Branisa, Boris; Ziegler, Maria; Klasen, Stephan

Conference Paper

The Institutional Basis of Gender Inequality

Proceedings of the German Development Economics Conference, Hannover 2010, No. 16

Provided in cooperation with:
Verein für Socialpolitik

Suggested citation: Branisa, Boris; Ziegler, Maria; Klasen, Stephan (2010) : The Institutional Basis of Gender Inequality, Proceedings of the German Development Economics Conference, Hannover 2010, No. 16, http://hdl.handle.net/10419/39975

The Institutional Basis of Gender Inequality: The Social Institutions and Gender Index (SIGI)*

Boris Branisa† **Maria Ziegler**† **Stephan Klasen**†

This version: November 17, 2009

Abstract. In this paper we construct the Social Institutions and Gender Index (SIGI) and its five subindices Family code, Civil liberties, Physical integrity, Son Preference and Ownership rights using variables of the OECD Gender, Institutions and Development database. Instead of measuring gender inequality in education, health, economic or political participation, these indices allow a new perspective on gender issues in developing countries. The SIGI and the subindices measure long-lasting social institutions which are mirrored by societal practices and legal norms that frame gender-relevant meanings and form the basis of gender roles. The subindices measure each one dimension of the concept and the SIGI combines the subindices into a multidimensional index of deprivation of women caused by social institutions. Methodologically, the SIGI is inspired by the Foster-Greer-Thorbecke poverty measures. It offers a new way of aggregating gender inequality in several dimensions, penalizing high inequality in each dimension and allowing only for partial compensation between dimensions. The SIGI and the subindices are useful tools to identify countries and dimensions of social institutions that deserve attention. Empirical results confirm that the SIGI provides additional information to that of other well-known gender-related indices.

Keywords: SIGI, Composite index, Gender inequality, Social institutions, OECD-GID database.

***JEL* codes:** D63, I39, J16

* We thank Walter Zucchini, Oleg Nenadić, Carola Grün and Axel Dreher from the University of Goettingen, as well as members of the International Working Group on Gender, Macroeconomics and International Economics (GEM-IWG), participants at the 2009 Far East and South Asia Meeting of the Econometric Society and at the 2009 Singapore Economic Review Conference for valuable comments and discussion. The usual disclaimer applies.

†University of Goettingen, Department of Economics, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany.

1 Introduction

Gender inequality is a major problem for development. First, the affected women are deprived of their basic freedoms (Sen, 1999). Second, going beyond this intrinsic feature of gender inequality, it implies high costs for society in the form of lower human capital, worse governance, and lower growth (e.g. World Bank, 2001; Klasen, 2002). Although the intrinsic and instrumental value of gender equality is known and set as a goal on the development agenda (e.g., Millennium Development Goal 3 "Promote gender equality and empower women"), gender inequality remains a pervasive phenomenon.

To measure the extent of this problem at the cross-country level several gender-related indices have been proposed, e.g. the Gender-Related Development Index (GDI) and the Gender Empowerment Measure (GEM) (United Nations Development Programme, 1995), the Global Gender Gap Index from the World Economic Forum (Lopez-Claros and Zahidi, 2005), the Gender Equity Index developed by Social Watch (2005) or the African Gender Status Index proposed by the Economic Commission for Africa (2004). These measures focus on gender inequality in well-being or in agency and they are typically outcome-focused (Klasen, 2006, 2007).

Focusing only on outcomes neglects the question of where gender inequality comes from. Gender inequality is mainly the result of human behavior. How people behave and interact is influenced by institutions. From an economics perspective, institutions are conceived as the result of collective choices in a society to achieve efficiency, solve collective action dilemmas and reduce transaction costs (e.g. North, 1990). Other social sciences emphasize legitimacy and appropriateness instead of efficiency. Institutions influence the preferences of actors and provide role models that are internalized by them (Hall and Taylor, 1996; De Soysa and Jütting, 2007).

There is a particular type of institutions that is relevant for gender inequality, *social institutions related to gender inequality*. Social institutions related to gender inequality are long-lasting norms, values and codes of conduct that find expression in traditions, customs and cultural practices, informal and formal laws. They influence human behavior as they frame gender-relevant meanings, form the basis of gender roles and become guiding principles in everyday life. Influencing the distribution of power between men and women in the private sphere of the family, in the economic sphere, and in public life, they constrain the opportunities of men and women and their capabilities to live the life they value (Sen, 1999). Accounting for these social institutions is necessary to understand outcome gender inequality and the deprivation women experience. Additionally, neglecting them implies neglecting a major factor that might be related to development.

There are three measures that from a human rights perspective deal with the question of how women are treated in society: the Women's Political Rights index (WOPOL), the Women's Economic Rights index (WECON), and the Women's Social Rights index (WOSOC) of the CIRI Human Rights Data Project.[1] These indices measure on a yearly basis whether a number of internationally recognized rights for women are included in law and whether government enforces them. They proxy somehow the type of institutions we are concerned about, but also cover outcomes of these institutions. From the three indices, WOSOC is the most encompassing measure covering social relations (Bjornskov, Dreher, and Fischer, 2009). However, it does not allow to differentiate between different dimensions of social institutions. For example, it is important to distinguish between what happens within the family and what happens in public and social life. Furthermore, all three indices can only take four values from 0 (no rights) to 3 (legally guaranteed and enforced rights) which makes it difficult to compare and rank countries as there are many ties in the data.

This paper centers on the measurement of social institutions related to gender inequality. We propose new composite measures that proxy social institutions related to gender inequality in non-OECD countries based on variables of the OECD Gender, Institutions and Development database (Morrison and Jütting, 2005; Jütting, Morrison, Dayton-Johnson, and Drechsler, 2008). We aggregate the variables into five subindices that measure each one dimension of social institutions related to gender inequality (Family code, Civil liberties, Physical integrity, Son preference and Ownership rights). We combine the subindices into the Social Institutions and Gender Index (SIGI) as a multidimensional measure of deprivation of women.

In general, the construction of composite measures requires several decisions, for example about the weighting scheme and the method of aggregation (e.g. Nardo, Saisana, Saltelli, Tarantola, Hoffman, and Giovannini, 2005). The subindices as one-dimensional measures are built using the method of polychoric PCA to extract the common information of the variables corresponding to a subindex. When we combine the subindices to construct the SIGI, we use a reasonable methodology to capture the multidimensional deprivation of women caused by social institutions. The formula of the SIGI is inspired by the Foster-Greer-Thorbecke poverty measures (Foster, Greer, and Thorbecke, 1984) and offers a new way of aggregating gender inequality in several dimensions measured by the subindices. It is transparent and easy to understand, it penalizes high inequality in each dimension and allows only for partial compensation between dimensions.

The SIGI and the subindices are useful tools to compare the societal situation of women

[1] Information is available on the webpage of the project http://ciri.binghamton.edu/.

in over 100 non-OECD countries from a new perspective, allowing the identification of problematic countries and dimensions of social institutions that deserve attention by policy makers and need to be scrutinized in detail. Empirical results show that the SIGI provides additional information to that of other well-known gender-related indices. Moreover, regression analysis shows that the SIGI is related to indices that measure outcome gender inequality, even if one controls for region, religion and level of economic development.

This paper is organized as follows. In section 2, we describe the OECD Gender, Institutions and Development Database. Then, in sections 3 and 4 we focus on the construction of the subindices and of the SIGI. In section 5, we present empirical results by country, interesting regional patterns and a comparison between the SIGI and other gender-related measures. Furthermore, using regression analysis we illustrate the relevance of the SIGI for explaining outcome gender inequality. The last section concludes with a discussion of the strengths and weaknesses of the proposed measures.

2 The OECD Gender, Institutions and Development Database

As input for the composite measures we use variables from the OECD Gender, Institutions and Development Database (Morrison and Jütting, 2005; Jütting et al., 2008). This is a cross-country database covering about 120 countries with more than 20 variables measuring social institutions related to gender inequality.[2] These variables proxy social institutions through prevalence rates, legal indicators or indicators of social practices. We assume that the concept social institutions related to gender inequality is multidimensional. Following previous work done by the OECD (Jütting et al., 2008) we choose twelve variables that are assumed to measure each one of four dimensions of social institutions.

The *Family code* dimension refers to the private sphere with institutions that influence the decision-making power of women in the household. Family code is measured by the following four variables. *Parental authority* measures whether women have the right to be the legal guardian of a child during marriage, and whether women have custody rights over a child after divorce. *Inheritance* is based on formal inheritance rights of spouses. *Early marriage* measures the percentage of girls between 15 and 19 years of age who

[2] The data are available at the web-pages http://www.wikigender.org and http://www.oecd.org/dev/gender/gid.

are/were ever married. *Polygamy* measures the acceptance of polygamy in the population. Countries where this information is not available are assigned scores based on the legality of polygamy.[3]

The public sphere is measured by the *Civil liberties* dimension that captures the freedom of social participation of women and includes the following two variables. *Freedom of movement* indicates the freedom of women to move outside the home. *Freedom of dress* is based on the obligation of women to use a veil or burqa to cover parts of their body in public.

The *Physical integrity* dimension comprises different indicators on violence against women. The variable *violence against women* indicates the existence of laws against domestic violence, sexual assault or rape, and sexual harassment. *Female genital mutilation* is the percentage of women who have undergone female genital mutilation. *Missing women* measures gender bias in mortality. Countries were coded based on estimates of gender bias in mortality for a sample of countries (Klasen and Wink, 2003) and on sex ratios of young people and adults.

The *Ownership rights* dimension covers the economic sphere of social institutions proxied by the access of women to several types of property. *Women's access to land* indicates whether women are allowed to own land. *Women's access to bank loans* measures whether women are allowed to access credits. *Women's access to property other than land* covers mainly access to real property such as houses, but also any other property.

Concerning the *missing women* variable in the *Physical integrity* dimension, it could be argued that it reflects another dimension of gender inequality. Missing women is an extreme manifestation of son preference under scarce resources. 100 million women are not alive who should be alive if women were not discriminated against (Sen, 1992; Klasen and Wink, 2003). The other components of *Physical integrity*, *violence against women* and *female genital mutilation*, measure particularly the treatment of women which is not only motivated by economic considerations. In the next section, we check with statistical methods if *missing women* measures another dimension as the variables *violence against women* and *female genital mutilation*.

These twelve variables are between 0 and 1. The value 0 means no or very low inequality and the value 1 indicates high inequality. Three of the variables (*early marriage, female genital mutilation and violence against women*) are continuous. The other indi-

[3] Acceptance of polygamy in the population might proxy actual practices better than the formal indicator legality of polygamy and, moreover, laws might be changed faster than practices. Therefore, the acceptance variable is the first choice for the subindex Family code. The reason for using legality when acceptance is missing is to increase the number of countries.

cators measure social institutions on an ordinal categorical scale. The chosen variables cover around 120 non-OECD countries from all regions in the world except North America.[4] The choice of the variables is also guided by the availability of information so that as many countries as possible can be ranked by the SIGI. Within our sample 102 countries have information for all twelve variables.

3 Construction of the Subindices

The objective of the subindices is to provide a summary measure for each dimension of social institutions related to gender inequality. In every subindex we want to combine variables that are assumed to belong to one dimension. The first step is to check the statistical association between the variables. The second step consists in aggregating the variables with a reasonable weighting scheme.

3.1 Measuring the Association between Categorical Variables

To check the association between variables, and as most of them are ordinal, we use Kendall Tau b and Multiple Joint Correspondence Analysis (Greenacre, 2007; Nenadić, 2007).

Kendall Tau b is a rank correlation coefficient. These measures are useful when the data are ordinal and thus the conditions for using Pearson's correlation coefficient are not fulfilled. For each variable, the values are ordered and ranked. Then the correspondence between the rankings is measured.[5] Taking into account tied pairs, the formula for Kendall Tau b is

$$\tau_b = \frac{C - D}{\sqrt{\frac{n(n-1)}{2-T_x}\frac{n(n-1)}{2-T_y}}}, \tag{1}$$

where C is the number of concordant pairs, D is the number of discordant pairs, n is the number of observations, $\frac{n(n-1)}{2}$ is the number of all pairs, T_x is the number of pairs tied on

[4] The OECD Gender, Institutions and Development Database does not contain variables that capture relevant social institutions related to gender inequality in OECD countries.

[5] For calculating Kendall Tau, one counts the number of concordant and discordant pairs of two rankings, builds the difference and divides this difference by the total number of pairs. A value of 1 means total correspondence of rankings, i.e. the rankings are the same. A value of -1 indicates reverse rankings or a negative association between rankings. A value of 0 means independence of rankings. Kendall Tau b is a variant of Kendall tau that corrects for ties, which are frequent in the case of discrete data (Agresti, 1984, chap. 9). We consider Kendall Tau b to be the appropriate measure of rank correlation to find out whether our data are related.

the variable x and T_y is the number of pairs tied on the variable y. The notation is taken from Agresti (1984).

As a second method to check the association between variables we examine the graphics produced by Multiple Joint Correspondence Analysis (MJCA) (Greenacre, 2007; Nenadić, 2007), after having discretized the three continuous variables. Correspondence Analysis is a method for analyzing and representing the structure of contingency tables graphically. We use MJCA to find out whether variables seem to measure the same.[6]

The results for Kendall tau b (Tables 1- 5) are reported in Appendix 1. A significant positive value of Kendall tau b is a sign for a positive association between two variables. This is the case for all variables belonging to one dimension, except *missing women* in the subindex *Physical integrity*. The graphs produced with MJCA are available upon request.[7] The results of MJCA also confirm that within every dimension all the variables seem to measure the same dimension, with the exception of *missing women* in the dimension *Physical integrity*. These results support the argumentation in section 2.

We decide to use the variable *missing women* as a fifth subindex called *Son preference*. The artificially higher female mortality is one of the most important and cruel aspects of gender inequality and should not be neglected, as over 100 million women that should be alive are missing (Sen, 1992; Klasen and Wink, 2003). Missing women is the "starkest manifestation of the lack of gender equality" (Duflo, 2005).

3.2 Aggregating Variables to Build a Subindex

The five subindices *Family code*, *Civil liberties*, *Son preference*, *Physical integrity* and *Ownership rights* use the twelve variables as input that were mentioned in the previous section. Each subindex combines variables that measure one dimension of social institutions related to gender inequality. In the case of Son preference, the subindex takes the

[6] Correspondence Analysis is an exploratory and descriptive method to analyze contingency tables. Instead of calculating a correlation coefficient to capture the association of variables, the correspondence of conditional and marginal distributions of either rows or columns - also called row or column profiles - is measured using a χ^2-statistic, that captures the distance between them. These row or column profiles then are plotted in a low-dimensional space, so that the distances between the points reflect the dissimilarities between the profiles. Multiple Joint Correspondence Analysis is an extended procedure for the analysis of more than two variables and considers the cross-tabulations of the variables against each other in a so-called Burt matrix but with modified diagonal sub-tables. This facilitates to figure out whether variables are associated. This is the case when they have similar deviations from homogeneity, and therefore get a similar position in a profile space (Greenacre, 2007; Nenadić, 2007).

[7] The graphs produced with MJCA can be interpreted in the following way. In most cases, one of the axes represents whether there is inequality and the other axe represents the extent of inequality. If one connects the values of a variable one obtains a graphical pattern. If this is similar to the pattern obtained for another variable, then both variables are associated.

value of the variable missing women. In all other cases, the computation of the subindex values involves two steps.

In the first step, the method of polychoric principal component analysis is used to extract the common information of the variables corresponding to a subindex. Principal component analysis (PCA) is a method of dimensionality reduction that is valid for normally distributed variables (Jolliffe, 1986). This assumption is violated in this case, as the data include variables that are ordinal, and hence the Pearson correlation coefficient is not appropriate. Following Kolenikov and Angeles (2004, 2009) we use polychoric PCA, which relies on polychoric and polyserial correlations. These are estimated with maximum likelihood, assuming that there are latent normally distributed variables that underly the ordinal categorical data. We use the First Principal Component (FPC) as a proxy for the common information contained by the variables corresponding to the subindices, measuring each one of the dimensions of social institutions related to gender inequality. The first principal component is the weighted sum of the standardized original variables that captures as much of the variance in the data as possible.[8] The standardization of the original variables is done as follows. In the case of continuous variables, one subtracts the mean and then divides by the standard deviation. In the case of ordinal categorical variables, the standardization uses results of an ordered probit model. The weight that each variable gets in these linear combinations is obtained by analyzing the correlation structure in the data. The weights are shown in Table 6.

In the second step, the subindex value is obtained rescaling the FPC so that it ranges from 0 to 1 to ease interpretation. A country with the best possible performance (no inequality) is assigned the value 0 and a country with the worst possible performance (highest inequality) the value 1. Hence, the subindex values of all countries are between 0 and 1. Using the score of the FPC the subindex is calculated using the following transformation. Country *X* corresponds to a country of interest, Country *Worst* corresponds to a country with worst possible performance and Country *Best* is a country with best possible performance.

$$\begin{aligned} \text{Subindex(Country X)} &= \frac{\text{FPC(Country X)}}{\text{FPC(Country Worst)} - \text{FPC(Country Best)}} \\ &- \frac{\text{FPC(Country Best)}}{\text{FPC(Country Worst)} - \text{FPC(Country Best)}} \end{aligned} \tag{2}$$

[8] The proportion of explained variance by the first principal component is 70% for *Family code*, 93% for *Civil liberties*, 60% for *Physical integrity* and 87% for *Ownership rights*.

Every subindex is intended to measure a different dimension of social institutions related to gender inequality. To check whether the subindices are empirically non-redundant, so that they provide each additional information, we conduct an empirical analysis of the statistical association between them. In the case of well-being measures, McGillivray and White (1993) suggest using two explicit thresholds to separate redundancy from non-redundancy, that is a correlation coefficient of 0.90 and 0.70. Based on this suggestion we use the threshold 0.80. In Table 7 we present Kendall tau b as a measure of the statistical association between the five subindices. In all cases, the subindices are positively correlated, showing that they all measure social institutions related to gender inequality. It must be noted, however, that the correlation is not always statistically significant. Kendall tau b is lower than 0.80 in all cases, which means that each subindex measures a distinct aspect of social institutions related to gender inequality.

4 The Social Institutions and Gender Index (SIGI)

With the subindices described in the last section as input, we build a multidimensional composite index named Social Institutions and Gender Index (SIGI) which reflects the deprivation of women caused by social institutions related to gender inequality. The proposed index is transparent and easy to understand. As in the case of the variables and of the subindices, the index value 0 corresponds to no inequality and the value 1 to complete inequality.

The SIGI is an unweighted average of a non-linear function of the subindices. We use equal weights for the subindices, as we see no reason for valuing one of the dimensions more or less than the others.[9] The non-linear function arises because we assume that inequality in gender-related social institutions leads to deprivation experienced by the affected women, and that deprivation increases more than proportionally when inequality increases. Thus, high inequality is penalized in every dimension. The non-linearity also means that the SIGI does not allow for total compensation among subindices, but permits partial compensation. Partial compensation implies that high inequality in one dimension, i.e. subindex, can only be partially compensated with low inequality on another dimension.[10]

[9] Empirically, even in the case of equal weights the ranking produced by a composite index is influenced by the different variances of its components. The component that has the highest variance has the largest influence on the composite index. In the case of the SIGI the variances of the five components are reasonably close to each other, *Ownership rights* having the largest and *Physical integrity* having the lowest variance.

[10] Other approaches have been also proposed in the literature, e.g. the non-compensatory approach by

For our specific five subindices, the value of the index the SIGI is then calculated as follows.

$$\begin{aligned} \text{SIGI} &= \frac{1}{5}\,(\text{Subindex Family Code})^2 + \frac{1}{5}\,(\text{Subindex Civil Liberties})^2 \\ &+ \frac{1}{5}\,(\text{Subindex Physical Integrity})^2 + \frac{1}{5}\,(\text{Subindex Son preference})^2 \\ &+ \frac{1}{5}\,(\text{Subindex Ownership Rights})^2 \end{aligned} \tag{3}$$

Using a more general notation, the formula for the SIGI $I(X)$, where X is the vector containing the values of the subindices x_i with $i = 1,\ldots,n$, is derived from the following considerations. For any subindex x_i, we interpret the value 0 as the goal of no inequality to be achieved in every dimension. We define a deprivation function $\phi(x_i,0)$, with $\phi(x_i,0) > 0$ if $x_i > 0$ and $\phi(x_i,0) = 0$ if $x_i = 0$ (e.g. Subramanian, 2007). Higher values of x_i should lead to a penalization in $I(X)$ that should increase with the distance x_i to zero. In our case the deprivation function is the square of the distance to 0 so that deprivation increases more than proportionally as inequality increases.

$$SIGI = I(X) = \frac{1}{n}\sum_{i=1}^{n}\phi(x_i,0) = \frac{1}{n}\sum_{i=1}^{n}(x_i-0)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i)^2. \tag{4}$$

The formula is inspired by the Foster-Greer-Thorbecke (FGT) poverty measures (Foster et al., 1984). The general FGT formula is defined for $y_i \leq z$ as:

$$FGT(Y,\alpha,z) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{z-y_i}{z}\right)^{\alpha}, \tag{5}$$

where Y is the vector containing all incomes, y_i with $i = 1,\ldots,n$ is the income of individual i, z is the poverty line, and $\alpha > 0$ is a penalization parameter.

To compute the SIGI, the value 2 is chosen for α as the square function has the advantage of easy interpretation. With $\alpha = 2$ the *transfer principle* is satisfied (Foster et al., 1984). In the context of poverty this principle means that a transfer from a person below the poverty line to a person less poor will raise poverty if the set of poor remains

Munda and Nardo (2005a,b).

unchanged. In the case of the SIGI, the transfer principle means that an increase in inequality in one dimension and a decrease of inequality in another dimension of the same magnitude will raise the SIGI.

Some differences between the SIGI and the FGT measures must be highlighted. In the case of the SIGI, we are aggregating across dimensions and not over individuals. Moreover, in contrast to the income case, a lower value of x_i is preferred, and the normalization achieved when dividing by the poverty line z is not necessary as $0 \leq x_i \leq 1,\ i = 1, \ldots, n$.

The SIGI fulfills several properties. For a formal presentation of the properties and the proofs, see Appendix 2.

- *Support and range*: The value of the index can be computed for any values of the subindices, and it is always between 0 and 1.
- *Anonymity*: Neither the name of the country nor the name of the subindex have an impact on the value of the index.
- *Unanimity or Pareto Optimality*: If a country has values for every subindex that are lower than or equal to those of another country, then the index value for the first country is lower than or equal to the one for the second country.
- *Monotonicity*: If one country has a lower value for the index than a second country, and a third country has the same values for the subindices as the first country, except for one subindex which is lower, then the third country has a lower index value than the second country.
- *Penalization of dispersion*: For two countries with the same average value of the subindices, the country with the lowest dispersion of the subindices gets a lower value for the index.
- *Compensation*: Although the SIGI is not conceived for changes over time this property is more intuitively understood in the following way. If a country experiences an increase in inequality by a given amount on a subindex, then the country can only have the same value of the index as before, if there is a decrease in inequality on another subindex that is higher in absolute value than the increase.

To highlight the effects of partial compensation as compared to total compensation we computed the statistical association between the SIGI and a simple arithmetic average of the five subindices that allows for total compensation and compared the country rankings

of both measures in Appendix 3.[11] The Pearson correlation coefficient between the SIGI and the simple arithmetic average of the five subindices shows a high and statistically significant correlation between both measures (Table 8). However, when we compare the ranks of the SIGI with those obtained using a simple arithmetic average of the five subindices in Table 9, we observe that there are noticeable differences in the rankings of the 102 included countries. Examples are China and Nepal. China ranks in position 55 using the simple average, but worsens to place 83 in the SIGI ranking. Nepal has place 84 considering the simple average, and improves to rank 65 using the SIGI. For China, this is due to the high value on the subindex *Son preference*, which in the SIGI case cannot be fully compensated with relatively low values for the other subindices. For Nepal we observe the opposite case as all subindices have values reflecting moderate inequality.

5 Results

5.1 Country Rankings and Regional Patterns

In Appendix 4, the results for the SIGI and its five subindices are presented. Among the 102 countries considered by the SIGI[12] (Table 10) Paraguay, Croatia, Kazakhstan, Argentina and Costa Rica have the lowest levels of gender inequality related to social institutions. Sudan is the country that occupies the last position, followed by Afghanistan, Sierra Leone, Mali and Yemen, which means that gender inequality in social institutions is a major problem there.

Rankings according to the subindices are as follows. For *Family code* 112 countries can be ranked. Best performers are China, Jamaica, Croatia, Belarus and Kazakhstan. Worst performers are Mali, Chad, Afghanistan, Mozambique and Zambia. In the dimension *Civil liberties* 123 countries are ranked. Among them 83 share place 1 in the ranking. Sudan, Saudi Arabia, Afghanistan, Yemen and Iran occupy the last five positions of high inequality. 114 countries can be compared with the subindex *Physical Integrity*. Hong Kong, Bangladesh, Chinese Taipei, Ecuador, El Salvador, Paraguay and Philippines are at the top of the ranking while Mali, Somalia, Sudan, Egypt and Sierra Leone are at the bottom. In the dimension *Son preference* 88 out of 123 countries rank at the top as they

[11] We cannot compare the SIGI with the results of the non-compensatory index as proposed by Munda and Nardo (2005a,b). The algorithm used for calculating non-compensatory indices compares pairwise each country for each subindex. However, as our dataset includes many countries with equal values on several subindices, the numerical algorithm cannot provide a ranking.

[12] The subindices are computed for countries that have no missing values on the relevant input variables. In the case of the SIGI only countries that have values for every subindex are considered.

do not have problems with missing women. The countries that rank worst are China, Afghanistan, Papua New Guinea, Pakistan, India and Bhutan. Finally, 122 countries are ranked with the subindex *Ownership rights*. 42 countries share position 1 as they have no inequality in this dimension. On the other hand the four worst performing countries are Sudan, Sierra Leone, Chad and the Democratic Republic of Congo.

To find out whether apparent regional patterns in social institutions related to gender inequality are systematic, we divide the countries in quintiles following the scores of the SIGI and its subindices (Table 11 in Appendix 5). The first quintile includes countries with lowest inequality, and the fifth quintile countries with highest inequality.

For the SIGI, no country of Europe and Central Asia (ECA) or Latin America and the Caribbean (LAC) is found in the two quintiles reflecting social institutions related to high gender inequality. In contrast, most countries in South Asia (SA), Sub-Saharan Africa (SSA), and Middle East and North Africa (MENA) rank in these two quintiles. It is interesting to note that in the most problematic regions two countries rank in the first two quintiles. These are Mauritius (SSA) and Tunisia (MENA). East Asia and Pacific (EAP) has countries in all five quintiles with Philippines, Thailand, Hong Kong and Singapore in the first quintile and China in the fifth quintile.

Going on with the subindices the patterns are similar to the one of the SIGI. As more information is available for the subindices, the number of countries covered by every subindex is different and higher than for the SIGI. In the following some interesting facts are highlighted, especially countries whose scores are different than the average in the region.

- *Family code*: No country in ECA, LAC or EAP shows high inequality. SA, MENA and SSA remain problematic with countries with social institutions related to high gender inequality. Exceptions are Bhutan in SA, Mauritius in SSA, and Tunisia and Israel in MENA.

- *Civil liberties*: Only three groups of countries using the quintile analysis can be generated with the first group including the first three quintiles. In SSA over one-half of the countries are now in the first group. Also in MENA there are some countries with good scores (Israel, Morocco and Tunisia). No country in SA is found in the first three quintiles of low and moderate inequality.

- *Physical integrity*: Most problematic regions are SSA and MENA. Exceptions in these regions are Botswana, Mauritius, South Africa and Tanzania (SSA), and Morocco and Tunisia (MENA).

- *Son preference*: Again only three groups of countries can be built by quintile analysis, with the first group including the first three quintiles. As in the case of Civil liberties most of the countries in SSA do not show problems. Missing women is mainly an issue in SA and MENA. But in both regions there are countries that rank in the first group. These are Sri Lanka in SA, and Israel, Lebanon and Occupied Palestinian Territory in MENA.

- *Ownership rights*: Most problematic regions are SA, SSA and MENA. Nevertheless, there are cases in these regions that rank in the first quintile. These are Egypt, Israel, Kuwait and Tunisia (MENA), Bhutan (SA), and Eritrea and Mauritius (SSA).

5.2 Simple Correlation with other Gender-related Indices

The SIGI is an important measure to understand gender inequality as it measures institutions that influence the basic functioning of society and explain gender inequality in outcomes. From this perspective, the SIGI has an added value to other gender-related measures irrespective from an empirical redundancy perspective, i.e. whether it provides additional information as compared to other measures.

Nevertheless, one can check whether the index is empirically redundant with an empirical analysis of the statistical association between the SIGI and other well-known gender-related indices. Relying on McGillivray and White (1993) we use a correlation coefficient of 0.80 in absolute value as the threshold to separate redundancy from non-redundancy.

We calculated Pearson correlation coefficient and Kendall tau b as a measure of rank correlation between the SIGI and each of the following indices: the Gender-related Development Index (GDI) and the Gender Empowerment Measure (GEM) from United Nations Development Programme (2006), the Global Gender Gap Index (GGG) from Hausmann, Tyson, and Zahidi (2007) and the Women's Social Rights Index.[13] As the GDI and the GEM have been criticized in the literature (e.g. Klasen, 2006; Schüler, 2006), we also do the analysis for two alternative measures, the Gender Gap Index Capped (GGI) and a revised Gender Empowerment Measure (GEM2) based on income shares proposed by Klasen and Schüler (2009).[14] For all the indices considered both measures of statistical

[13] Data obtained from http://ciri.binghamton.edu/.

[14] The Gender Gap Index Capped (GGI) is a geometric mean of the ratios of female to male achievements in the dimensions health, education and labor force participation. Capped means that every component is capped at one before calculating the geometric mean. This is necessary as a better relative performance of women, e.g. in the dimension health can be due to a risky behavior of men that should not be rewarded. GGI can be more directly interpreted as a measure of gender inequality while the GDI measures human development penalizing gender inequality. The GEM has three components, political representation, representation in senior positions in the economy, and power over economic resources.

association are lower than 0.80 in absolute value and statistically significant. We conclude that the SIGI is related to these gender measures but is non-redundant. These results as well as the comparison of the country rankings of the SIGI and these other measures can be found in Tables 12 and 13 (Appendix 6).

5.3 Regression Analysis

The SIGI is aimed to measure the institutional basis of gender inequality. To explore whether the SIGI is associated with gender inequality in outcomes we use linear regressions with two well-known measures as dependent variables and the SIGI as regressor. The first is the Global Gender Gap Index (GGG) that captures gaps in outcome variables related to basic rights such as health, economic participation and political empowerment. The second measure is the ratio of GDI to HDI as composite measure of gender inequality in the dimensions health, education and income.[15] In both regressions we control for the level of economic development using the log of per capita GDP in constant prices (US$, PPP, base year: 2005) (World Bank, 2008); for religion using a Muslim majority and a Christian majority dummy, the left-out category being countries that have neither a majority of Muslim nor a majority of Christian population (Central Intelligence Agency, 2009); and for geography and other unexplained heterogeneity that might go together with region using region dummies, the left-out category being Sub-Saharan Africa. As the number of observations is lower than 100, we use HC3 robust standard errors proposed by Davidson and MacKinnon (1993) to account for possible heteroscedasticity in our data.

The regression using GGG as dependent variable is presented in Table **??**. It includes 72 countries and the coefficient of determination R^2 is 0.66. the SIGI is negatively associated with GGG and significant at the 1% level. The second regression with the ratio of GDI to HDI as dependent variable is shown in Table **??**. The sample consists of 78 countries and R^2 is 0.50. The SIGI is again negatively associated with the response variable and this association is statistically significant at the 1% level. The results suggest that gender inequality in well-being and empowerment is strongly associated with social institutions that shape gender roles.

Even if we include control variables in the regressions we cannot rule out omitted variable bias, but as we consider that social institutions related to gender inequality are rela-

The most problematic component is power over economic resources proxied by earned incomes. This component measures female and male earned incomes using income levels adjusted by gender gaps but not the gender gaps themselves. The revised version GEM2 uses income shares of males and females.

[15] As the GDI is not a measure of gender inequality, UNDP recommends using the ratio of GDI to HDI (http://hdr.undp.org/en/statistics/indices/gdi_gem/).

tively stable and long-lasting, we consider that endogeneity does not pose a major problem. To check that our findings are not driven by observations that have large residuals and/or high leverage, we also run robust regressions obtaining similar results.[16]

6 Conclusion

In this paper we present composite indices that offer a new way to approach gender inequality that has been neglected in the literature and by other gender measures that focus mainly on well-being and agency. Instead of measuring gender inequality in education, health, economic or political participation and other dimensions, the proposed measures proxy the underlying social institutions that are mirrored by societal practices and legal norms that might produce inequalities between women and men in developing countries.

Based on 12 variables of the OECD Gender, Institutions and Development (GID) Database (Morrison and Jütting, 2005; Jütting et al., 2008) we construct five subindices capturing each one dimension of social institutions related to gender inequality: *Family code*, *Civil liberties*, *Physical integrity*, *Son preference* and *Ownership rights*. The Social Institutions and Gender Index (SIGI) combines the subindices to a multidimensional index of deprivation of women caused by social institutions related to gender inequality. With these measures over 100 developing countries can be compared and ranked.

When constructing composite indices one is always confronted with decisions and trade-offs concerning for example the choice and treatment of the variables included, the weighting scheme and the aggregation method. We try to be transparent in our choices. As the subindices are intended to proxy each one dimension of social institutions, we use the method of polychoric PCA to extract the common element of the included variables (Kolenikov and Angeles, 2009). The methodology for constructing the multidimensional SIGI is based on the assumption that in each dimension deprivation of women increases more than proportionally when inequality increases, and that each dimension should be weighted equally. The formula of the SIGI is inspired by the FGT poverty measures (Foster et al., 1984) and has the advantage of penalizing high inequality in each dimension and only allowing for partial compensation among the five dimensions. We consider that the formula to compute the SIGI is easy to understand and to communicate.

[16] Results are available upon request. The type of robust regression we perform uses iteratively reweighted least squares and is described in Hamilton (1992). A regression is run with ordinary least squares, then case weights based on absolute residuals are calculated, and a new regression is performed using these weights. The iterations continue as long as the maximum change in weights remains above a specified value.

However, some limitations of the subindices and the SIGI must be noted. First, a composite index depends on the quality of the data used as input. Social institutions related to gender inequality are hard to measure and the work accomplished by the OECD building the GID database is an important step forward. It is worth to continue this endeavor and invest more resources in the measurement of social institutions related to gender inequality. This includes data coverage, coding schemes and the refinement of indicators. It would be useful to exploit data available, for example from Demographic and Health Surveys (DHS)[17] that specifically address the perception that women have of violence against women, and to finance surveys in countries where data is not available.

Second, by aggregating variables and subindices, some information is inevitably lost. Figures and rankings according to the SIGI and the subindices should not substitute a careful investigation of the variables from the database. Furthermore, to understand the situation in a given country additional qualitative information could be valuable.

Third, one should keep in mind that OECD countries are not included in our sample as social institutions related to gender inequality in these countries are not well captured by the 12 variables used for building the composite measures. This does not mean that this phenomenon is not relevant for OECD countries, but that further research is required to develop appropriate measures.

Nonetheless, the SIGI and its subindices offer a new perspective to understand gender inequality. Empirical results show that the SIGI is statistically non-redundant and adds new information to other well-known gender-related measures. The SIGI and the five subindices can help policy-makers to detect in what developing countries and in which dimensions of social institutions problems need to be addressed. For example, according to the SIGI scores, regions with highest inequality are South Asia, Sub-Saharan Africa, and Middle East and North Africa. The composite measures can be valuable instruments to generate public discussion. Moreover, the SIGI and its subindices have the potential to influence current development thinking as they highlight social institutions that affect overall development. As it is shown in the literature (e.g. Klasen, 2002; Klasen and Lamanna, 2009) gender inequality in education negatively affects overall development. Economic research investigating these outcome inequality should consider social institutions related to gender inequality as possible explanatory factors. Results from regression analysis show that the SIGI is related to gender inequality in well-being and empowerment, even after controlling for region, religion and the level of economic development.

[17] Information is available on the webpage http://www.measuredhs.com/.

References

Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.

Alkire, S. and J. E. Foster (2008). Counting and multidimensional poverty measurement. OPHI Working Paper 7, Oxford Poverty and Human Development Initiative, Queen Elizabeth House, University of Oxford.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*(3), 244–263.

Bjornskov, C., A. Dreher, and J. A. Fischer (2009). On gender inequality and life satisfaction: Does discrimination matter? Mimeo.

Central Intelligence Agency (2009). The world factbook. Electronic publication. https://www.cia.gov/library/publications/the-world-factbook/.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*. New York: Oxford Univ. Press.

De Soysa, I. and J. Jütting (2007). Informal institutions and development: How they matter and what makes them change. In J. Jütting, D. Drechsler, S. Bartsch, and I. de Soysa (Eds.), *Informal Institutions. How Social Norms Help or Hinder Development*, pp. 29–43. Paris: OECD Development Centre.

Deaton, A. (1997). *The analysis of household surveys: A microeconometric approach to development policy*. Baltimore: John Hopkins University Press.

Duflo, E. (2005). Gender equality in development. BREAD Policy Paper 011, Bureau for Research in Economic Analysis of Development.

Economic Commission for Africa (2004). *The African Gender and Development Index*. Addis Ababa: ECA.

Foster, J. E., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica 52*, 761–766.

Greenacre, M. (2007). *Correspondence Analysis in Practice* (second ed.). Interdisciplinary Statistics. Boca Raton: Chapman and Hall.

Hall, P. A. and R. C. R. Taylor (1996). Political science and the three New Institutionalisms. *MPIFG Discussion Paper 96/6*.

Hamilton, L. C. (1992). How robust is robust regression? *Stata Technical Bulletin 1*(2).

Hausmann, R., L. D. Tyson, and S. Zahidi (2007). *The Global Gender Gap Report 2007*. Geneva: World Economic Forum.

Jolliffe, I. T. (1986). *Principal component analysis*. New York, NY: Springer.

Jütting, J., C. Morrison, J. Dayton-Johnson, and D. Drechsler (2008). Measuring gender (In)Equality: The OECD gender, institutions and development data base. *Journal of Human Development 9*(1), 65–86.

Kakwani, N. C. (1984). Issues in measuring poverty. In R. L. Basmann and J. G. F. Rhodes (Eds.), *Advances in Econometrics*, Volume 3. Greenwich, CT and London: JAI Press.

Klasen, S. (2002). Low schooling for girls, slower growth for all? *World Bank Economic Review 16*(3), 345–373.

Klasen, S. (2006). UNDP's gender-related measures: Some conceptual problems and possible solutions. *Journal of Human Development 7*(2), 243–274.

Klasen, S. (2007). Gender-related indicators of well-being. In M. McGillivray (Ed.), *Human Well-Being: Concept and Measurement*, Studies in Development Economics and Policy, Chapter 7, pp. 167–192. New York, NY: Palgrave Macmillan.

Klasen, S. and F. Lamanna (2009). The impact of gender inequality in education and employment on economic growth in developing countries: Updates and extensions. *Feminist Economics 15*(3), 91–132.

Klasen, S. and D. Schüler (2009). Reforming the gender-related development index (GDI) and the gender empowerment measure (GEM): Some specific proposals. Ibero-America Institute for Economic Research (IAI) Discussion Papers 186, University of Goettingen.

Klasen, S. and C. Wink (2003). Missing women: Revisiting the debate. *Feminist Economics 9*, 263–300.

Kolenikov, S. and G. Angeles (2004). The use of discrete data in PCA: Theory, simulations, and applications to socioeconomics indices. CPC/MEASURE Working paper WP-04-85, Carolina Population Center.

Kolenikov, S. and G. Angeles (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth 55*(1), 128–165.

Lopez-Claros, A. and S. Zahidi (2005). *Women's Empowerment: Measuring the Global Gender Gap*. Davos: World Economic Forum.

McGillivray, M. and H. White (1993). Measuring development? The UNDP's Human Development Index. *Journal of International Development 5*(2), 183–192.

Morrison, C. and J. P. Jütting (2005). Women's discrimination in developing countries: A new data set for better policies. *World Development 33*(7), 1065–1081.

Munda, G. and M. Nardo (2005a). Constructing consistent composite indicators: The issue of weights. Technical Report EUR 21834 EN, European Commission.

Munda, G. and M. Nardo (2005b). Non-compensatory composite indicators for ranking countries: A defensible setting. Technical Report EUR 21833 EN, European Commission.

Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, and E. Giovannini (2005). Handbook on constructing composite indicators: Methodology and user guide. Technical Report 2005/3, OECD.

Nenadić, O. (2007). *An Implementation of Correspondence Analysis in R and its Application in the Analysis of Web Usage*. Ph. D. thesis, University of Goettingen, Goettingen.

North, D. C. (1990). *Institutions, Institutional Change, and Economic Performance*. New York: Cambridge University Press.

Ravallion, M. (1994). *Poverty Comparisons: A Guide to Concepts and Methods*. Chur: Harwood Academic.

Schüler, D. (2006). The uses and misuses of the Gender-related Development Index and the Gender Empowerment Measure: A review of the literature. *Journal of Human Development 7*(2), 161–182.

Sen, A. (1992). Missing women. *British Medical Journal 304*, 586–7.

Sen, A. (1999). *Develoment as Freedom*. New York, NY: Alfred A. Knopf.

Shorrocks, A. F. and J. E. Foster (1987). Transfer sensitive inequality measures. *Review of Economic Studies 54*(1), 485–497.

Social Watch (2005). *Roars and Whispers Gender and Poverty: Promises versus Action*. Montevideo: Social Watch.

Subramanian, S. (2007). Indicators of inequality and poverty. In M. McGillivray (Ed.), *Human Well-Being: Concept and Measurement*, Studies in Development Economics and Policy, Chapter 6, pp. 135–166. New York, NY: Palgrave Macmillan.

United Nations Development Programme (1995). *Human Development Report*. New York: Oxford University Press.

United Nations Development Programme (2006). *Human Development Report*. New York: Oxford University Press.

World Bank (2001). *Engendering Development: Through Gender Equality in Rigths, Resources and Voices*. New York, NY: World Bank/Oxford University Press.

World Bank (2008). World development indicators. Technical report, World Bank.

Appendix 1: Building the Subindices

Kendall tau and Weights from Polychoric PCA

Kendall tau b: Dimension Family Code

Table 1:

		earmarr	polyg	parauth	inher
earmarr	Kendall tau b	1			
	Number of obs.	112			
	p-Value				
polyg	Kendall tau b	0.2950	1		
	Number of obs.	112	112		
	p-Value	0.0001			
parauth	Kendall tau b	0.2884	0.4792	1	
	Number of obs.	112	112	112	
	p-Value	0.0001	0.0000		
inher	Kendall tau b	0.234	0.5964	0.5742	1
	Number of obs.	112	112	112	112
	p-Value	0.0020	0.0000	0.0000	

earmarr stands for the variable Early marriage, polyg for Polygamy, parauth is the variable Parental authority and inher is the variable inheritance. For a description of these variables, see section 2. The p-values correspond to the null hypothesis that the two variables are independent.

Kendall tau b: Dimension Civil Liberties

Table 2:

		obliveil
freemov	Kendall tau b	0.613
	Number of obs.	123
	p-Value	0.0000

freemov stands for the variable Freedom of movement. obliveil is the variable Obligation to wear a veil in public. For a description of these variables, see section 2. The p-value correspond to the null hypothesis that two variables are independent.

Kendall tau b: Dimension Physical Integrity with Missing Women

Table 3:

		femmut	vio	misswom
femmut	Kendall tau b	1		
	Number of obs.	114		
	p-Value			
vio	Kendall tau b	0.1584	1	
	Number of obs.	114	114	
	p-Value	0.0382		
misswom	Kendall tau b	-0.1041	0.1098	1
	Number of obs.	114	114	114
	p-Value	0.2160	0.1634	

femmut stands for the variable Female Genital Mutilation, vio for Violence against women and misswom is the variable Missing women. For a description of these variables, see section 2. The p-values correspond to the null hypothesis that the two variables are independent.

Kendall tau b: Dimension Physical Integrity without Missing Women

Table 4:

		vio
femmut	Kendall tau b	0.1584
	Number of obs.	114
	p-Value	0.0382

femmut stands for the variable Female Genital Mutilation and vio for Violence against women. For a description of these variables, see section 2. The p-value correspond to the null hypothesis that two variables are independent.

Kendall tau b: Dimension Ownership Rights

Table 5:

		womland	womloans	womprop
womland	Kendall tau b	1		
	Number of obs.	122		
	p-Value			
womloans	Kendall tau b	0.5943	1	
	Number of obs.	122	122	
	p-Value	0.0000		
womprop	Kendall tau b	0.6438	0.5975	1
	Number of obs.	122	122	122
	p-Value	0.0000	0.0000	

womland stands for the variable Women's access to land. womloans is the variable Women's access to loans and womprop is the variable Women's access to property other than land. For a description of these variables, see section 2. The p-values correspond to the null hypothesis that the two variables are independent.

Weights from Polychoric PCA

Table 6:

	Weights
Family code	
Parental authority	0.5212
Inheritance	0.5404
Early marriage	0.3877
Polygamy	0.5348
Civil liberties	
Freedom of movement	0.7071
Obligation to wear a veil	0.7071
Physical integrity	
Female genital mutilation	0.7071
Violence against women	0.7071
Ownership rights	
Woment's access to land	0.5811
Woment's access to loans	0.5665
Woment's access to other property	0.5843

Kendall tau b between Subindices

Table 7:

		Family code	Civil liberties	Physical integrity	Son preference	Ownership rights
Family code	Kendall tau b	1				
	Number obs.	112				
Civil liberties	Kendall tau b	0.3844	1			
	Number obs.	112	123			
	p-value	0.0000				
Physical integrity	Kendall tau b	0.4367	0.2648	1		
	Number obs.	103	113	114		
	p-value	0.0000	0.0005			
Son preference	Kendall tau b	0.1603	0.4264	0.0272	1	
	Number obs.	112	122	114	123	
	p-value	0.0317	0.0000	0.7220		
Ownership rights	Kendall tau b	0.5484	0.3047	0.3937	0.1039	1
	Number obs.	111	121	112	121	122
	p-value	0.0000	0.0001	0.0000	0.181	

Appendix 2: Objectives, Properties and Proofs

In this section, we present the objectives and properties that we consider relevant for any composite index related to social institutions related to gender inequality. Moreover, we show that the proposed index fulfills all of them.

We use the following notation. Let X^j, with $j = A, B$, be the vector containing the the values of the subindices x_i^j, with $i = 1, ..., n$, for the country j[18]. $I(X)$ represents the composite index.

Objectives of the Index

The objectives of the index are the following:

1. The index $I(X)$ should represent the level of gender inequality, so that countries can be ranked.

2. The interpretation of $I(X)$ should be straightforward. As in the case of the subindices x_i, the value 0 should correspond to no inequality and the value 1 to complete inequality.

3. For any subindex x_i, we interpret the value 0, i.e. no inequality, as the goal to be achieved. The value zero can be thought of as a poverty line (see Ravallion, 1994; Deaton, 1997; Subramanian, 2007, and references therein). We define a deprivation function $\phi(x_i, 0)$, with $\phi(x_i, 0) > 0$ if $x_i > 0$, and $\phi(x_i, 0) = 0$ if $x_i = 0$. Higher values of x_i should lead to a penalization in $I(X)$ that should increase with the distance x_i to zero, i.e. $\frac{\partial I(X)}{\partial x_i} > 0$, and $\frac{\partial^2 I(X)}{\partial x_i^2} > 0$.

4. $I(X)$ should not allow for total compensation among variables, but permit partial compensation. This somehow relates to the transfer axioms that should be fulfilled by inequality as well as poverty measures. A decrease in x_i, i.e. less inequality, is rewarded more in $I(X)$ than an equivalent increase in another variable x_k (see Atkinson, 1970; Kakwani, 1984; Shorrocks and Foster, 1987; Subramanian, 2007; Alkire and Foster, 2008, and references therein).

5. $I(X)$ should be easy to compute and transparent.

Properties of the Index

Some of the properties that any index should fulfill are:

1. **Support and range of** $I(X)$:

   - $I(X)$ must be defined for $0 \leq x_i \leq 1,\ i = 1, \ldots, n$.

[18] In what follows, the superscript j will only be used if it is necessary to distinguish countries.

- $0 \leq I(X) \leq 1$ must hold for any X.
- If $x_i = 0\ \forall i$, then $I(X) = 0$. If $x_i = 1\ \forall i$, then $I(X) = 1$.

2. **Anonymity (symmetry)**: The value of $I(X^j)$ does not depend either on the names of the subindices nor on the name of the country (j).

3. **Unanimity (Pareto Optimality)**: If $x_i^A \leq x_i^B\ \forall i$, then $I(X^A) \leq I(X^B)$.

4. **Monotonicity**: If considering X^A and X^B country A is preferred to country B, and only x_i^A improves (i.e. decreases) for a given i, while $x_i^B\ \forall i$ remains unchanged, then country A should still be preferred over country B.

5. **Penalization of inequality in the case of equal means**: Let the mean of X^A be equal to the mean of X^B. If the dispersion of X^A is smaller than the dispersion of X^B, then $I(X^A) < I(X^B)$.

6. **Compensation property**: In a two-variable example, $\triangle x_1 \leq 1 - x_1$, and $\triangle x_2 \leq 1 - x_2$.

   a) If x_1 increases by $|\triangle x_1|$ and x_2 decreases by $|\triangle x_2|$ and $|\triangle x_1| = |\triangle x_2|$, then $I(X)$ must increase.

   b) For $I(X)$ to remain unchanged, we must have $|\triangle x_2| > |\triangle x_1|$.

Proofs

The composite index $I(X)$ is defined as

$$I(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2.$$

The index proposed fulfills all the stated properties.

1. **Support and range of $I(X)$**

   - $I(X)$ is defined for $0 \leq x_i \leq 1,\ i = 1, \ldots, n$.
   - For any X, we have that $0 \leq I(X) \leq 1$.
   - If $x_i = 0\ \forall i$, then $I(X) = 0$. If $x_i = 1\ \forall i$, then $I(X) = 1$.

2. **Anonymity (symmetry)**

   The value of $I(X^j)$ does not depend either on the names of the subindices nor on the name of the country (j).

3. **Unanimity (Pareto Optimality)**

If we assume that $\forall i$

$$x_i^A \leq x_i^B,$$

then we can show that

$$\begin{aligned} (x_i^A)^2 &\leq (x_i^B)^2 \\ \frac{1}{n}\sum_{i=1}^{n}(x_i^A-0)^2 &\leq \frac{1}{n}\sum_{i=1}^{n}(x_i^B-0)^2 \\ I(X^A) &\leq I(X^B). \end{aligned}$$

4. **Monotonicity**

We assume that

$$\begin{aligned} I(X^A) &\leq I(X^B) \\ \frac{1}{n}\sum_{i=1}^{n}(x_i^A-0)^2 &\leq \frac{1}{n}\sum_{i=1}^{n}(x_i^B-0)^2. \end{aligned}$$

Let us suppose, without loss of generality, that subindex x_1 improves (decreases) by $\delta > 0$ for country A. Then we have that

$$\frac{1}{n}(x_1^A-\delta-0)^2+\frac{1}{n}\sum_{i=2}^{n}(x_i^A-0)^2 \leq \frac{1}{n}\sum_{i=1}^{n}(x_i^A-0)^2,$$

and hence

$$\frac{1}{n}(x_1^A-\delta-0)^2+\frac{1}{n}\sum_{i=2}^{n}(x_i^A-0)^2 \leq \frac{1}{n}\sum_{i=1}^{n}(x_i^B-0)^2.$$

This means that

$$I(X^{A^*}) \leq I(X^B)$$

with X^{A^*} defined as the vector corresponding to country A with only one variable having improved (decreased) by δ.

5. **Penalization of inequality in the case of equal means**

If we assume equal means, so that

$$\mu = \frac{1}{n}\sum_{i=1}^{n}(x_i^A) = \frac{1}{n}\sum_{i=1}^{n}(x_i^B),$$

then we also have

$$\sum_{i=1}^{n}(x_i^A) = \sum_{i=1}^{n}(x_i^B).$$

If we assume that the variance of X^A is smaller than the variance of X^B so that

$$\frac{1}{n}\sum_{i=1}^{n}(x_i^A-\mu)^2 \quad < \quad \frac{1}{n}\sum_{i=1}^{n}(x_i^B-\mu)^2,$$

we can show that

$$\begin{aligned}
\sum_{i=1}^{n}\left[(x_i^A)^2-2\mu x_i^A+\mu^2)\right] &< \sum_{i=1}^{n}\left[(x_i^B)^2-2\mu x_i^B+\mu^2)\right], \\
\sum_{i=1}^{n}(x_i^A)^2-2\mu\sum_{i=1}^{n}x_i^A+n\mu^2 &< \sum_{i=1}^{n}(x_i^B)^2-2\mu\sum_{i=1}^{n}x_i^B+n\mu^2.
\end{aligned}$$

As $\sum_{i=1}^{n}(x_i^A) = \sum_{i=1}^{n}(x_i^B)$, we have that

$$\begin{aligned}
\sum_{i=1}^{n}(x_i^A)^2 &< \sum_{i=1}^{n}(x_i^B)^2 \\
\frac{1}{n}\sum_{i=1}^{n}(x_i^A-0)^2 &< \frac{1}{n}\sum_{i=1}^{n}(x_i^B-0)^2 \\
I(X^A) &< I(X^B).
\end{aligned}$$

6. **Compensation property**

In a two-variable example, let $\triangle x_1 \leq 1-x_1$, and $\triangle x_2 \leq 1-x_2$.

a) We can show that if $\triangle x_1 = \triangle x_2 = \delta > 0$, then

$$\begin{aligned}
x_2 &< x_1+\delta \\
0 &< x_1-x_2+\delta \\
0 &< 2\delta(x_1-x_2+\delta) \\
x_1^2+x_2^2 &< x_1^2+x_2^2+2\delta(x_1-x_2+\delta) \\
\frac{1}{2}\left(x_1^2+x_2^2\right) &< \frac{1}{2}\left(x_1^2+2\delta x_1+\delta^2+x_2^2-2\delta x_2+\delta^2\right) \\
\frac{1}{2}\left(x_1^2+x_2^2\right) &< \frac{1}{2}\left[(x_1^2+\delta)^2+(x_2^2-\delta)^2\right] \\
I(x_1,x_2) &< I(x_1+\delta,x_2-\delta),
\end{aligned}$$

and hence we have shown that if x_1 increases by δ and x_2 decreases by δ, then $I(X)$ must increase.

b) Let $x_1 = x_2 = x > 0$. We will show that if x_1 increases by $\triangle x_1$ and x_2 decreases by

$\triangle x_1$ and the value of the index remains unchanged, the increase of x_1 must be smaller than the absolute value of the decrease in x_2.

$$\begin{aligned} I(x_1,x_2) &= I(x_1+\triangle x_1, x_2-\triangle x_2) \\ \frac{1}{2}\left(x_1^2+x_2^2\right) &= \frac{1}{2}\left[(x_1+\triangle x_1)^2+(x_2-\triangle x_2)^2\right] \\ x_1^2+x_2^2 &= x_1^2+2x_1\triangle x_1+(\triangle x_1)^2+x_2^2-2x_2\triangle x_2+(\triangle x_2)^2 \\ 0 &= 2x_1\triangle x_1+(\triangle x_1)^2-2x_2\triangle x_2+(\triangle x_2)^2 \end{aligned}$$

Using the fact that $x_1 = x_2 = x$, we can rewrite this as

$$\begin{aligned} 0 &= 2x\triangle x_1+(\triangle x_1)^2-2x\triangle x_2+(\triangle x_2)^2 \\ 0 &= 2x(\triangle x_1-\triangle x_2)+(\triangle x_1)^2+(\triangle x_2)^2. \end{aligned}$$

As $2x > 0$, $(\triangle x_1)^2 > 0$, and $(\triangle x_2)^2 > 0$, we must have that

$$\begin{aligned} \triangle x_1-\triangle x_2 &< 0 \\ \triangle x_1 &< \triangle x_2. \end{aligned}$$

Appendix 3: Comparison of SIGI with the Simple Average of the Subindices

Pearson Correlation Coefficient (ρ) between the SIGI and the Simple Average of the Five Subindices

Table 8:

ρ	0.9593
Number obs.	102
p-value	0.0000

Comparison of the SIGI and the Simple Average of the Subindices

Table 9:

	SIGI		Simple Aver.		Simple Aver. Rank
Country	**Ranking**	**Value**	**Ranking**	**Value**	**minus SIGI rank**
Paraguay	1	0.0024832	2	0.0312943	1
Croatia	2	0.00333	1	0.0273771	-1
Kazakhstan	3	0.0034778	3	0.0314302	0
Argentina	4	0.0037899	4	0.0354832	0
Costa Rica	5	0.0070934	5	0.0502099	0
Russian Federation	6	0.0072524	11	0.0538114	5
Philippines	7	0.0078831	15	0.0603212	8
El Salvador	8	0.0082581	16	0.0647861	8
Ecuador	9	0.0091447	18	0.0700484	9
Ukraine	10	0.00969	6	0.051376	-4
Mauritius	11	0.009759	7	0.0521866	-4
Moldova	12	0.0098035	8	0.052673	-4
Bolivia	13	0.0098346	9	0.0529972	-4
Uruguay	14	0.0099167	10	0.0538078	-4
Venezuela, RB	15	0.0104259	13	0.0578608	-2
Thailand	16	0.010677	17	0.0652957	1
Peru	17	0.0121323	14	0.0586566	-3
Colombia	18	0.012727	24	0.0828911	6
Belarus	19	0.0133856	12	0.0563755	-7
Hong Kong, China	20	0.0146549	19	0.07076	-1
Singapore	21	0.0152573	20	0.0714613	-1

Continued on next page

Table 9 – continued from previous page

Country	SIGI Ranking	SIGI Value	Simple Aver. Ranking	Simple Aver. Value	Simple Aver. Rank minus SIGI rank
Cuba	22	0.0160304	22	0.0750193	0
Macedonia, FYR	23	0.0178696	23	0.0818509	0
Brazil	24	0.0188021	21	0.073534	-3
Tunisia	25	0.0190618	29	0.1012313	4
Chile	26	0.0195128	31	0.106534	5
Cambodia	27	0.0220188	27	0.0886198	0
Nicaragua	28	0.0225149	32	0.1117536	4
Trinidad and Tobago	29	0.0228815	34	0.1143368	5
Kyrgyz Republic	30	0.0292419	36	0.12716	6
Viet Nam	31	0.0300619	25	0.0837526	-6
Armenia	32	0.0301177	26	0.0845632	-6
Georgia	33	0.0306926	28	0.0902375	-5
Guatemala	34	0.0319271	35	0.124404	1
Tajikistan	35	0.0326237	37	0.137724	2
Honduras	36	0.0331625	33	0.1122453	-3
Azerbaijan	37	0.0339496	30	0.1058964	-7
Lao PDR	38	0.0357687	39	0.1416411	1
Mongolia	39	0.0391165	43	0.1680587	4
Dominican Republic	40	0.0398379	40	0.1440229	0
Myanmar	41	0.0462871	42	0.1553233	1
Jamaica	42	0.0484293	38	0.1399837	-4
Morocco	43	0.0534361	45	0.1973177	2
Fiji	44	0.0545044	41	0.1551223	-3
Sri Lanka	45	0.059141	47	0.2106919	2
Madagascar	46	0.0695815	44	0.1938462	-2
Namibia	47	0.0750237	49	0.241875	2
Botswana	48	0.0810172	46	0.2027736	-2
South Africa	49	0.0867689	53	0.2565411	4
Burundi	50	0.1069056	52	0.2488075	2
Albania	51	0.1071956	58	0.2715919	7
Senegal	52	0.1104056	50	0.2424129	-2
Tanzania	53	0.1124419	51	0.2445237	-2
Ghana	54	0.112694	54	0.2568415	0
Indonesia	55	0.1277609	57	0.2692867	2
Eritrea	56	0.1364469	48	0.2288967	-8
Kenya	57	0.1370416	56	0.2673039	-1
Cote d'Ivoire	58	0.1371181	59	0.2862332	1
Syrian Arab Republic	59	0.1381059	74	0.3619356	15
Malawi	60	0.1432271	65	0.330963	5

Continued on next page

Table 9 – continued from previous page

Country	SIGI Ranking	SIGI Value	Simple Aver. Ranking	Simple Aver. Value	Simple Aver. Rank minus SIGI rank
Mauritania	61	0.1497032	68	0.3336183	7
Swaziland	62	0.1565499	70	0.3456205	8
Burkina Faso	63	0.1616069	60	0.3030649	-3
Bhutan	64	0.162508	63	0.3196661	-1
Nepal	65	0.1672252	84	0.3973769	19
Rwanda	66	0.1685859	61	0.3059172	-5
Niger	67	0.1755873	72	0.3537308	5
Equatorial Guinea	68	0.1759719	76	0.3676708	8
Gambia, The	69	0.1782978	62	0.3177497	-7
Central African Republic	70	0.1843973	67	0.3323123	-3
Kuwait	71	0.1860213	79	0.3723096	8
Zimbabwe	72	0.1869958	78	0.3685864	6
Uganda	73	0.1871794	80	0.3735746	7
Benin	74	0.1889945	66	0.3319663	-8
Algeria	75	0.190244	87	0.4123239	12
Bahrain	76	0.1965476	89	0.4310629	13
Mozambique	77	0.1995442	82	0.3808849	5
Togo	78	0.202518	69	0.343517	-9
Congo, Dem. Rep.	79	0.2044817	64	0.3276955	-15
Papua New Guinea	80	0.2093579	83	0.3843125	3
Cameroon	81	0.2165121	85	0.4013174	4
Egypt, Arab Rep.	82	0.2176608	81	0.3779768	-1
China	83	0.2178559	55	0.2605644	-28
Gabon	84	0.2189224	86	0.4038617	2
Zambia	85	0.2193876	71	0.3526082	-14
Nigeria	86	0.2199123	92	0.4540078	6
Liberia	87	0.2265095	75	0.3629022	-12
Guinea	88	0.2280293	77	0.3678226	-11
Ethiopia	89	0.2332508	73	0.3559035	-16
Bangladesh	90	0.2446482	91	0.4491116	1
Libya	91	0.260187	94	0.5057952	3
United Arab Emirates	92	0.2657521	96	0.5082552	4
Iraq	93	0.2752427	97	0.522977	4
Pakistan	94	0.2832434	95	0.5062053	1
Iran, Islamic Rep.	95	0.3043608	98	0.5252544	3
India	96	0.318112	99	0.5295102	3
Chad	97	0.3225771	93	0.4733184	-4
Yemen	98	0.3270495	100	0.5567938	2
Mali	99	0.339493	88	0.422655	-11

Continued on next page

Table 9 – continued from previous page

Country	SIGI Ranking	SIGI Value	Simple Aver. Ranking	Simple Aver. Value	Simple Aver. Rank minus SIGI rank
Sierra Leone	100	0.3424468	90	0.4488637	-10
Afghanistan	101	0.5823044	101	0.746126	0
Sudan	102	0.6778067	102	0.800509	0

The data are sorted according to the value of the SIGI.

Appendix 4: Rankings of Countries according to the SIGI and its Subindices

Ranking according to the SIGI and the Five Subindices

Table 10:

	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
Country	**Ranking**	**Value**	**Ranking**	**Value**	**Ranking**	**Value**	**Ranking**	**Value**	**Ranking**	**Value**	**Ranking**	**Value**
Paraguay	1	0.00248	19	0.06890	1	0	3	0.08757	1	0	1	0
Croatia	2	0.00333	3	0.00811	1	0	9	0.12878	1	0	1	0
Kazakhstan	3	0.00348	5	0.02837	1	0	9	0.12878	1	0	1	0
Argentina	4	0.00379	13	0.04864	1	0	9	0.12878	1	0	1	0
Costa Rica	5	0.00709	23	0.08106	1	0	15	0.16999	1	0	1	0
Russian Federation	6	0.00725	35	0.14028	1	0	9	0.12878	1	0	1	0
Philippines	7	0.00788	8	0.04053	1	0	3	0.08757	1	0	53	0.17351
El Salvador	8	0.00826	17	0.06485	1	0	3	0.08757	1	0	43	0.17151
Ecuador	9	0.00914	24	0.08917	1	0	3	0.08757	1	0	53	0.17351
Ukraine	10	0.00969	8	0.04053	1	0	23	0.21635	1	0	1	0
Mauritius	11	0.00976	11	0.04458	1	0	23	0.21635	1	0	1	0
Moldova	12	0.00980	12	0.04701	1	0	23	0.21635	1	0	1	0
Bolivia	13	0.00983	13	0.04864	1	0	23	0.21635	1	0	1	0
Uruguay	14	0.00992	15	0.05269	1	0	23	0.21635	1	0	1	0
Venezuela, RB	15	0.01043	21	0.07295	1	0	23	0.21635	1	0	1	0
Thailand	16	0.01068	41	0.15649	1	0	15	0.16999	1	0	1	0
Peru	17	0.01213	15	0.05269	1	0	33	0.24059	1	0	1	0
Colombia	18	0.01273	21	0.07295	1	0	15	0.16999	1	0	43	0.17151
Belarus	19	0.01339	4	0.02432	1	0	34	0.25756	1	0	1	0
Hong Kong, China	20	0.01465	26	0.10380	1	0	1	0	89	0.25	1	0
Singapore	21	0.01526	25	0.09975	1	0	34	0.25756	1	0	1	0

Continued on next page

Table 10 – continued from previous page

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Cuba	22	0.01603	28	0.11754	1	0	34	0.25756	1	0	1	0
Macedonia, FYR	23	0.01787	39	0.15169	1	0	34	0.25756	1	0	1	0
Brazil	24	0.01880	19	0.06890	1	0	48	0.29877	1	0	1	0
Tunisia	25	0.01906	32	0.12738	1	0	9	0.12878	89	0.25	1	0
Chile	26	0.01951	34	0.13909	1	0	23	0.21635	1	0	56	0.17723
Cambodia	27	0.02202	38	0.14433	1	0	48	0.29877	1	0	1	0
Nicaragua	28	0.02251	33	0.12970	1	0	34	0.25756	1	0	43	0.17151
Trinidad and Tobago	29	0.02288	39	0.15169	1	0	15	0.16999	89	0.25	1	0
Kyrgyz Republic	30	0.02924	42	0.15980	1	0	48	0.29877	1	0	56	0.17723
Viet Nam	31	0.03006	6	0.03242	1	0	60	0.38634	1	0	1	0
Armenia	32	0.03012	7	0.03648	1	0	60	0.38634	1	0	1	0
Georgia	33	0.03069	17	0.06485	1	0	60	0.38634	1	0	1	0
Guatemala	34	0.03193	27	0.10538	1	0	54	0.34513	1	0	43	0.17151
Tajikistan	35	0.03262	47	0.25955	1	0	34	0.25756	1	0	43	0.17151
Honduras	36	0.03316	44	0.21610	1	0	54	0.34513	1	0	1	0
Azerbaijan	37	0.03395	37	0.14314	1	0	60	0.38634	1	0	1	0
Lao PDR	38	0.03577	51	0.32034	1	0	23	0.21635	1	0	43	0.17151
Mongolia	39	0.03912	30	0.12001	1	0	48	0.29877	89	0.25	43	0.17151
Dominican Republic	40	0.03984	28	0.11754	1	0	34	0.25756	1	0	58	0.34502
Myanmar	41	0.04629	35	0.14028	1	0	60	0.38634	89	0.25	1	0
Jamaica	42	0.04843	1	0.00405	1	0	54	0.34513	1	0	76	0.35074
Morocco	43	0.05344	48	0.26279	1	0	9	0.12878	89	0.25	58	0.34502
Fiji	44	0.05450	8	0.04053	1	0	60	0.38634	1	0	66	0.34874
Sri Lanka	45	0.05914	46	0.23404	98	0.30069	15	0.16999	1	0	66	0.34874

Continued on next page

Table 10 – continued from previous page

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Madagascar	46	0.06958	70	0.41138	1	0	60	0.38634	1	0	43	0.17151
Namibia	47	0.07502	58	0.35307	1	0	34	0.25756	89	0.25	66	0.34874
Botswana	48	0.08102	53	0.32163	1	0	15	0.16999	1	0	79	0.52225
South Africa	49	0.08677	73	0.42326	84	0.29808	23	0.21635	1	0	58	0.34502
Burundi	50	0.10691	57	0.33545	1	0	60	0.38634	1	0	79	0.52225
Albania	51	0.10720	31	0.12288	1	0	60	0.38634	101	0.5	66	0.34874
Senegal	52	0.11041	99	0.60250	1	0	45	0.26455	1	0	58	0.34502
Tanzania	53	0.11244	81	0.49886	1	0	22	0.20151	1	0	79	0.52225
Ghana	54	0.11269	61	0.36621	1	0	80	0.39575	1	0	79	0.52225
Indonesia	55	0.12776	59	0.35405	103	0.59876	79	0.39362	1	0	1	0
Eritrea	56	0.13645	76	0.45538	1	0	106	0.68910	1	0	1	0
Kenya	57	0.13704	63	0.37027	1	0	46	0.28152	1	0	111	0.68473
Cote d'Ivoire	58	0.13712	79	0.49012	1	0	85	0.43455	1	0	77	0.50650
Syrian Arab Republic	59	0.13811	68	0.40269	98	0.30069	34	0.25756	101	0.5	66	0.34874
Malawi	60	0.14323	60	0.36087	84	0.29808	88	0.47362	1	0	79	0.52225
Mauritania	61	0.14970	71	0.42056	98	0.30069	103	0.60183	1	0	58	0.34502
Swaziland	62	0.15655	86	0.52144	84	0.29808	60	0.38634	1	0	79	0.52225
Burkina Faso	63	0.16161	88	0.53939	1	0	104	0.63092	1	0	58	0.34502
Bhutan	64	0.16251	43	0.20513	84	0.29808	54	0.34513	118	0.75	1	0
Nepal	65	0.16723	62	0.36779	84	0.29808	48	0.29877	101	0.5	79	0.52225
Rwanda	66	0.16859	56	0.32974	1	0	91	0.51512	1	0	111	0.68473
Niger	67	0.17559	104	0.64882	1	0	99	0.52482	89	0.25	58	0.34502
Equatorial Guinea	68	0.17597	82	0.50291	84	0.29808	91	0.51512	1	0	79	0.52225
Gambia, The	69	0.17830	103	0.64303	1	0	102	0.59698	1	0	66	0.34874

Continued on next page

Table 10 – continued from previous page

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Central African Republic	70	0.18440	92	0.55902	1	0	101	0.58029	1	0	79	0.52225
Kuwait	71	0.18602	83	0.50523	103	0.59876	34	0.25756	101	0.5	1	0
Zimbabwe	72	0.18700	80	0.49075	84	0.29808	59	0.36937	1	0	111	0.68473
Uganda	73	0.18718	102	0.63697	84	0.29808	81	0.41058	1	0	79	0.52225
Benin	74	0.18899	84	0.50633	1	0	87	0.46877	1	0	111	0.68473
Algeria	75	0.19024	69	0.40501	103	0.59876	60	0.38634	101	0.5	43	0.17151
Bahrain	76	0.19655	52	0.32147	103	0.59876	60	0.38634	101	0.5	66	0.34874
Mozambique	77	0.19954	109	0.69776	84	0.29808	60	0.38634	1	0	79	0.52225
Togo	78	0.20252	96	0.58833	1	0	86	0.44452	1	0	111	0.68473
Congo, Dem. Rep.	79	0.20448	66	0.39038	1	0	81	0.41058	1	0	119	0.83752
Papua New Guinea	80	0.20936	50	0.27697	1	0	60	0.38634	118	0.75	78	0.50825
Cameroon	81	0.21651	89	0.54344	84	0.29808	90	0.48332	1	0	109	0.68175
Egypt, Arab Rep.	82	0.21766	49	0.26647	98	0.30069	111	0.82273	101	0.5	1	0
China	83	0.21786	1	0.00405	1	0	48	0.29877	122	1	1	0
Gabon	84	0.21892	107	0.68387	84	0.29808	91	0.51512	1	0	79	0.52225
Zambia	85	0.21939	108	0.69197	1	0	60	0.38634	1	0	111	0.68473
Nigeria	86	0.21991	71	0.42056	103	0.59876	89	0.47847	89	0.25	79	0.52225
Liberia	87	0.22651	87	0.53470	1	0	107	0.75756	1	0	79	0.52225
Guinea	88	0.22803	105	0.67140	1	0	105	0.64546	1	0	79	0.52225
Ethiopia	89	0.23325	55	0.32726	1	0	109	0.77424	1	0	108	0.67801
Bangladesh	90	0.24465	95	0.58334	103	0.59876	2	0.04121	101	0.5	79	0.52225
Libya	91	0.26019	67	0.39285	103	0.59876	91	0.51512	101	0.5	79	0.52225
United Arab Emirates	92	0.26575	93	0.56197	103	0.59876	100	0.53180	101	0.5	66	0.34874
Iraq	93	0.27524	77	0.47391	103	0.59876	98	0.51997	101	0.5	79	0.52225

Continued on next page

Table 10 – continued from previous page

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Pakistan	94	0.28324	64	0.37821	103	0.59876	47	0.28180	118	0.75	79	0.52225
Iran, Islamic Rep.	95	0.30436	91	0.55792	119	0.78099	91	0.51512	89	0.25	79	0.52225
India	96	0.31811	100	0.60655	103	0.59876	15	0.16999	118	0.75	79	0.52225
Chad	97	0.32258	111	0.79330	98	0.30069	84	0.43212	1	0	120	0.84049
Yemen	98	0.32705	97	0.59439	119	0.78099	60	0.38634	101	0.5	79	0.52225
Mali	99	0.33949	112	0.79735	1	0	114	0.97091	1	0	58	0.34502
Sierra Leone	100	0.34245	98	0.60159	1	0	110	0.79849	1	0	121	0.84424
Afghanistan	101	0.58230	110	0.71598	121	0.81777	91	0.51512	122	1	109	0.68175
Sudan	102	0.67781	106	0.67981	122	1	111	0.82273	101	0.5	122	1
Angola		NA	89	0.54344	1	0		NA	89	0.25	79	0.52225
Bosnia and Herzegovina		NA		NA	1	0	34	0.25756	1	0	1	0
Chinese Taipei		NA		NA	1	0	3	0.08757	101	0.5	1	0
Congo, Rep.		NA	101	0.62450	1	0		NA	1	0	79	0.52225
Guinea-Bissau		NA		NA		NA	107	0.75756	1	0	111	0.68473
Haiti		NA	65	0.37837	1	0	54	0.34513	1	0		NA
Israel		NA	45	0.22712	1	0		NA	1	0	1	0
Jordan		NA	85	0.51739	103	0.59876		NA	101	0.5	79	0.52225
Korea, Dem. Rep.		NA		NA	84	0.29808	91	0.51512	1	0	1	0
Lebanon		NA		NA	103	0.59876	60	0.38634	1	0	53	0.17351
Lesotho		NA	94	0.57149	84	0.29808		NA	1	0	79	0.52225
Malaysia		NA	53	0.32163	103	0.59876		NA	1	0	1	0
Occupied Palestinian Territory		NA	78	0.48607	103	0.59876		NA	1	0	66	0.34874
Oman		NA	74	0.45364	84	0.29808		NA	101	0.5	66	0.34874
Panama		NA		NA	1	0	8	0.11181	1	0	1	0

Continued on next page

Table 10 – continued from previous page

Country	SIGI		Family code		Civil liberties		Physical integrity		Son preference		Ownership rights	
	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value
Puerto Rico		NA		NA	1	0	23	0.21635	1	0		NA
Saudi Arabia		NA	74	0.45364	122	1		NA	101	0.5	79	0.52225
Serbia and Montenegro		NA		NA	1	0		NA		NA	43	0.17151
Somalia		NA		NA	103	0.59876	113	0.84213	1	0	111	0.68473
Timor-Leste		NA		NA	1	0	83	0.42755	89	0.25	79	0.52225
Turkmenistan		NA		NA	1	0	60	0.38634	1	0	79	0.52225
Uzbekistan		NA		NA	1	0	60	0.38634	1	0	1	0

Appendix 5: Regional Pattern of the Composite Index and Subindices

Table 11:

	ECA	LAC	EAP	SA	SSA	MENA	Total
SIGI							
Quintile 1	6	10	4	0	1	0	**21**
Quintile 2	6	8	5	0	0	1	**20**
Quintile 3	1	1	2	1	14	2	**21**
Quintile 4	0	0	1	2	13	4	**20**
Quintile 5	0	0	1	4	10	5	**20**
Total	**13**	**19**	**13**	**7**	**38**	**12**	**102**
Family Code							
Quintile 1	7	11	4	0	1	0	**23**
Quintile 2	5	8	6	1	0	2	**22**
Quintile 3	1	1	4	3	9	5	**23**
Quintile 4	0	0	0	0	15	7	**22**
Quintile 5	0	0	0	3	16	3	**22**
Total	**13**	**20**	**14**	**7**	**41**	**17**	**112**
Civil Liberties							
Quintile 1, 2, 3	17	22	14	0	27	3	**83**
Quintile 4	0	0	1	3	12	3	**19**
Quintile 5	0	0	2	4	3	12	**21**
Total	**17**	**22**	**17**	**7**	**42**	**18**	**123**
Physical Integrity							
Quintile 1	5	13	5	3	4	2	**32**
Quintile 2	4	4	1	0	3	2	**14**
Quintile 3	7	5	7	3	6	4	**32**
Quintile 4	0	0	3	1	13	2	**19**
Quintile 5	0	0	0	0	14	3	**17**
Total	**16**	**22**	**16**	**7**	**40**	**13**	**114**
Missing Women							
Quintile 1, 2, 3	15	21	10	1	38	3	**88**
Quintile 4	0	1	4	0	4	3	**12**
Quintile 5	1	0	3	6	1	12	**23**
Total	**16**	**22**	**17**	**7**	**43**	**18**	**123**
Ownership Rights							
Quintile 1	12	12	11	1	2	4	**42**
Quintile 2	2	4	2	0	1	1	**10**
Quintile 3	2	3	2	1	8	7	**23**
Quintile 4	1	1	2	4	18	6	**32**
Quintile 5	0	0	0	1	14	0	**15**
Total	**17**	**20**	**17**	**7**	**43**	**18**	**122**

ECA stands for Europe and Central Asia, LAC for Latin America and the Caribbean, EAP for East Asia and Pacific, SSA for Sub-Saharan Africa, and MENA for Middle East and North Africa.

Appendix 6: Comparison with other Gender-related Indices

Statistical Association between the SIGI and other Gender-related Measures

Table 12:

GDI	Kendall tau b	-0.501	Pearson Corr. Coeff.	-0.5852
Number obs. 79	p-value	0.0000	p-value	0.0000
GGI (capped)	Kendall tau b	-0.5088	Pearson Corr. Coeff.	-0.7169
Number obs. 85	p-value	0.0000	p-value	0.0000
GEM	Kendall tau b	-0.425	Pearson Corr. Coeff.	-0.7024
Number obs. 33	p-value	0.0005	p-value	0.0000
GEM (revised)	Kendall tau b	-0.4402	Pearson Corr. Coeff.	-0.7507
Number obs. 33	p-value	0.0003	p-value	0.0000
GGG	Kendall tau b	-0.4741	Pearson Corr. Coeff.	-0.7295
Number obs. 73	p-value	0.0000	p-value	0.0000
WOSOC	Kendall tau b	-0.4861	Pearson Corr. Coeff.	-0.5266
Number obs. 99	p-value	0.0000	p-value	0.0000

Data for the Gender-related development Index (GDI) and the Gender Empowerment Measure (GEM) are from United Nations Development Programme (2006) and are based on the year 2004. The Gender Gap Index (GGI) capped and the revised Gender Empowerment Measure (GEM revised) are taken from Klasen and Schüler (2009) based on the year 2004. Data for the Global Gender Gap Index (GGG) are from Hausmann et al. (2007). The Women's Social Rights Index (WOSOC) data correspond to the year 2007 and are obtained from `http://ciri.binghamton.edu/`. The p-values correspond to the null hypothesis that the SIGI and the corresponding measure are independent.

Comparison of Ranks: the SIGI and other Gender-related Indices

Table 13:

Country	SIGI	GDI	GGI (capped)	GEM	GEM (revised)	GGG	WOSOC
Paraguay	1					32	19
Croatia	2	6	16	6	7	3	19
Kazakhstan	3	18	1			10	19
Argentina	4	2	21	2	3	11	3
Costa Rica	5	7	40	3	2	8	3
Russian Federation	6	10	6	22	22	18	19
Philippines	7	22	30	10	8	1	19
El Salvador	8	29	35	13	14	20	19
Ecuador	9			14	11	17	19
Ukraine	10	19	7	23	23	25	19
Mauritius	11	12	46			44	3
Moldova	12						
Bolivia	13	35	24	19	15	41	3
Uruguay	14	5	17	15	17	39	19
Venezuela, RB	15	17	23	11	13	24	
Thailand	16	16	8	20	18	22	19
Peru	17	23	24	8	6	37	3
Colombia	18	15	11	16	16	7	3
Belarus	19	11	3			6	3
Hong Kong, China	20						
Singapore	21			1	11	38	19
Cuba	22		37			5	1
Macedonia, FYR	23	13	32	9	9	13	19
Brazil	24	14	20	20	19	36	3
Tunisia	25	26	72			55	64
Chile	26	3	44	16	20	45	3
Cambodia	27	45	10	28	26	52	3
Nicaragua	28	37	56			49	19
Trinidad and Tobago	29	9	33	4	5	19	1
Kyrgyz Republic	30	34	11			33	19
Viet Nam	31	31	2			15	19
Armenia	32	20	4			34	19
Georgia	33			24	24	30	19
Guatemala	34	39	64			58	19
Tajikistan	35	40	19			40	19
Honduras	36	38	36	12	10	31	19
Azerbaijan	37	28	4			26	19

Continued on next page

Table 13 – continued from previous page

Country	SIGI	GDI	GGI (capped)	GEM	GEM (revised)	GGG	WOSOC
Lao PDR	38	47	45				3
Mongolia	39	36	27	25	25	27	3
Dominican Republic	40	25	38			29	19
Myanmar	41		14				64
Jamaica	42	30	18			14	3
Morocco	43						19
Fiji	44						3
Sri Lanka	45	24	51	29	28	2	19
Madagascar	46	53	15			48	19
Namibia	47	43	33	5	4	9	19
Botswana	48	46	59	18	21	23	64
South Africa	49	41	42			4	19
Burundi	50	72	24				64
Albania	51						19
Senegal	52						64
Tanzania	53	66	27	7	1	12	19
Ghana	54	48	27			28	19
Indonesia	55	32	39			42	19
Eritrea	56						19
Kenya	57	57	42			43	64
Cote d'Ivoire	58	68	80				64
Syrian Arab Republic	59	33	63			56	64
Malawi	60	70	41			46	19
Mauritania	61	60	48			60	64
Swaziland	62	59	82				64
Burkina Faso	63	76	50			66	64
Bhutan	64						3
Nepal	65	51	61			70	64
Rwanda	66	63	9				3
Niger	67	79	78				19
Equatorial Guinea	68	42	62				19
Gambia, The	69					50	19
Central African Republic	70	75	67				19
Kuwait	71	1	48			51	64
Zimbabwe	72	58	57			47	19
Uganda	73	54	31			21	19
Benin	74	67	73			69	64
Algeria	75						64
Bahrain	76	4	76			64	64

Continued on next page

Table 13 – continued from previous page

Country	SIGI	GDI	GGI (capped)	GEM	GEM (revised)	GGG	WOSOC
Mozambique	77	71	47			16	64
Togo	78	61	70				64
Congo, Dem. Rep.	79	73	60				64
Papua New Guinea	80	50	22				19
Cameroon	81	55	54			65	64
Egypt, Arab Rep.	82			32	31	68	64
China	83	20	13			35	64
Gabon	84						64
Zambia	85	69	64			54	64
Nigeria	86	64	66			59	64
Liberia	87		68				19
Guinea	88	65	58				19
Ethiopia	89					62	64
Bangladesh	90	49	52	27	27	53	64
Libya	91		69				64
United Arab Emirates	92	8	74	30	32	57	64
Iraq	93		84				64
Pakistan	94	51	81	26	28	71	64
Iran, Islamic Rep.	95	27	54	31	30	67	64
India	96	44	77			63	19
Chad	97	74	75			72	64
Yemen	98	62	83	33	33	73	64
Mali	99	77	53			61	19
Sierra Leone	100	78	71				64
Afghanistan	101		85				19
Sudan	102	56	79				64
Number of obs.	102	79	85	33	33	73	99

Data for the Gender-related development Index (GDI) and the Gender Empowerment Measure (GEM) are from United Nations Development Programme (2006) and are based on the year 2004. The Gender Gap Index (GGI) capped and the revised Gender Empowerment Measure (GEM revised) are taken from Klasen and Schüler (2009) based on the year 2004. Data for the Global Gender Gap Index (GGG) are from Hausmann et al. (2007). The Women's Social Rights Index (WOSOC) data correspond to the year 2007 and are obtained from `http://ciri.binghamton.edu/`.

Appendix 7: Results from Regression Analysis

Linear Regression with Dependent Variable Global Gender Gap Index 2007

Table 14: Linear Regression with Dependent Variables GGG and Ratio GDI to HDI

	GGG coef/se	**Ratio GDI to HDI** coef/se
SIGI	-0.284***	-0.054***
	(0.089)	(0.017)
GDP	0.014*	0.004
	(0.008)	(0.003)
SA	-0.006	-0.001
	(0.032)	(0.008)
ECA	-0.012	0.007
	(0.017)	(0.005)
LAC	-0.040**	-0.000
	(0.017)	(0.005)
MENA	-0.043	0.001
	(0.028)	(0.011)
EAP	0.005	0.010**
	(0.022)	(0.005)
Muslim	-0.001	-0.002
	(0.018)	(0.006)
Christian	0.026	0.002
	(0.017)	(0.004)
constant	0.570***	0.960***
	(0.063)	(0.020)
Number of obs.	73	79
Adjusted R2	0.617	0.438
Prob F	0.000	0.000

note: *** p<0.01, ** p<0.05, * p<0.1