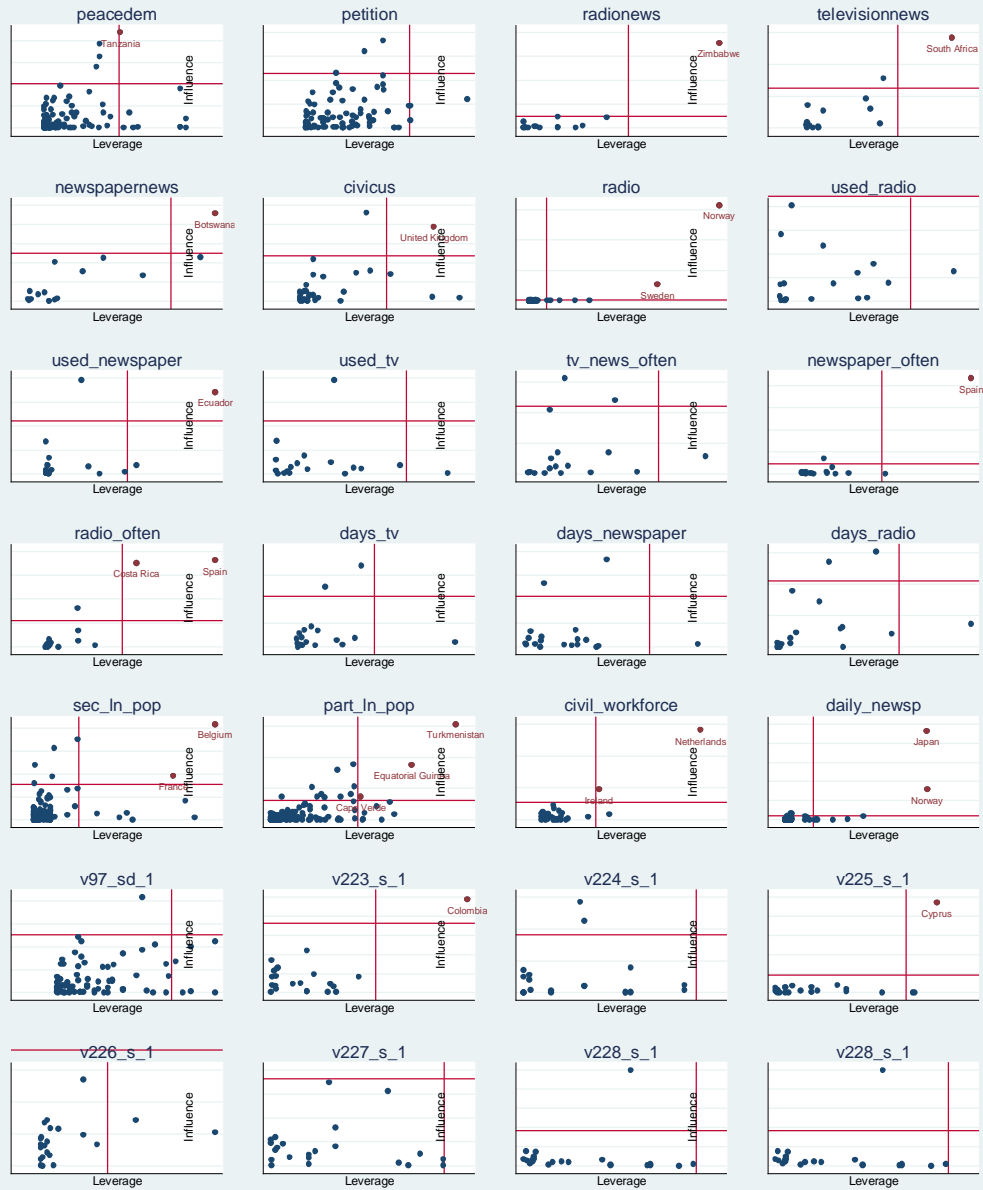
$$D_i = \frac{(\hat{\beta}_{(i)} - \hat{\beta})' X X' (\hat{\beta}_{(i)} - \hat{\beta})}{p \cdot MSE}$$

- where MSE is the mean squared error of the regression. The cutoff for Cook's Distance is $4/n$, and anything larger merits further inspection.

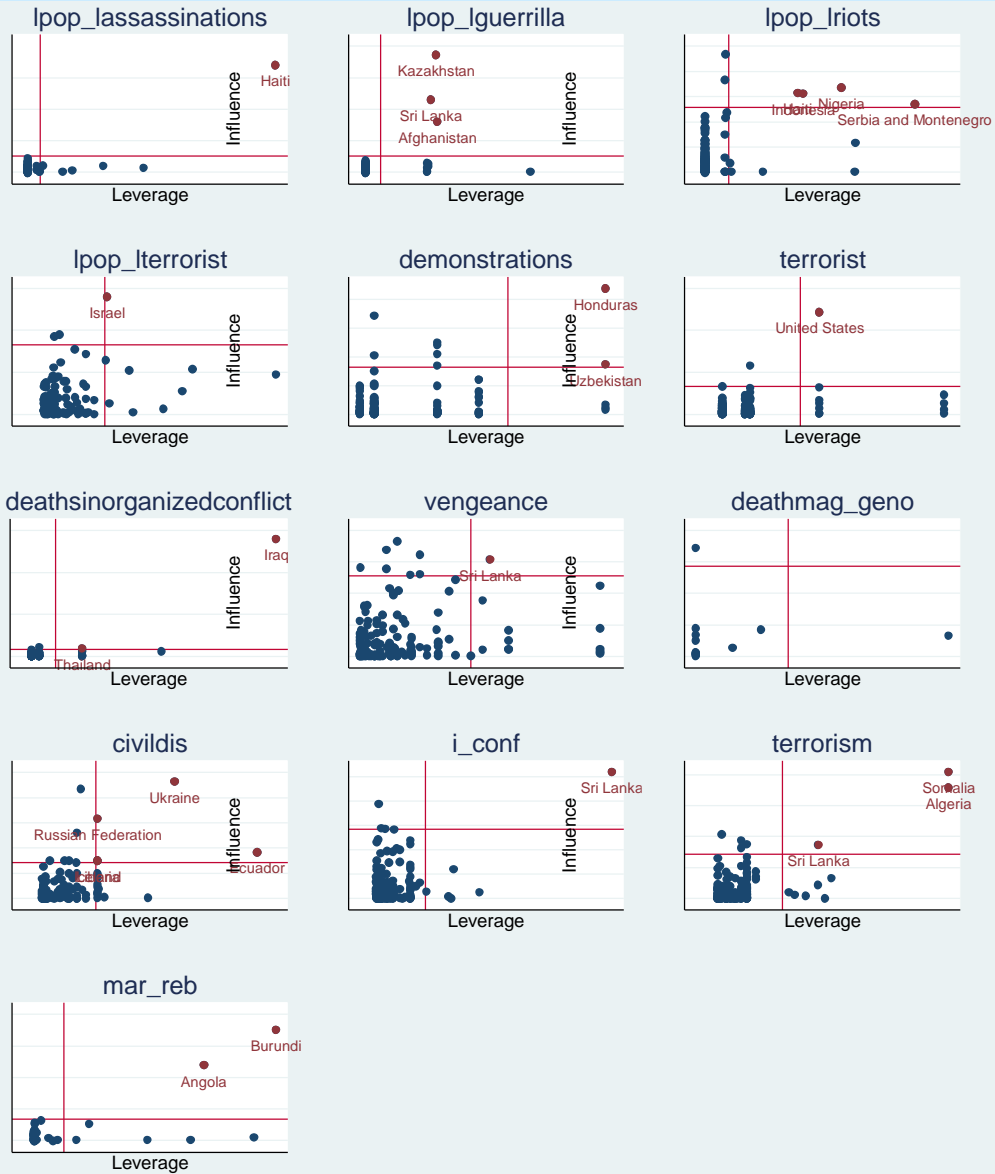
7.3.1.2.3 Graphical Analysis of Leverage and Influence

Though influence statistics are more useful a leverage statistics because they highlight observation that actually do affect regression results, points which have both high leverage and high influence merit further inspection as potential outliers. Consequently, these two diagnostics are usefully graphed together. The following sets of scatter plots, grouped by index, show leverage and influence for each of the index's component indicators. Countries which fall above the thresholds are labeled.

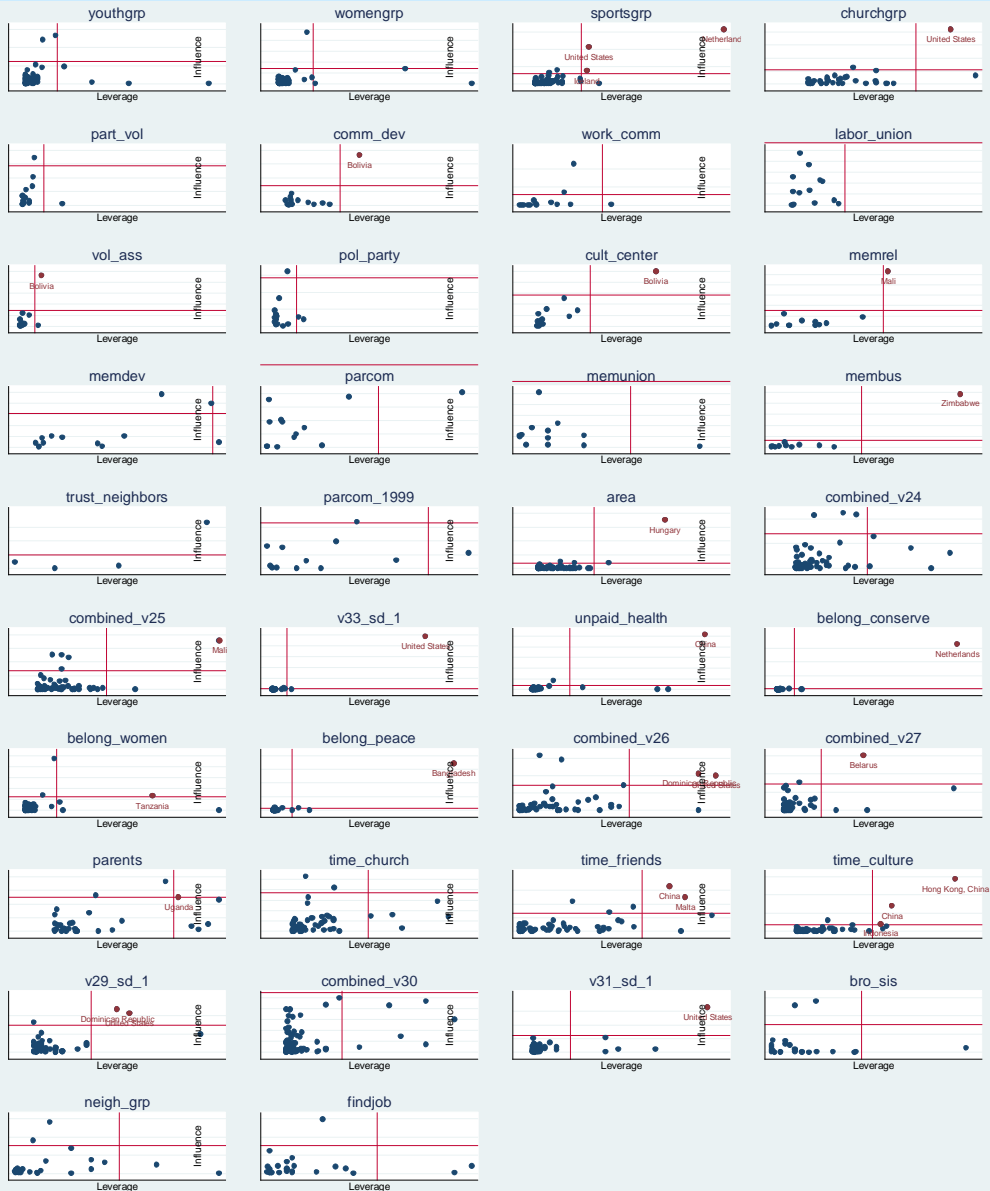
Civic Variables – Leverage and Influence



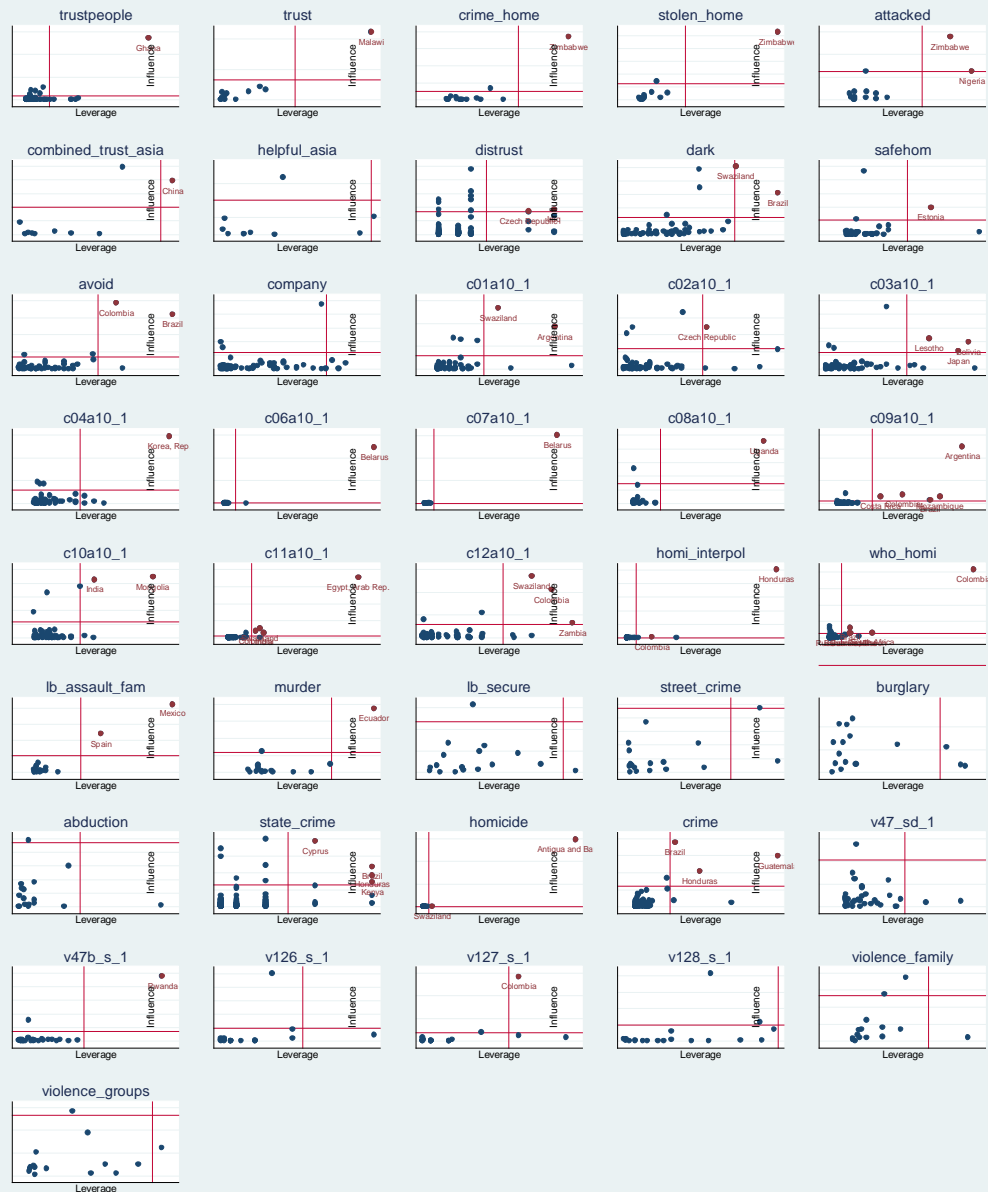
Cohesion Variables – Leverage and Influence



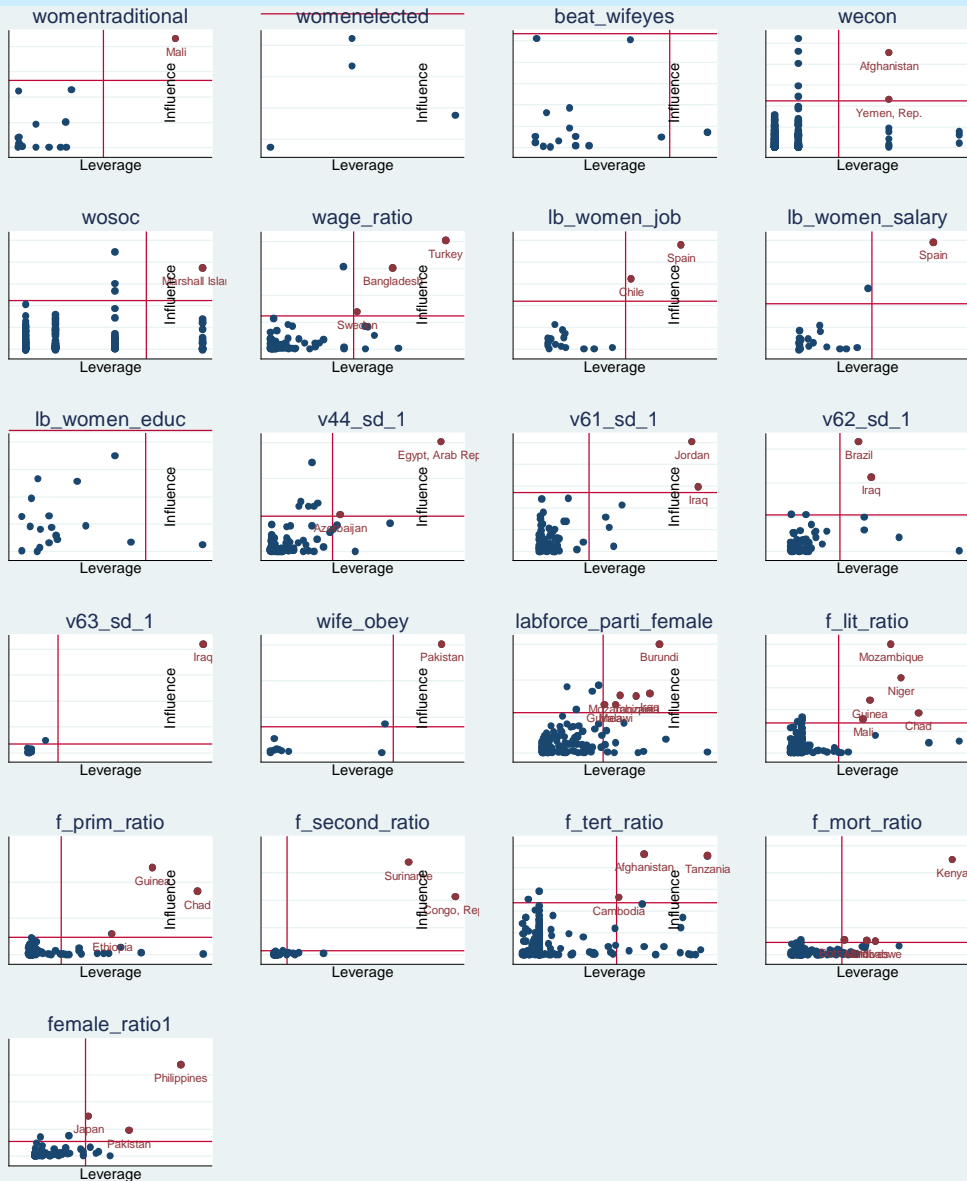
Voluntary Association Variables – Leverage and Influence



Interpersonal Safety and Trust Variables – Leverage and Influence



Gender Equity Variables – Leverage and Influence



7.3.2 Results

Having compared the results of the standard deviation diagnostics and the leverage and influence diagnostics, we find no single source of data that is consistently unreliable, though we do highlight the following observations as potentially replaceable with missing values, on account of specific concerns with reported country data in Table 5. We drop questionable cases, though the impact of these changes is minimal, and unlikely to substantially affect final country results.

Table 5. Country Observations with Outlier Leverage/Influence

Variable	Country	Problematic Entry	Cause
Over the past five years, did anyone actually get into your house or flat without permission and steal or try to steal something?	Belarus	Value of 100% found in dataset; value not possible.	Probable data entry error.
Over the past five years, do you have any evidence that someone tried to get into your house or flat unsuccessfully -- for example, damage to locks, doors or windows, or scratches around the lock?	Belarus	Value of 100% found in dataset; value not possible.	Probable data entry error.
"Could you tell me whether you are an active member, an inactive member or not a member of [other voluntary organizations]?"	United States	Value of 90% in 2005 far higher than previous years (circa. 30%) and other countries	Possible miscoding; early release of data.
UNCJIN homicide rate	Antigua and Barbuda	Value of 330 per 100,000 well above other countries; other estimates of homicide rate in Antigua circa. 6.9/100,000	Probable data entry error.
Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?	Ghana	Value of 90% trust improbable; next highest (Sweden) is circa. 60%	Possible miscoding of polarities; early release of data.
"Could you tell me whether you are an active member, an inactive member or not a member of [other voluntary organizations]?"	Armenia	Value of 100% found in dataset; value not possible.	Possible miscoding; early release of data.
"On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors." People of a different race	Hong Kong	Value of 73% improbable;	Possible miscoding of polarities; early release of data.

"On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors." Immigrants or foreign workers	Hong Kong	Value of 79% improbable;	Possible miscoding of polarities; early release of data.
---	-----------	--------------------------	---

7.4 What's Where?

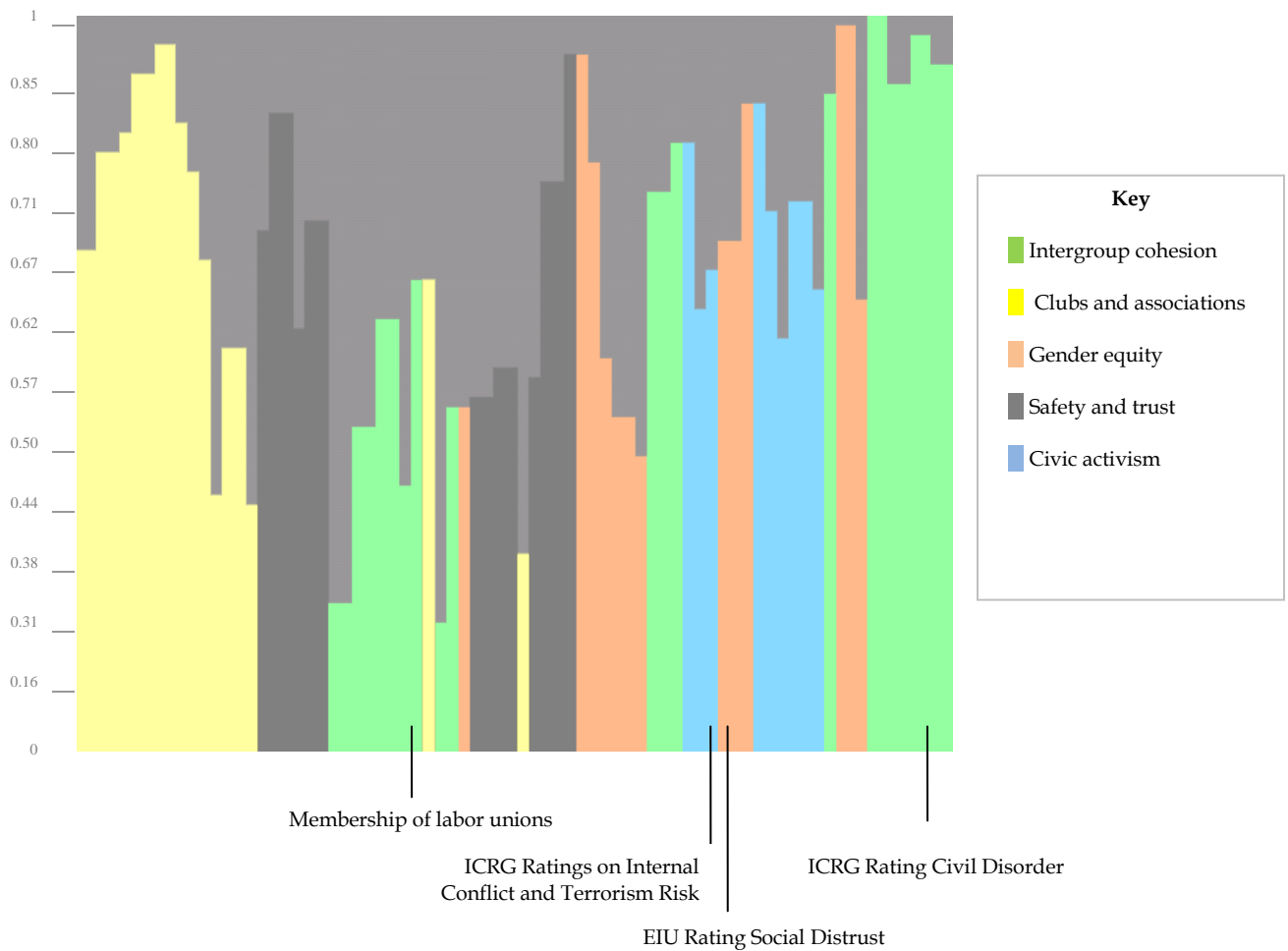
In designing the initial set of indices of social development, 200 indicators were assigned *a priori* into five categories. In this section of the diagnostics, therefore, we consider more closely the issue of whether indicators are correctly assigned: that is, whether the indicators used in each cluster ‘fit’ empirically with the other indicators in that same grouping, or constitute outliers to be reassigned or removed. We do so by deploying a range of diagnostic techniques that include factor and cluster analysis.. As a prerequisite to these analyses, we employ a regression imputation methodology for the missing data, used only for these diagnostic analyses and not for the generation of the indices themselves.

7.4.1 Hierarchical Cluster Analysis

Our first method for detecting the ‘natural’ statistical grouping of the indicators is hierarchical cluster analysis. Hierarchical cluster analysis groups items (variables) based on similarities using a decision rule which dictates threshold levels for inclusion. We use Pearson correlations to express similarity between individual items or items within groups. This analysis employs alternative decision rules known as the “complete link” and the “average” rules. Complete link dictates that the threshold correlation for clustering two groups (or two items or an item to an existing group of items) is the Pearson correlation of the two most disparate elements with the groups—that is the two items with the lowest correlation.. The average rule results in a threshold of group clustering at the average correlation of all elements between the two groups. Clusters are grown from twigs to trunk by joining the most similar items first and then searching for the next most similar item or group of items using the decision rule.

Cluster analysis is often presented graphically as either a similarity tree or as a “cityscape” diagram. They are usefully read from trunk to twigs. Low-level breaks indicate conceptually separate items. High-level matches indicate item redundancy. Cohesively colored groupings give a strong endorsement for our agglomeration of the index items, and stripy groupings and ‘misplaced’ items indicate areas in which our index holds together less well.

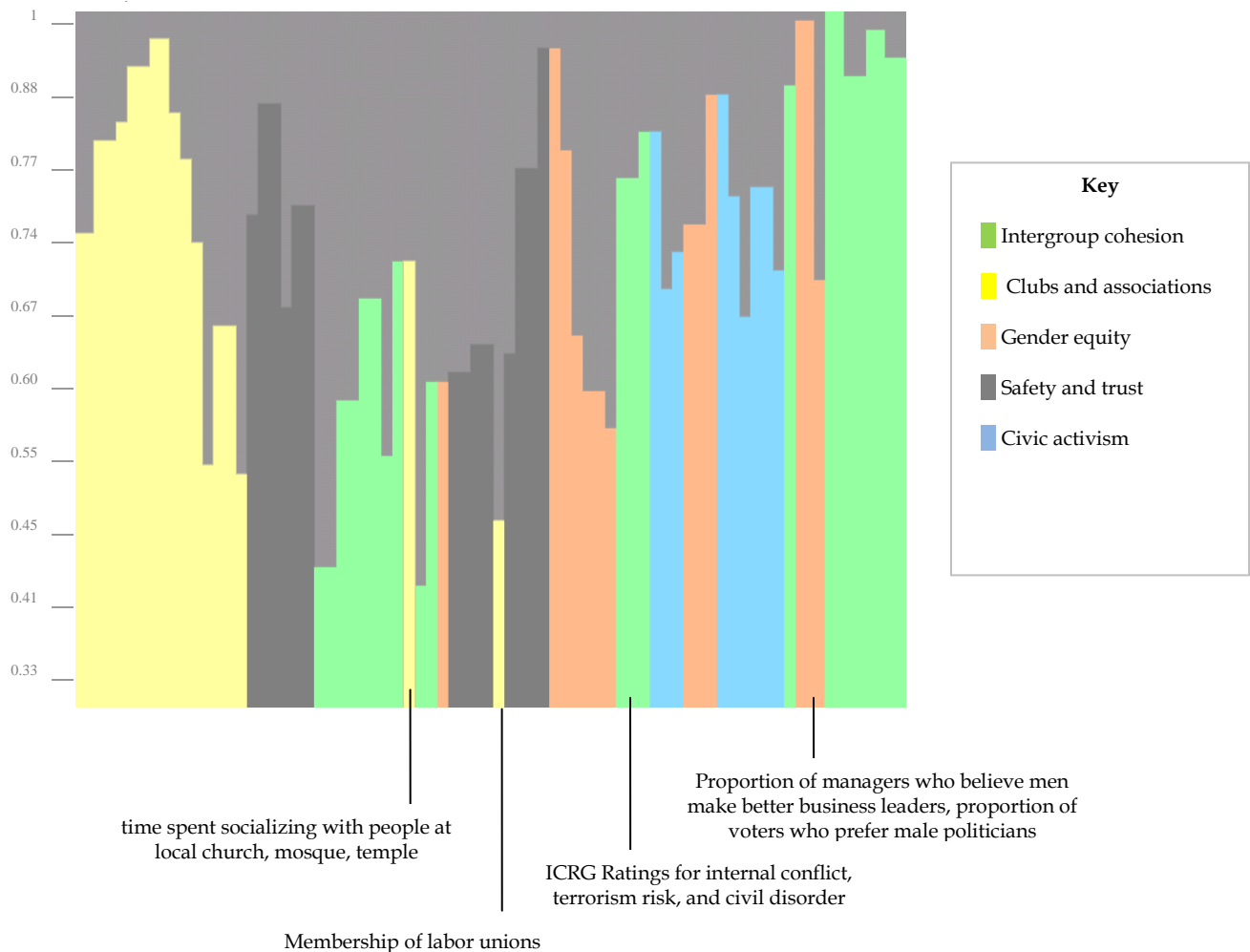
Figure 2 Hierarchical Cluster Analysis, Average Rule



Each vertical line represents one indicator (variable). Scale of values on y-axis not intervally spaced.

Figure 2 above shows the results of a hierarchical cluster analysis using the average rule, using all indicators of the Social Development Indicators Project for which data exist for at least 30 countries. It can be seen that there is strong clustering among the safety and trust, clubs and associations, and civic activism groupings. Indicators for intergroup cohesion and gender equity are more dispersed. Using this method, we are able to identify several indicators that do not belong in their clusters, highlighted in the figure above.

Figure 3 *Hierarchical Cluster Analysis, Complete Link Rule*



Each vertical line represents one indicator (variable). Scale of values on y-axis not intervally spaced.

Figure 3 shows the results of a complete link hierarchical cluster analysis, using all indicators of the Social Development Indicators Project for which data exist for at least 30 countries. It can be seen that there is strong clustering among the clubs and associations and civic activism groupings. Indicators for gender equity are widely dispersed; safety and trust and intergroup cohesion seem to divide into two distinct clusters. Again, indicators that fall outside of their clusters are highlighted in the figure. Two sets of indicators emerge as consistent outliers using the Hierarchical Cluster Analysis method: the rate of membership of labor unions, and the International Country Risk Guide (ICRG) ratings for internal conflict, the risk of terrorism, and civil disorder.

7.4.2 Factor Analysis

Our second method for identifying the fit of the indicators is factor analysis. Factor analysis is used to uncover the latent structure in a set of variables. The Principal Factor Analysis (used here) seeks the least number of factors which can explain the greatest amount of shared variance among the variables.

We use Exploratory Factor Analysis to investigate a potential structure for how the variables group together, as well as to identify whether any variables simply don't belong in our framework of social development. In exploratory factor analysis no priors are imposed on which elements are latent to a given construct, and the factor analysis is used to indicate which variables in the item space should be combined together to form an index. It also provides a mathematical indication on the appropriate number of indices to be formed; that is, it can reveal the dimensionality of the latent construct. Exploratory factor analysis can also help in identifying outlier variables that do not fit within particular subindices, through examination of the factor loadings; where a variable is not loaded highly in common with other variables in a component, this is indicative of weak convergent validity. Finally, factor analysis can help inform insofar as factor analysis extracts the latent components that underlie a set of data, and this can include useful information on the redesign of indices, both in terms of content indicators, but also the conceptual framework that may explain why certain indicators pattern together in a common fashion and others do not.

We begin by using the results of the factor analysis to identify outlier variables within the data. Results of the factor analysis, using the Principle Axis method, are displayed below in Table 3.0. 6 factors were extracted, using the varimax rotation to obtain optimal results. As indicators have not been uniformly repolarized, no meaning should be attributed to the presence of a positive or a negative sign on the factor loading. Factor loadings have been highlighted to reflect the association of a particular cluster with that factor: blue cells represent indicators in the civic activism cluster, grey the interpersonal safety and trust cluster, yellow the clubs and associations cluster, orange the gender equity cluster and green the intergroup cohesion clusters. To better illustrate patterns, items in bold show the highest correlation across the factors for each indicator; loadings of less than or equal to 0.2 are blanked.

Table 6. Results of Factor Analysis

Principle axis factoring, varimax rotation, first 6 factors only withdrawn.

Variable	Factor					
	1	2	3	4	5	6
CSO workers as percentage of workforce	0.55	0.35	0.43	0.24		
Civicus civil society ratings	0.11		0.52	0.35		
Percentage who would 'join boycott'	0.64		0.46			
Percentage who would 'sign petition'	0.3	0.69	0.26			0.32
Percentage who would 'demonstrate'	0.03	0.7		0.31		
Newspaper circulation, per capita	0.65		0.5			0.3
International NGO secretariats / capita	0.76			0.35		
International NGO membership / capita	0.19		0.59			
Radios per capita	0.47		0.62	0.28		
Proportion who 'feel safe in their home'	0.57		0.58		0.24	0.29
Perceived level of social fairness		0.87	0.12			
Proportion who 'avoid places' in their area		0.36	0.85			
Proportion who feel 'safe' after dark	0.47	0.23	0.7			
Burglary rate			0.85			
Robbery rate			0.89			
Proportion feel crime is business constraint	0.22		0.13			0.59
Proportion who 'take company' going out	0.52	0.29	0.57	0.34		
Car theft rate (of owners)			0.81			0.24
Theft from car rate (of owners)		0.51	0.58			0.23
Damage to car rate (of owners)			0.66			0.26
Motor theft rate (of owners)		0.44	0.73			
Attempted burglary rate			0.86			
Personal theft rate			0.65			0.52

Variable	Factor				
Rate of sexual offence		0.9	0.06		
Assault rate			0.09	0.3	0.21
Perceived level of social trust	0.24		0.45	0.34	
Homicide rate (Interpol)	0.4	0.67	0.37		
homicide rate (UN figure)		0.87	0.21	0.24	
Rating on level of social trust (EIU)		0.7	0.12	0.22	0.45
State Department crime rating		0.72	0		0.41
Homicide (WHO figures)		0.71	0.2	0.49	
Frequency of socializing with relatives	0.38	0.59	0.08	0.44	
Membership of voluntary health assoc.	0.51		0.37	0.39	
Membership of other voluntary assoc.	0.15	0.31		0.45	0.28
Membership of women's groups	0.38	0.73	0.36		
Membership of peace groups	0.52	0.55		0.32	0.25
Membership of environment groups (2)	0.33	0.35	0.4	0.56	
Membership of youth clubs	0.11	0.63	0.62		
Membership of sports clubs (1)	0.54			0.68	
Membership of arts, cultural clubs	0.5	0.31		0.29	
Membership of labor unions	0.33			0.79	
Frequency socializing in religious context	0.08		0.76		
Frequency socializing in clubs	0.39	0.39	0.44		0.44
Proportion living in 'helpful neighborhood'	0.75	0.21	0.34		
Membership of environment groups	0.58	0.26			0.27
Membership of religious groups	0.82			0.28	
Membership of sports clubs (2)	0.62		0.49		
Frequency socializing with friends	0.68			0.3	
Membership of human rights organizations	0.58	0.33		0.24	
Membership of professional groups	0.76	0.3			0.22
Male-female wage gap		0.63			
Voters saying 'men make better politicians'		0.66			
Proportion of female administrators (%)		0.24		0.62	
% parents 'boy more right to education'		0.24		0.63	
% managers 'men more right to a job'	0.64	0.44		0.22	0.24
Proportion of female professionals (%)		0.59			
Male-female secondary enrolment ratio	0.56	0.03	0.41	0.21	
Male-female literacy rate ratio	0.29	0.02	0.28		0.42
Male-female tertiary enrolment ratio		0.24		0.3	
Women's economic rights, rating	0.54	0.26			0.37
Women's social rights, rating		0.31		0.24	0.64
Male-female primary completion ratio	0.21	0.06			0.66
Females as percentage of the labor force	0.39	0.42			0.43
Male-female mortality rate ratio	0.34	0.21			0.22
Prop. who refuse neighbor of other religion	0.4	0.39			0.53
Prop. who refuse a neighbor with AIDS	0.38	0.3			0.54
Prop. refuse neighbor of other race or caste	0.24	0.3	0.21		0.43
Prop. refuse neighbor who is a migrant	0.53			0.48	0.29
Prop. refuse neighbor who is homosexual	0.44	0.25	0.26		

Variable	Factor				
Minority discrimination (MAR dataset)	0.65		0.4		
Minority exclusion (MAR dataset)	0.58	0.26			
Minority rebellion (MAR dataset)	0.68	0.21			
Terrorist acts (EIU rating)	0.53			0.44	
Risk of violent demonstrations (EIU rating)	0.54	0.35		0.37	
Terrorism risk (EIU rating)	0.15	0.62			0.36
Number of civil conflict deaths	0.1	0.24		0.23	0.23
Civil disorder rating (ICRG)	0.04			0.45	
Internal conflict rating (ICRG)	0.02				
Terrorism risk (ICRG)	0.67		0.39		
Level of ethnic tensions (ICRG)	0.46	0.23		0.21	
Level of religious tensions (ICRG)	0.55			0.34	
Uneven development along group lines	0.21	0.59			0.26
Legacy of intergroup vengeance (FFP)	0.59		0.32		
Number of reported assassinations / capita	0.69				
Number of reported guerrilla acts / capita	0.05		0.22	0.59	0.22
Number of reported violent riots / capita	0.63	0.24	0.5		

Low factor loadings within any cluster are indicative of weak fit within any given factor. The factor analysis also helps to identify indicators that are particularly strongly associated with a particular item cluster. Fifty percent of the highest correlations (those in bold) fall into our *ex ante* assignment of indicators, and seventy-five percent of the indicators in the pre-assigned groups are correlated with the assigned item with a factor loading of greater than 0.2. This broadly confirms that our groupings are reasonable. Select Indicators which are loaded especially strongly (>0.65) or weakly (<0.1) are given below in Table 7. Clearly our taxonomy isn't perfectly confirmed by the factor analysis, but it is quite favorable overall.

Table 7. Strong and Weak Indicators, Factor Analysis Results

Strong Indicators	
Civic Activism	INGO secretariats (LSE), daily newspaper circulation (UNESCO)
Interpersonal Safety and Trust	Burglary, robbery rates, and car theft rates, share of people who avoid certain places in their area
Clubs and Associations	Proportion of respondents living in a helpful neighborhood, membership of professional or religious groups, and frequency socializing with friends.
Gender Equity	Male-female wage gap, and voters who agree with the statement that “men make better politicians.”
Intergroup Cohesion	Minority discrimination and rebellion (MAR), risk of terrorism (ICRG), and the number of reported assassinations per capita.
Weak Indicators	
Civic Activism	Willingness to demonstrate
Interpersonal Safety and Trust	Sexual offense and assault rates, and the US State Department’s crime rate
Clubs and Associations	Frequency socializing in a religious context
Gender Equity	Male-female literacy, primary completion, and secondary enrolment ratios
Intergroup Cohesion	Civil disorder rating and Internal conflict rating,(ICRG), and the number of guerrilla attacks per capita

Confirmatory Factor Analysis

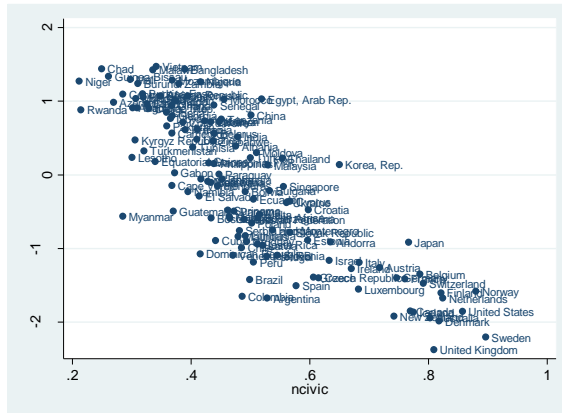
Our second use of the factor analysis results is to conduct confirmatory factor analysis, that is, to assess the appropriateness of the six indices that have been drawn *a priori* from the social development indicators database based on a purely statistical criterion.

These results are shown in figures 4.0 below. It can be seen that the factors extracted from the set of social development indicators replicate to a remarkable degree five of the six social development indices. The first factor correlates ($r = 0.82$) with the civic activism cluster, and also ($r = 0.78$) with the gender equity measure. The second factor correlates strongly with the safety and trust measure. The third factor correlates strongly with the clubs and associations measure. The fifth factor correlates strongly with the intergroup cohesion measure. Only the fourth and sixth factors do not correlate with the existing social development indices.

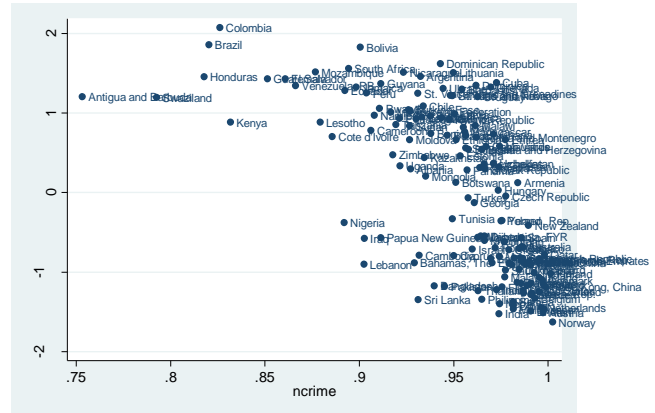
These results provide exceptionally strong support for the choice of categories drawn from the social development indicators. For the factors drawn schematically from the dataset to parallel so closely those drawn *a priori* from the data is an outcome that could certainly not be taken for granted at the outset, and provides strong confirmation for the assumption of the matching percentiles aggregation methodology, that each index reflects one single underlying latent dimension or process.

Figures 4.0 Factor Analysis Results and Social Development Indices Compared

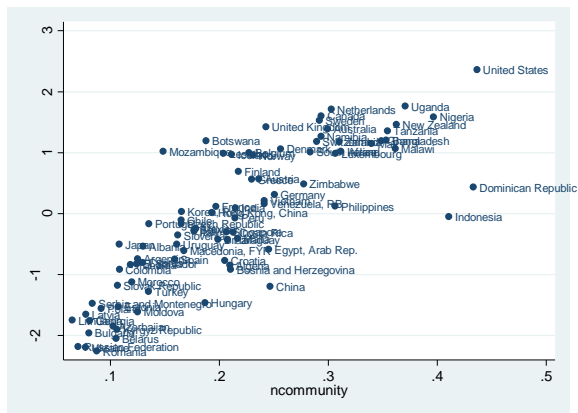
Factor 1 and Civic Activism



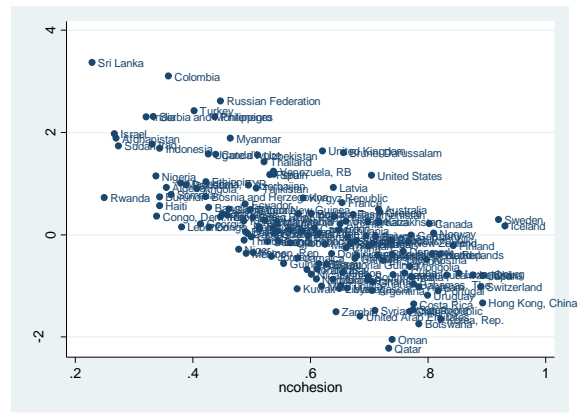
Factor 2 and Safety and Trust



Factor 3 and Clubs and Associations



Factor 5 and Inter-group Cohesion



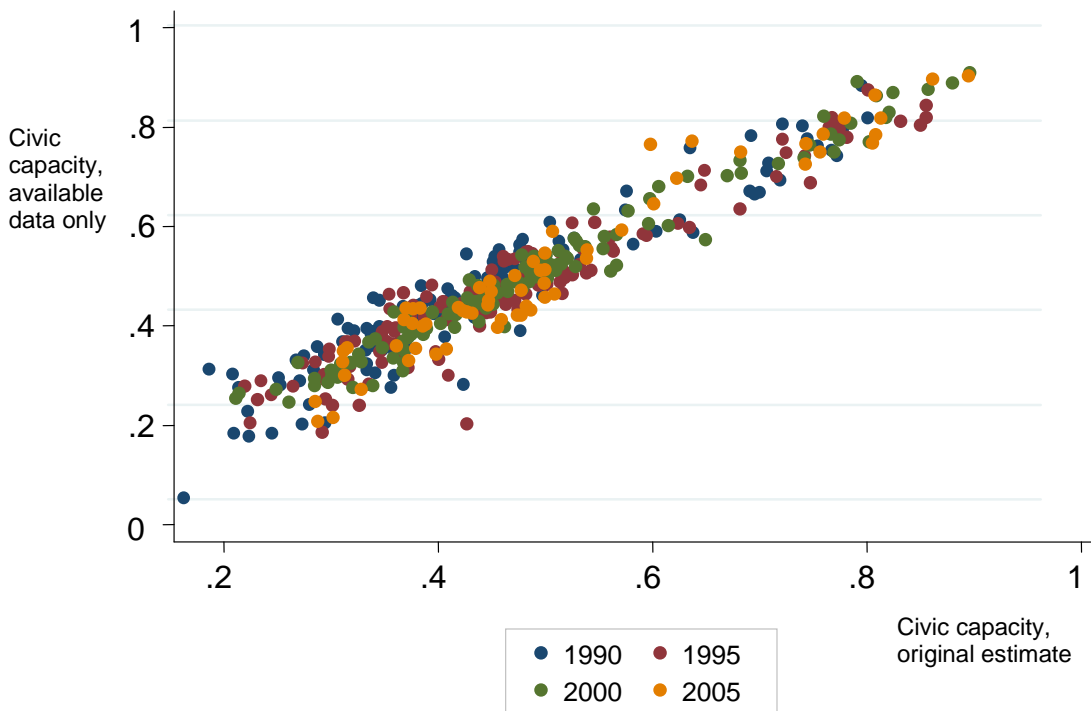
7.4.3 Testing the Robustness of the Aggregation Methodology – Sensitivity Analysis

We also need to test the robustness of the matching percentiles technique itself, in particular to ascertain whether it produces consistent results when testing the robustness of the indices to the inclusion of 'lagged' scores where data for a particular year is not available.

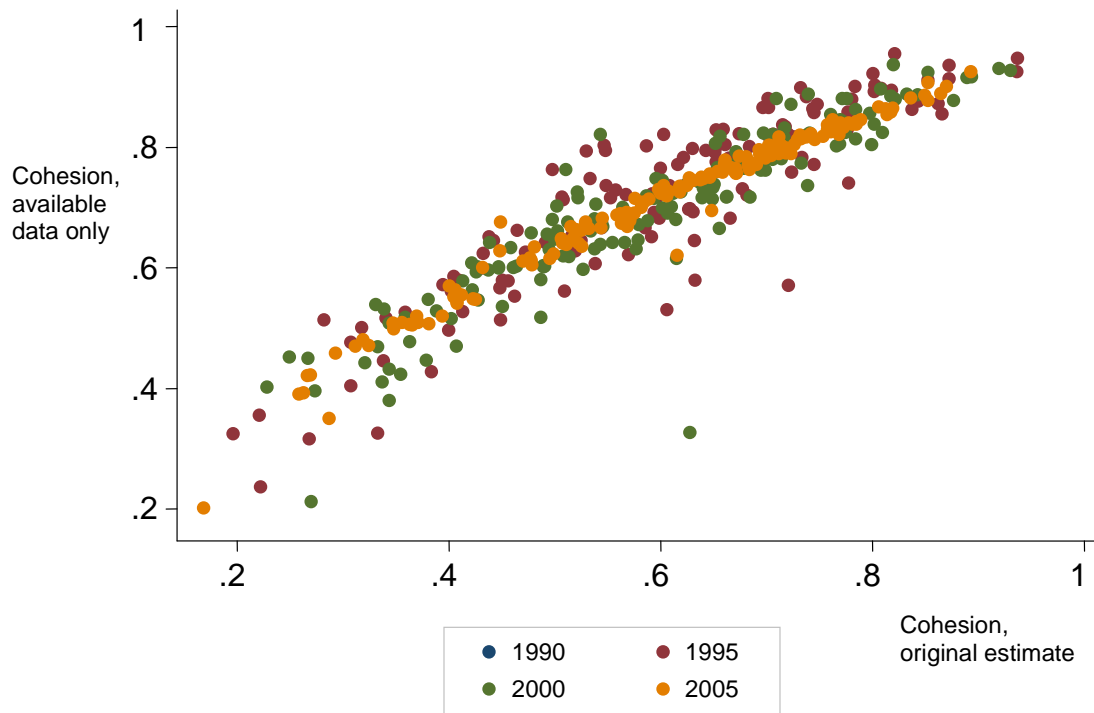
In the first set of sensitivity analysis tests below, we show the robustness of the indices to the inclusion of lagged data. In order to prevent volatility in index scores from one round to the next being produced by alterations in the selection of indicators, where data were not yet available for a country score on an indicator which had previously been available for that country, the country's most recent score was used instead, in addition to the new data from other sources. The tests below plot index scores for each of the four index year estimates (1990, 1995, 2000, 2005) both by the original estimation method, including lagged scores (x-axis) and without (y-axis), to check whether substantial biases are introduced.

7.4.3.1 A Comparison of Results With and Without 'Copying Back' Data

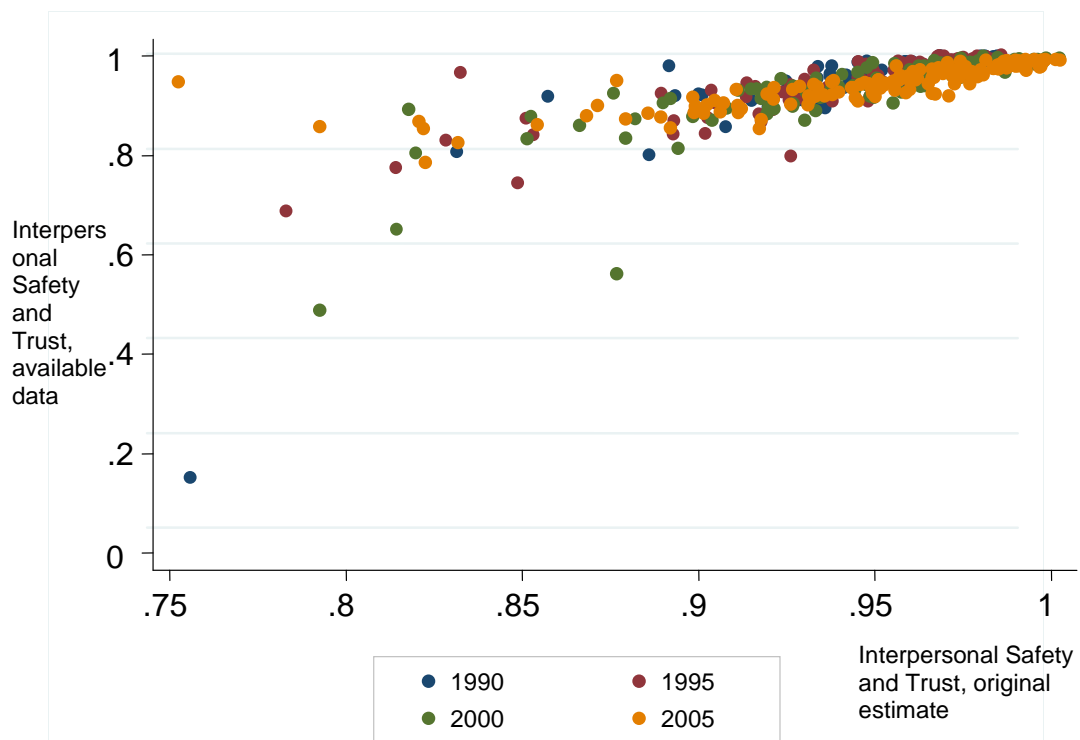
Civic Activism



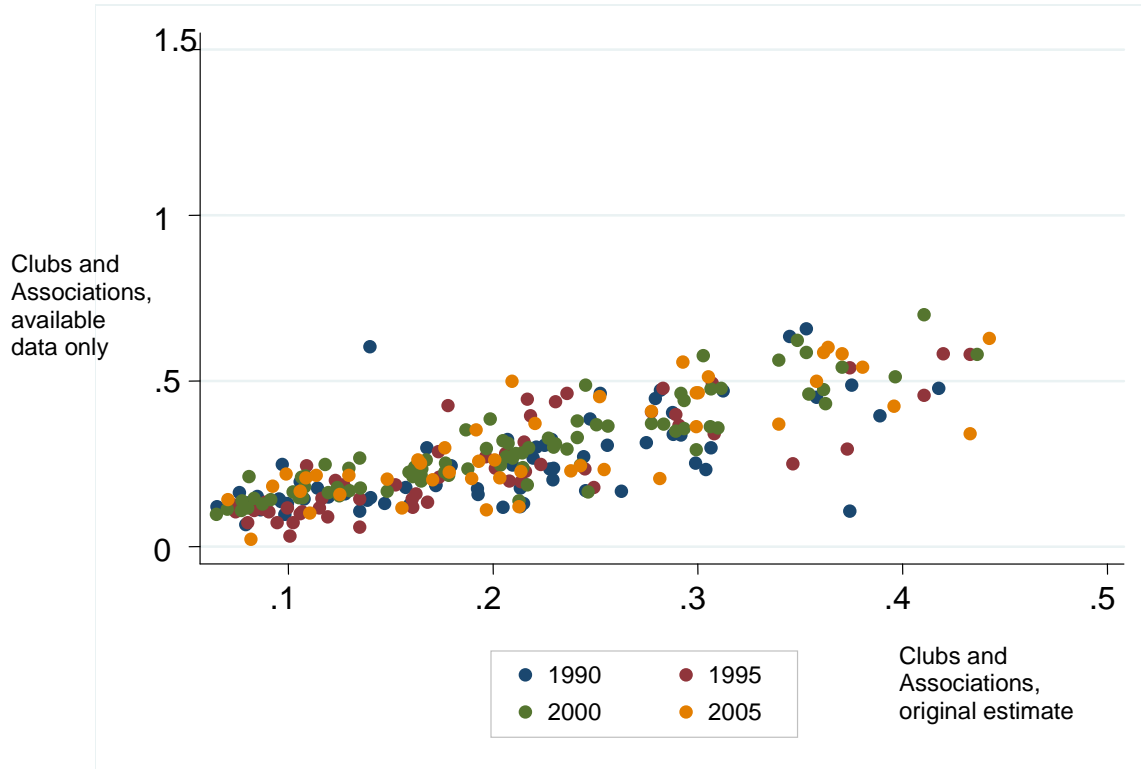
Intergroup Cohesion



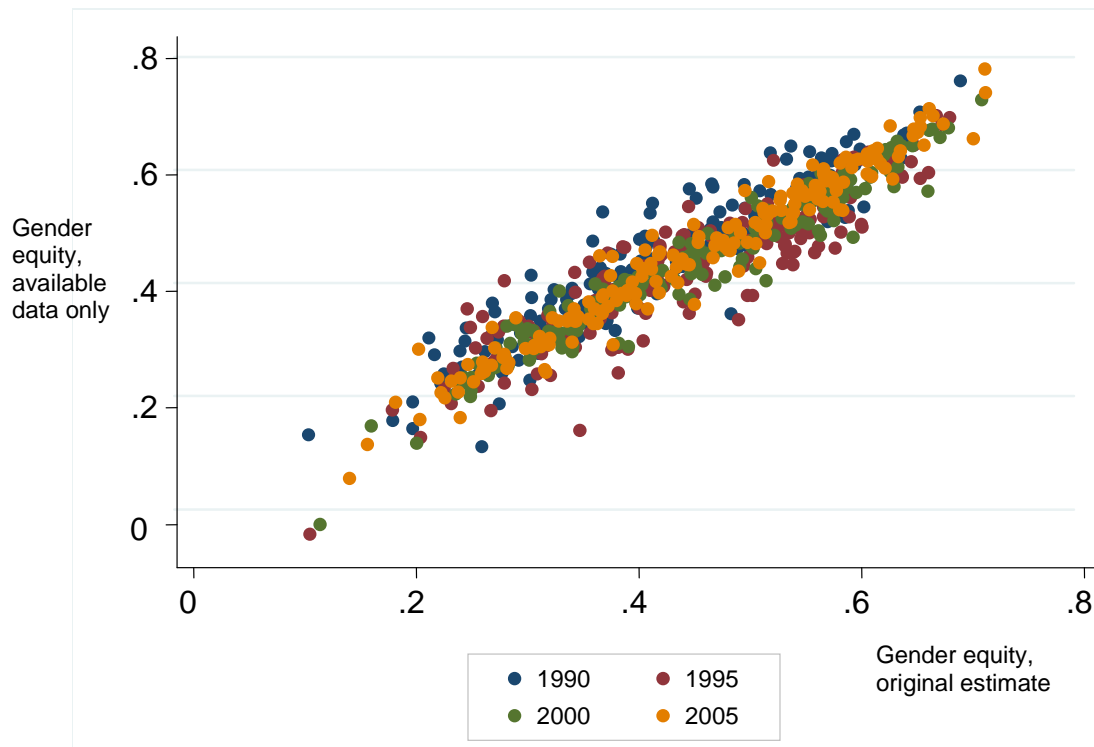
Interpersonal Safety and Trust



Clubs and Associations



Gender Equity



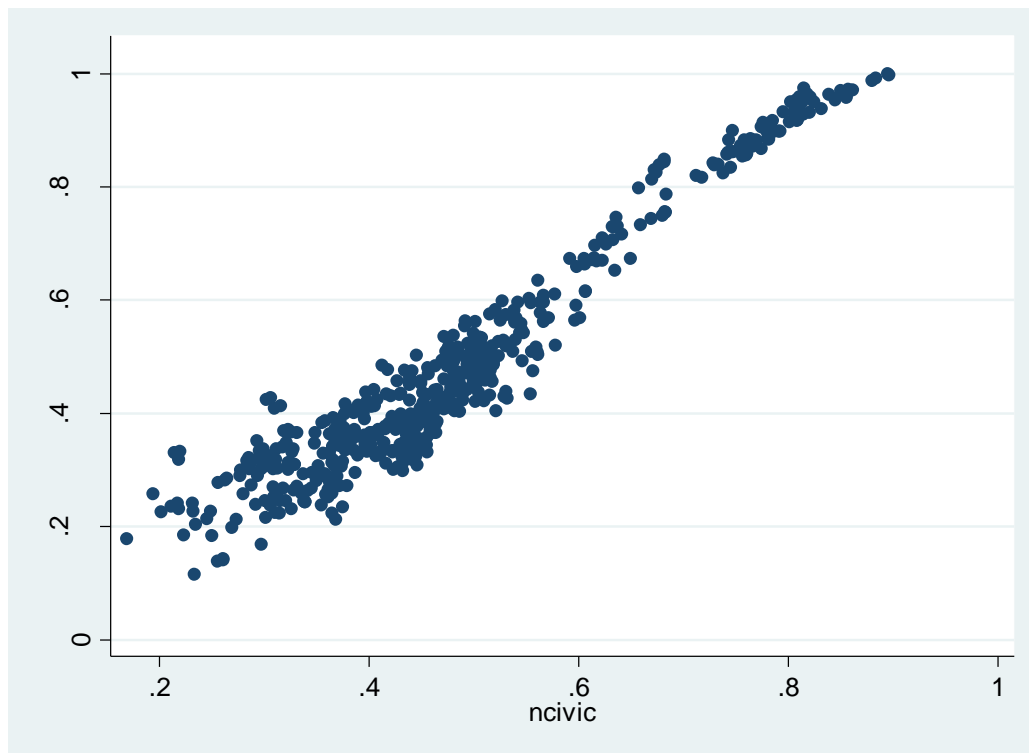
The results show that the two sets of estimates are highly correlated; introducing lagged data refines the country scores without substantially changing country positions. In addition, it is visible in most cases that the variance is widest among the early year estimates (1990 and 1995), for which relatively reduced levels of data are available, whereas by the most recent estimate (2005) correlation between the two estimates is very high. This is particularly the case with the intergroup cohesion and the gender equity indices. Only the clubs and associations measure shows greater volatility, reflecting the relative lack of underlying data for this measure.

7.4.3.2 Using Imputed Dataset to Generate the Indices

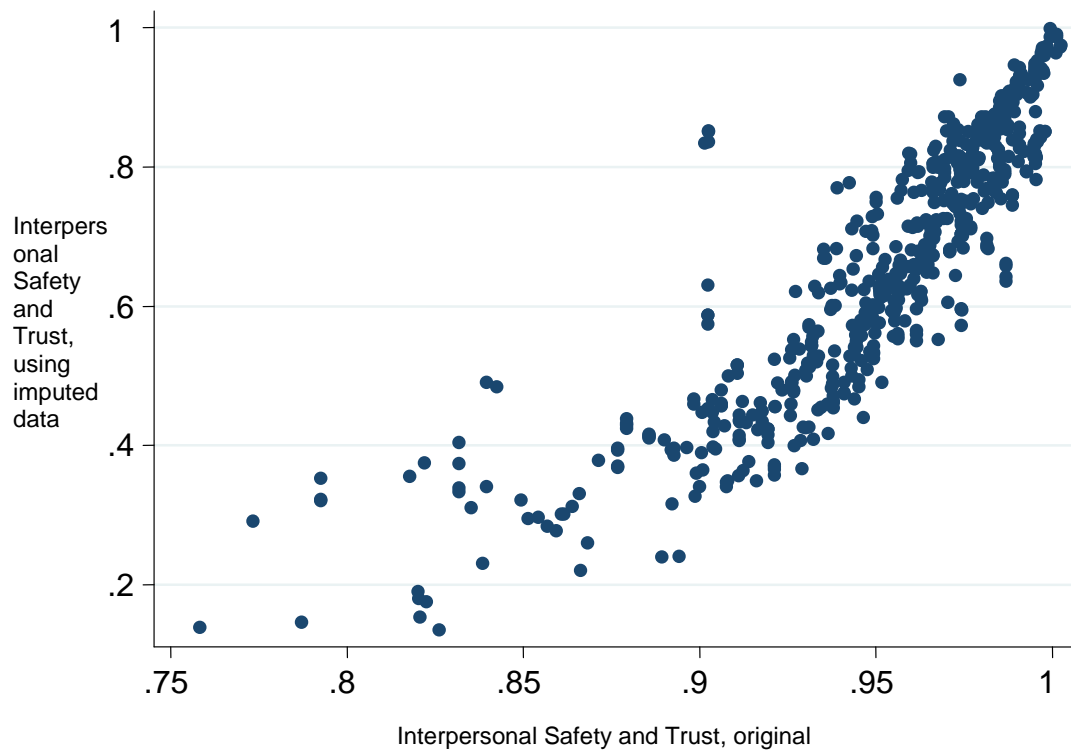
One of the key methodological advantages of the matching percentiles procedure is that it averts any imputation process, using only ‘real’ data for the countries that it scores. However, is this aggregation methodology as ‘reliable’ as aggregation after imputation? In this section, we provide a weak falsification test of the reliability of the matching percentiles method in this regard by comparing the results of country scores generated via the matching percentiles methodology, and the results that would be generated after imputing missing values. For each indicator used in the aggregation process, missing values were imputed using socioeconomic variables such as GDP per capita, the level of urbanization, literacy rates, as well as other indicators from within the social development indicators database. On condition that the imputation models are well-specified and the matching percentiles methodology is reliable, we should observe a close correspondence between these two sets of results. A weak or non-existent correspondence is indicative of either poor imputation results or a deeper flaw in the matching percentiles methodology itself.

In the following charts, the scores generated by the original matching percentiles methodology, using only actually recorded data, are shown on the x-axis, while scores generated after imputation of missing values are shown on the y-axis.

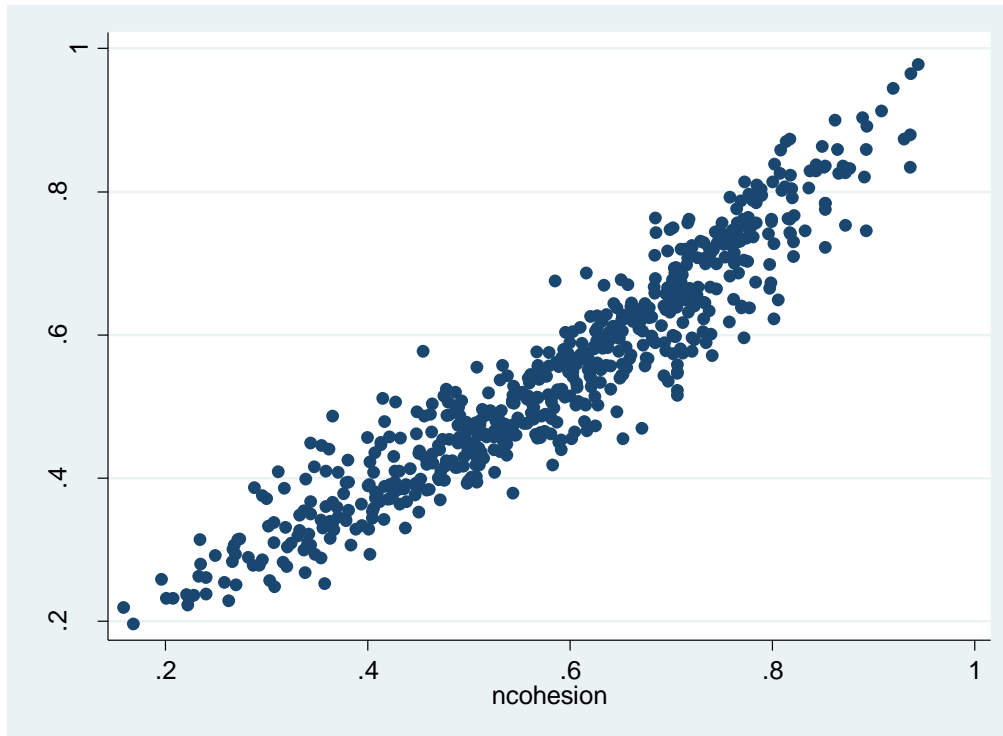
Civic Activism



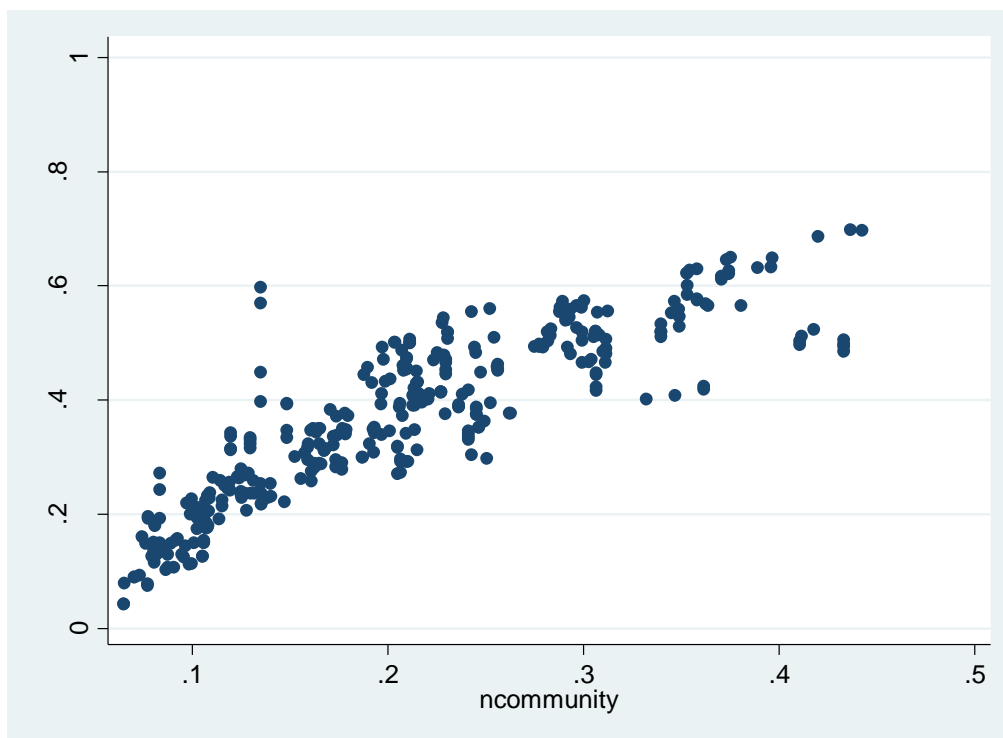
Interpersonal Safety and Trust



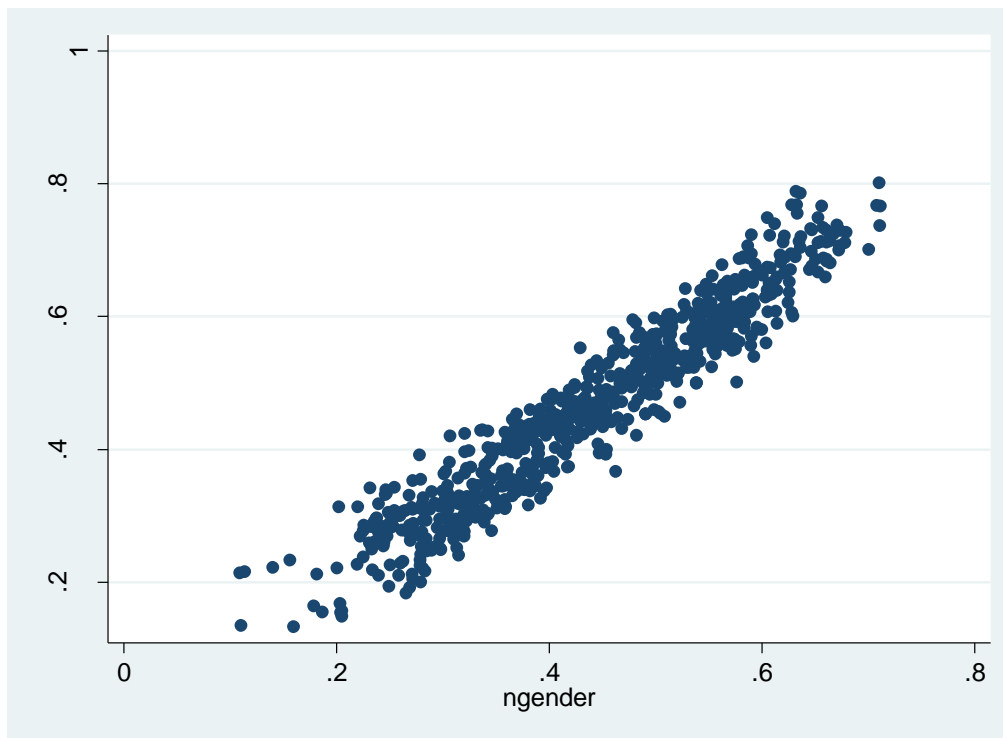
Intergroup Cohesion



Clubs and Associations



Gender Equity



Again, it can be seen that the scores generated following imputation of missing values are highly correlated with the original index scores. Note also that the scores generated following imputation are not necessarily more accurate than scores generated without imputation, and therefore even the small discrepancies between the two cannot be interpreted as deviations from the ‘real world’ index score. The outcome of this test therefore implies a satisfactory degree of accuracy for the original index estimates.

VIII. Conclusion

In recent years, there has been a dramatic expansion in the range of indicators available for the cross-country study of social institutions, such as the level of development of civil society, norms of intergroup cohesion, and discrimination against women and minorities. However, despite a wide range of individual studies that make use of individual indicators, there has been no comprehensive attempt to aggregate the new data into reliable and valid indices that would enable systematic cross-country testing. This paper has presented a

methodology for aggregating indicators, that of matching percentiles. The methodology, we contend, allows researchers to combine multiple indicators from different sources with different scales and different country samples, without the need to impute missing values or otherwise use data other than that which has been collected for individual countries.

The benefits of aggregating social indicators into composite indices along these lines are manifold. First, among the benefits of the matching percentiles method is the ability to estimate scores for a much larger number of countries than can be achieved by using a single indicator, with index coverage as high as 181 countries. Second, we have shown that the matching percentiles method provides a means of avoiding imputation, and therefore assigning scores to countries for indicators in which they lack data. Third, the matching percentiles method is relatively intuitive and easy to explain, in that it essentially assigns comparable scores across a series of indicators based on their rankings, before averaging them. And because matching percentiles uses only the ranking when assigning comparable scores, it dispenses with the linearity assumption, an assumption which may be at best unproven and at worst false.

We have provided evidence suggesting that the resultant indicators exhibit construct validity, understood as the extent to which measures accurately represent their concepts, and produce an observation distinct from that produced by other measures. Applying a range of diagnostic tests to the five composite indices created as part of the Indices of Social Development, we find substantial evidence confirming the clusters developed these five indices, as well as specific assessments of each of the 200 indicators assembled by the project.

Among the many diagnostic tests presented in this article, two in particular provide important support for the division of indicators into five indices of social development, and their aggregation using the matching percentiles method outlined here. The first is the confirmatory factor analysis, which shows that the indicators naturally group into the five categories assigned *a priori* to the data, namely civic activism, clubs and associations, interpersonal safety and trust, intergroup cohesion and gender equity. The *a priori* testing of indicator clustering suggests that these five dimensions do refer to stable and reliable

underlying aspects of social organization. The second are the sensitivity analyses outlined in the final part of this article, which show that the matching percentiles method is robust to a range of alterations in the aggregation rule, including restrictions on use of lagged data (using data from prior years to substitute missing observations), and using fully imputed data rather than skipping missing country-year observations.

Together, these diagnostic tests suggest that there are stable, underlying patterns in social institutions, and that social science concepts such as ‘civil society’, ‘social cohesion’, or ‘gender inclusion’ are not simply loose abstractions, but rather, refer to important latent features of social organization whose presence is indicated by the coherence of a range of associated indicators. They also suggest that we are able to summarize their level across different societies by aggregating an appropriate range of indicators, and provide a score that allows for comparison across countries and benchmarking of progress over time.

References

- Acemoglu, D., Johnson, S., and Robinson, J. (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation". *American Economic Review*, December 2001.
- Acemoglu, D. and Robinson, J. (2006). *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.
- Almond, G. and Verba, S. (1963). *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.
- Arrow, K. (1974). *The Limits of Organization*. Norton, New York.
- Bandura and Martin del Campo. (2006). 'Indices of National Performance: A Survey'. New York: UNDP.
- Barro, R. (1991). "Economic Growth in a Cross-Section of Countries". *Quarterly Journal of Economics*. 6, 407-444.
- Bebbington, A., Guugenheim, S., and Woolcock, M. (2006). "Concepts: Their Context and Their Consequences", In Anthony J. Bebbington, Michael Woolcock, Scott Guggenheim, and Elizabeth A. Olson, eds., *The Search for Empowerment: Social Capital as Idea and Practise in the World Bank*. Connecticut: Kumarian Press.
- Berkman, L. and Syme, L. (1979). "Social Networks, Host Resistance, and Mortality: a Nine-Year Follow-Up Study of Alameda County Residents." *American Journal of Epidemiology*. 109, no. 2: 186-204.
- Beugelsdijk, S., de Groot, H., and van Schaik, A. (2004). "Trust and economic growth: a robustness analysis", *Oxford Economic Papers*, 56, 118-134.
- Bockstette, V., Chanda, A. and Putterman, L.. (2002). "States and Markets: The Advantage of an Early Start," *Journal of Economic Growth* 7, 347-369.
- Carmines. E. G. and R. Z. Zeller. (1979). *Reliability and Validity Assessment*. Series: Quantitative Applications in the Social Sciences 07-017.
- Chu, Yun-han, Larry Diamond, Andrew J. Nathan, and Doh Chull Shin. (2008). "Comparative Perspectives on Democratic Legitimacy in East Asia," in Yun-han Chu, Larry Diamond, Andrew J. Nathan, and Doh Chull Shin, eds., *How East Asians View Democracy*. New York: Columbia University Press, 1-38.
- De, P. (2010). Governance, Institutions, and Regional Infrastructure in Asia. Asian Development Bank Institute Working Paper Series.

- Dempster, A., Rubin, D. (1983). "Incomplete data in sample surveys", in Madow, W.G., Olkin, I., Rubin, D. (Eds), *Sample Surveys Vol. II: Theory and Annotated Bibliography*, Academic Press, New York, NY, pp.3-10.
- Doing Business. (2005). *Doing Business Indicators*, World Bank: Washington DC.
- Donsbach and Traugott (2007). 'Editorial' in *International Journal of Public Opinion Research*, 19(1):3-4.
- Durkheim, E. [1897] 1951. *Suicide*. New York: Free Press.
- Esteve-Volart, B. 2004. "Gender discrimination and growth: Theory and evidence from India." DEDPS Working Paper No. 42. London: London School of Economics and Political Science.
- Etzioni, A. (1997). *The New Golden Rule: Community and Morality in a Democratic Society*. London: Profile Books.
- Fukuda-Parr and Kumar (eds.) 2004. *Readings in Human Development*. Oxford: Oxford University Press.
- Fukuyama, F. (1995). *Trust: Social Virtues and the Creation of Prosperity*. NY: Free Press.
- Guggenheim, S. (2006). "Crises and Contradictions: Understanding the Origins of a Community Development Project in Indonesia". In Bebbington, Woolcock, Guggenheim and Olson (eds.), *The Search for Empowerment*. CT: Kumarian Press.
- Hall, R. and Jones, C. (1999). "Why do Some Countries Produce so Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114, 83-116.
- Huntington, S. (1968). *Political Order in Changing Societies*. New Haven: Yale University Press.
- Kaufmann, D., Kraay, A., and Zoido-Lobaton, P. (1999a). "Governance Matters." World Bank Policy research Working Paper No. 2195, Washington, D.C.
- Kaufmann, D., Kraay, A., and Zoido-Lobaton, P. (1999b). "Aggregating Government Indicators." World Bank Policy research Working Paper No. 2195, Washington, D.C.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2007). "Governance Matters VI: Aggregate and Individual Governance Indicators 1996-2006", World Bank Policy Research Working Paper No. 4280, Washington, D.C.
- King, G. Honaker, J. Joseph, A. and K. Scheve. (2001). "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation," *American Political Science Review*, Vol. 95, No. 1.

- King, A. and E. Hill. (1993). "Women's Education in the Third World: An Overview," in E. M. King and M. A. Hill (eds.) *Women's Education in Developing Countries: Barriers, Benefits, and Policy*. Baltimore, MD: Johns Hopkins University Press.
- Knack, S. (2002). "Social Capital and the Quality of Government: Evidence from the States", *American Journal of Political Science* 46 (4): 772-785.
- Knack, S., and Keefer, P. (1997). "Does Social Capital have an Economic Payoff? A Cross-Country Investigation", *The Quarterly Journal of Economics* 112 (4):1251-1288.
- Inglehart, R. (1997). *Modernization and Postmodernization*. Princeton: Princeton University Press.
- Inglehart, R. and Baker, W. (2000). "Modernization, Cultural Change, and the Persistence of Traditional Values," *American Sociological Review* 65, 19-51.
- Inglehart, R. and Welzel, C. (2005). *Modernization, Cultural Change and Democracy*. Cambridge University Press.
- Lambsdorff, J. G. (2006). "The Methodology of the Corruptions Perceptions Index", Transparency International and the University of Passau, permanent url: http://www.icgg.org/downloads/CPI_2006_Methodology.pdf
- Levi, M. Karen S. Cook, Russell Hardin. (2007). *Cooperation Without Trust?* New York: Russell Sage Foundation.
- Narayan, D. and Cassidy, M. (2001). "A Dimensional Approach to Measuring Social Capital: Development and Validation of a Social Capital Inventory" in *Current Sociology* 49(2): 59-102.
- Narayan, D. and Pritchett, L. (1999). "Cents and Sociability: Household Income and Social Capital in Rural Tanzania", *Economic Development and Cultural Change* 47(4): 871-97.
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E. (2005) *Handbook on constructing composite indicators: methodology and users guide*, OECD-JRC joint publication, OECD Statistics Working Paper, STD/DOC(2005)3, JT00188147.
- Norris, P. (ed). (1999). *Critical Citizens: Global Support for Democratic Governance*. Oxford: Oxford University Press.
- North, D. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- North, D. (1991). "Institutions," *Journal of Economic Perspectives*, American Economic Association, vol. 5(1), pages 97-112.

- Paldman, M. (2000): "Social Capital: One or Many? Definition and Measurement", *Journal of Economics Surveys*, 14(5), pp. 29-653.
- Paxton, P. (2002). "Social Capital and Democracy: An Interdependent Relationship." *American Sociological Review*, 67:254-277.
- Putnam, R., Leonardi, R. and Nanetti, R. (1993) *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- Putnam, R. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.
- Putnam, R. (2007). "E Pluribus Unum: Diversity and Community in the Twenty-first Century", *Scandinavian Political Studies*, 30 (2), 137–174.
- Saltelli, A. 2007.' Composite indicators between analysis and advocacy' , *Social Indicators Research*, 81 , 65-77.Social Watch. 2007. *Social Watch Report*. Montevideo: Social Watch.
- Schultz, P. (2002). "Why governments should invest more to educate girls." *World Development*, 30(2): 207-225.
- Seeman, T. et al. (1994). Social ties and support and neuroendocrine function. MacArthur Studies of Successful Aging. *Ann Behavioral Medicine*, 16:95-106.
- Schuster T., Kessler R., and R. Aseltine. (1990). Supportive interactions, negative ineteractions and depressed mood. *Am J Community Psychology*, 18:423-438.
- Stone, W. (2001). *Measuring social capital: towards a theoretically informed measurement framework for researching social capital in family and community life* [Research Paper No. 24]. Australian Institute for Family Studies.
- Tilly, C. (2003). *The Politics of Collective Violence*. Cambridge: Cambridge University Press.
- Thomas, D., Strauss, J. and Henriques, M. 1991. "How Does Mother's Education Affect Child Height?" *Journal of Human Resources*, 26, no. 2:183-211.
- Tonnies, F. ([1887] 2001). *Community and Civil Society*. Cambridge: Cambridge University Press.
- Veenhoven, R. 2005. *World Database of Happiness: Continuous register of research on subjective enjoyment of life*. Available at the website of Erasmus University Rotterdam in the Netherlands: <http://worlddatabaseofhappiness.eur.nl>
- United Nations Development Program. 2007. *Human Development Report*. New York : United Nations.

- Urrea, F. J. (2007). Assessing Corruption - An analytical review of Corruption measurement and its problems: Perception, Error and Utility. Available on the United Nations Public Administration Network, www.unpan.org.
- Woolcock and Narayan. (2000). "Social Capital: Implications for Development Theory, Research, and Policy", *World Bank Research Observer*, Vol. 15, 2, pp. 225-249.
- World Bank. 2001. *Engendering development: Through Gender Equality in Rights, Resources, and Voice*. New York: Oxford University Press.

Appendix

Complete Indicator Summary, by Cluster

Civic Activism

Indicator	Source	Coverage
"How often do you get news from the following sources?" Percentage who have listened to radio news 'in the last day' or 'several times in the last week'	Afrobarometer	16
"How often do you get news from the following sources?" Percentage who have watched TV news 'in the last day' or 'several times in the last week'	Afrobarometer	16
"How often do you get news from the following sources?" Percentage who have read newspaper news 'in the last day' or 'several times in the last week'	Afrobarometer	16
Civics civil society rating — Structure	Civics	37
Civics civil society rating — Environment	Civics	37
Civics civil society rating — Values	Civics	37
Civics civil society rating — Impact	Civics	37
Radios per capita	ITU	197
Radios per household	ITU	197
"I am going to read out a political activity. I would like you to tell me, if you have ever done it, if you would ever do it, or if you would never do it. Taking part in authorized demonstrations. Percentage "have done" or "would do".	Latinobarometer	19
"I am going to read out a political activity. I would like you to tell me, if you have ever done it, if you would ever do it, or if you would never do it. Signing a petition. Percentage "have done" or "would do".	Latinobarometer	18
"How do you inform yourself about political affairs?" The radio. Percentage mentioned	Latinobarometer	18
"How do you inform yourself about political affairs?" The newspaper. Percentage mentioned	Latinobarometer	18
"How do you inform yourself about political affairs?" Television. Percentage mentioned	Latinobarometer	18
"How much attention did you pay to the political news on television?" Percentage saying a lot or quite a lot	Latinobarometer	18
How much attention did you pay to the political news in the newspaper? Percentage saying a lot or quite a lot	Latinobarometer	18
How much attention did you pay to the political news on the radio? Percentage saying a lot or quite a lot	Latinobarometer	18
"How many days during the last week did you watch the news on television?" Average number of days in country sample	Latinobarometer	19

"How many days during the last week did you read news in newspapers?" Average number of days in country sample	Latinobarometer	19
"How many days during the last week did you listen to news on the radio?" Average number of days in country sample	Latinobarometer	19
log total number of international organisation secretariats of international non-governmental organizations in given country, per log population	Global Civil Society Project	175
log extent to which organisations and individuals in each country are members of INGOs, number of INGOs with members in that country, per log population	Global Civil Society Project	176
Nonprofit sector workers as a percentage of the economically active population	SAIS	36
Newspapers per capita	UNESCO	107
Daily newspaper titles, per capita	UNESCO	
"I'm going to read out some different forms of political action that people can take, and I'd like you to tell me, for each one, whether you have actually done any of these things, whether you would do it, might do it, or would never, under any circumstances, do any of them". Signing a petition. Percentage "have done" or "might do".	World Values Surveys, Latinobarometer	89
"I'm going to read out some different forms of political action that people can take, and I'd like you to tell me, for each one, whether you have actually done any of these things, whether you would do it, might do it, or would never, under any circumstances, do any of them". Join a boycott. Percentage "have done" or "might do".	World Values Surveys	81
"I'm going to read out some different forms of political action that people can take, and I'd like you to tell me, for each one, whether you have actually done any of these things, whether you would do it, might do it, or would never, under any circumstances, do any of them". Attend a lawful demonstration. Percentage "have done" or "might do".	World Values Surveys, Afrobarometer, Latinobarometer	96
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". Daily newspaper, percentage mentioned	World Values Surveys	22
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". News broadcasts on radio or TV, percentage mentioned	World Values Surveys	22
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". Printed magazines, percentage mentioned	World Values Surveys	22
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". In depth reports on radio or TV, percentage mentioned	World Values Surveys	22
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". Books, percentage mentioned	World Values Surveys	22
"People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information". Internet, email, percentage mentioned	World Values Surveys	22

Intergroup Cohesion

Indicator	Source	Coverage
(Log) Number of reported incidents of violent riots, per capita	Databanks	189
(Log) Number of reported incidents of assassinations, per capita	Databanks	189
(Log) Number of reported incidents of terrorist acts, per capita	Databanks	189
(Log) Number of reported incidents of guerrilla activity, per capita	Databanks	121
Economist Intelligence Unit rating on likelihood of violent demonstrations	Economist Intelligence Unit	121
Economist Intelligence Unit rating on potential for terrorist acts	Economist Intelligence Unit	121
Fund for Peace rating on the "legacy of vengeance-seeking group grievance or group paranoia"	Fund for Peace	176
Fund for Peace rating on level of uneven economic development along group lines	Fund for Peace	176
Level of civil disorder, International Country Risk Guide rating	International Country Risk Guide	140
Level of internal conflict, International Country Risk Guide rating	International Country Risk Guide	140
Risk of terrorism, International Country Risk Guide rating	International Country Risk Guide	140
Level of ethnic tensions, International Country Risk Guide rating	International Country Risk Guide	140
Level of religious tensions, International Country Risk Guide rating	International Country Risk Guide	140
Level of ethnic minority rebellion in country, aggregated by group	Minorities at Risk	118
Level of economic and political discrimination against minorities in country, aggregated by group	Minorities at Risk	118
Economic and political disparities between minorities in country, aggregated by group	Minorities at Risk	118
"On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors." People of a different race or caste, percentage mentioned	World Values Survey	84
"On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors." People of a different religion, percentage mentioned	World Values Survey	50
"On this list are various groups of people. Could you please sort out any that you would not like to have as neighbors." People of a different language, percentage mentioned	World Values Survey	28
"I now want to ask you how much you trust various groups of people. Using the responses on this card, could you tell me how much you trust people of another religion?" Percentage who trust "not very much" or "not at all"	World Values Survey	22

"I now want to ask you how much you trust various groups of people. Using the responses on this card, could you tell me how much you trust people of another nationality?" Percentage who trust "not very much" or "not at all"	World Values Survey	21
On a scale of 1 to 10, where 1 is 'there is no discrimination' and 10 is 'there is a lot of discrimination', could you tell me if there is or is not discrimination against indigenous people in [this country] in the [workplace/courts/school system/political parties/police]?" Average level, among all respondents in country who identify as indigenous or mestizo	Latinobarometer	17
On a scale of 1 to 10, where 1 is 'there is no discrimination' and 10 is 'there is a lot of discrimination', could you tell me if there is or is not discrimination against indigenous people in [this country] in the [workplace/courts/school system/political parties/police]?" Average level, among all respondents in country who identify as black or mulatto	Latinobarometer	17
"As far as you know or have heard, which of the following groups is most discriminated against in this country - or are there no such groups?" Combined percentage citing: blacks, indigenous peoples, mulattos, mestizos, Asians, Arabs, Jews, immigrants, the disabled, those with AIDS.	Latinobarometer	18
Proportion of population reporting that their economic situation is the 'same' as other ethnic groups in country	Afrobarometer	16
Proportion of population reporting that their political situation is the 'same' as other ethnic groups in country	Afrobarometer	4
Proportion of population reporting that their ethnic group is 'never' treated unfairly in country	Afrobarometer	16

Clubs and Associations

Indicator	Source	Coverage
"What do you normally do in your free time?" Work in a voluntary organization, percentage mentioned	Latinbarometer	18
"How frequently do you do each of the following things?" Work for something affecting oneself or the community, percentage 'very frequently' or 'fairly frequently'	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Youth centers, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Mother's center/women's group, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Sports club, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Church organizations, percentage mentioned	Latinbarometer	18
"How frequently do you work for an issue that affects you or your community?" Percentage "very frequently" or "frequently"	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" [Labor] unions, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Voluntary associations, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Political party, percentage mentioned	Latinbarometer	18
"In which of the following organizations do you take part in, or don't you take part in any of them?" Cultural center, percentage mentioned	Latinbarometer	18
"Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an official leader, an active member, an inactive member, or not a member." Percentage who are leaders, active or inactive members of a "religious group".	Afrobarometer	16
"Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an official leader, an active member, an inactive member, or not a member." Percentage who are active or inactive member of a "development association".	Afrobarometer	16
Here is a list of actions that people sometimes take as citizens. For each of these, please tell me whether you, personally, have done any of these things during the past year. "Attended a community meeting". All respondents "Yes" or "Would do if had the chance"	Afrobarometer	16

"Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an active member, an inactive member, or not a member." Percentage who are active or inactive member of a "labor union"	Afrobarometer	16
"Now I am going to read out a list of groups that people join or attend. For each one, could you tell me whether you are an active member, an inactive member, or not a member." Percentage who are active or inactive member of a "business group"	Afrobarometer	16
Percentage respondents saying that people generally help one another in their neighborhood	International Crime Victim Survey	59
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Church or religious organization, percentage mentioned	World Values Survey	57
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Sports or recreational organization, percentage mentioned	World Values Survey	58
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Any other voluntary organizations, percentage mentioned	World Values Survey	56
"Which, if any, of the following do you belong to?" Youth work (e.g. Scouts, guides, youth clubs, etc), percentage mentioned	World Values Survey	58
"Which, if any, of the following do you belong to?" Sports or recreation, percentage mentioned	World Values Survey	58
"Do you currently do any unpaid voluntary work for any of these?" Voluntary organizations concerned with health, percentage mentioned	World Values Survey	54
"Which, if any, of the following do you belong to?" Conservation, environmental, or animal rights groups, percentage mentioned	World Values Survey	62
"Which, if any, of the following do you belong to?" Women's groups, percentage mentioned	World Values Survey	58
"Which, if any, of the following do you belong to?" Peace movement, percentage mentioned	World Values Survey	56
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Arts, music or educational organizations, percentage mentioned	World Values Survey	58
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Labor unions, percentage mentioned	World Values Survey	58
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Environmental organizations, percentage mentioned	World Values Survey	57

"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Professional organizations, percentage mentioned	World Values Survey	76
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Human rights organizations, percentage mentioned	World Values Survey	66
"Now I am going to read off a list of voluntary organizations; for each one, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?" Consumer groups, percentage mentioned	World Values Survey	22
I'm going to ask how of often you do various things. For each activity, would you say you do them every week or nearly every week; once or twice a month; only a few times a year; or not at all? Spent time socializing with friends, percentage "every week" or "once or twice a month"	World Values Survey	64
I'm going to ask how of often you do various things. For each activity, would you say you do them every week or nearly every week; once or twice a month; only a few times a year; or not at all? Spent time socializing with other members of arts or cultural association, percentage "every week" or "once or twice a month"	World Values Survey	58
Member of church or religious organization	International Social Survey	27
Member of neighborhood group	International Social Survey	27

Interpersonal Safety and Trust

Indicator	Source	Coverage
Felt unsafe in home, proportion saying 'never'	Afrobarometer	16
Had stuff stolen from home, proportion saying 'never'	Afrobarometer	16
Been attacked, proportion saying 'never'	Afrobarometer	16
Proportion of respondents who say that 'in general, most people can be trusted'	Asian Barometer	10
Proportion of respondents who say that most people try to be fair, rather than take advantage of you when given the chance	Asian Barometer	10
Economist Intelligence Unit rating on social distrust	Economist Intelligence Unit	121
Percentage respondents feel 'very safe' or 'fairly safe' walking alone in their area after dark	International Crime Victim Survey	64
Percentage respondents feel 'very safe' or 'fairly safe' while at home after dark	International Crime Victim Survey	37
Percentage respondents who avoid places when they go out	International Crime Victim Survey	56
Percentage respondents who take company with them when they go out	International Crime Victim Survey	67
Over the past five years have you or other members of your household had any of their cars/vans/trucks stolen? Vehicle owners only. Percentage "yes".	International Crime Victim Survey	67
Over the past five years have you or have members of your household been the victim of a theft of a car radio, or something else which was left in your car, or theft of a part of the car, such as a car mirror or wheel? Vehicle owners only. Percentage "yes".	International Crime Victim Survey	67
Apart from thefts, have parts of any of the cars/vans/trucks belonging to your household been deliberately damaged (vandalized) over the past five years? Vehicle owners only. Percentage "yes".	International Crime Victim Survey	67
"Over the past five years have you or other members of your household had any of their mopeds/scooters/motorcycles stolen?" Vehicle owners only. Percentage "yes".	International Crime Victim Survey	67
Over the past five years, did anyone actually get into your house or flat without permission and steal or try to steal something? I am not including here thefts from garages, sheds or lock-ups. Percentage "yes".	International Crime Victim Survey	66
Over the past five years, do you have any evidence that someone tried to get into your house or flat unsuccessfully. For example, damage to locks, doors or windows, or scratches around the lock? Percentage "yes".	International Crime Victim Survey	67
Over the past five years has anyone taken something from you, by using force, or threatening you? Or did anyone try to do so? Percentage "yes".	International Crime Victim Survey	66
Apart from theft involving force, there are many other types of theft of personal property, such as pickpocketing or the theft of a purse, wallet, clothing, jewellery, sports equipment	International Crime Victim Survey	67

at one's work, at school, in a pub, on public transport, on the beach or in the street. Over the past five years, have you personally been the victim of any of these thefts? Percentage "yes".

Apart from theft involving force, there are many other types of theft of personal property, such as pickpocketing or the theft of a purse, wallet, clothing, jewellery, sports equipment at one's work, at school, in a pub, on public transport, on the beach or in the street. Over the past five years, have you personally been the victim of any of these thefts? Percentage "yes".

International Crime
Victim Survey 67

Apart from the incidents just covered, have you over the past five years been personally attacked or threatened by someone in a way that really frightened you either at home or elsewhere, such as in a pub, in the street, at school, on public transport, on the beach, or at your workplace? Percentage "yes".

International Crime
Victim Survey 67

Interpol homicide rate

Interpol 124

Interpol rape rate

Interpol 54

Interpol rate of serious assault

Interpol 57

Interpol rate of aggravated theft

Interpol 52

Interpol rate of breaking and entering

Interpol 57

Interpol vehicle theft rate

Interpol 59

Interpol 'other theft' rate

Interpol 59

Interpol fraud rate

Interpol 59

Interpol consumer fraud rate

Interpol 56

WHO homicide rate

World Health
Organization 102

"Have you, or someone in your family, been assaulted, attacked, or been the victim of crime in the last 12 months?" Percentage mentioned

Latinobarometer 19

"How do you feel in the neighborhood in which you live?" Percentage of respondents who feel "secure"

Latinobarometer 17

"Have you been the victim of a crime?" Percentage replying "street robbery"

Latinobarometer 17

"Have you been the victim of a crime?" Percentage replying "burglary"

Latinobarometer 17

"Have you been the victim of a crime?" Percentage replying [attempted] "homicides or murders"

Latinobarometer 17

"Have you been the victim of a crime?" Percentage replying "kidnapping or disappearances"

Latinobarometer 17

State Department crime advisories, coded 1-5

US State Department 186

UNCJIN homicide rate

United Nations
Criminal Justice
Information Network 116

Percentage of managers surveyed for whom crime is a major business constraint	World Development Indicators	66
Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? Percentage replying that 'in general, most people can be trusted'	World Values Survey, Afrobarometer, Latinobarometer	86
"Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?" Proportion of respondents who say that most people would try to be fair	World Values Survey	40
"I now want to ask you how much you trust various groups of people. Using the responses on this card, could you tell me how much you trust your neighborhood?" Percentage who trust "not very much" or "not at all"	World Values Survey	22
"I now want to ask you how much you trust various groups of people. Using the responses on this card, could you tell me how much you trust people you know personally?" Percentage who trust "not very much" or "not at all"	World Values Survey	22
"I now want to ask you how much you trust various groups of people. Using the responses on this card, could you tell me how much you trust people you meet for the first time?" Percentage who trust "not very much" or "not at all"	World Values Survey	22

Gender Equity

Indicator	Source	Coverage
Percentage of men believing that a 'married man has a right to beat his wife and children'	Afrobarometer	15
Percentage of respondents who tend to agree or strongly agree that 'women have always been subject to traditional laws and customs, and should remain so'.	Afrobarometer	15
Percentage of respondents who tend to agree or strongly agree that 'women should have the same chance of being elected to political office as men'.	Afrobarometer	4
Rating on level of women's economic rights	CIRI	190
Rating on level of women's social rights	CIRI	190
Ratio of average female to male wages, across all available labor categories	International Labor Organization	65
Percentage of women who agree that women have the same chance as men to get a good job in their country	Latinobarometer	18
Percentage of women who agree that women have the chance to earn the same salary as men in their country	Latinobarometer	19
Percentage of women who agree that women have the same chance as men to get a good education in their country	Latinobarometer	18
"Do you agree or disagree with the following statements? When jobs are scarce, men have more right to a job than women." Proportion of employers and managers who agree or strongly agree	World Values Surveys	83
"People talk about the changing roles of men and women today. For each of the following statements I read out, can you tell me how much you agree with each. Do you agree strongly, agree, disagree, or disagree strongly? On the whole, men make better political leaders than women do". Proportion of those of voting age who agree or strongly agree	World Values Surveys	74
"People talk about the changing roles of men and women today. For each of the following statements I read out, can you tell me how much you agree with each. Do you agree strongly, agree, disagree, or disagree strongly? A university education is more important for a boy than a girl". Proportion of parents who agree or strongly agree	World Values Surveys	75
"For each of the following statements I read out, can you tell me how much you agree with each. Do you agree strongly, agree, disagree, or disagree strongly? On the whole, men make better business executives than women do". Proportion of employers and managers who agree or strongly agree	World Values Surveys	23
Percentage of labor force that is female	World Development Indicators	186
Ratio of females among legislators, senior officials and managers	International Labor	94

	Organization	
Ratio of females in professional jobs	International Labor Organization	136
Ratio between female and male primary school enrollment	World Development Indicators	175
Ratio between female and male secondary school enrollment	World Development Indicators	150
Ratio between female and male tertiary educational enrollment	World Development Indicators	173
Ratio between adult female and male literacy rates	World Development Indicators	146
Ratio between adult female and adult male mortality rates	World Development Indicators	195
