

Replication, Extension, and Revision of the GDI and GII

In this appendix, we provide details on the extension of the UNDP's new GDI and the GII back to 1990 and the sensitivity of the GDI index to replacement or re-estimation of some of the variables in the index, following suggestions by Klasen (2014) and Dijkstra (2002). We refer to our replicated and extended series as time consistent (TC) versions.

Construction of the GDI, TC version

The UNDP's newly introduced GDI begins in 2014, limiting the ability to do time series and panel data analysis with this index. Given that gender equality is a long-term objective and many of the index's indicators are available over a longer period, we have extended the series to provide a consistently constructed series. Dilli, et al. (2015) and Gonzales, et al. (2015) have done something similar, the former, with their own gender equality index and the latter, with the UNDP's GII. Table D1 provides information on the various indicators included in the GDI calculation. Although data for many of the indicators are available as far back as 1950, some indicators have limited data and, rather than impute data for variables other than wages, we extend the index back only to 1990.

Table D1. Overview of the Indicators Included in the GDI

Indicator	Range	Mean (s.d.)	Countries ¹	Years	Source
Life Expectancy (F)	22.7-86.7	67.33 (11.25)	230	1960-2013	World Development Indicators, World Bank
Mean Years of Schooling (F)	0-13.64	5.67 (3.46)	156	1950-2012	Barro and Lee (2014) and UNESCO Institute for Statistics
Expected Years of Schooling (F)	0.36-20.84	10.77 (4.13)	101	1970-2013	UNESCO Institute for Statistics
Ratio of Female to Male Wage	0.41-1.5	0.77 (.12)	27	1995-2011	ILO
Female Share of Economically Active Population	0.096-0.56	0.40 (.095)	184	1990-2014	ILO
GNI per capita PPP (2011 \$)	307-156,408	15,172.78 (19,408.85)	189	1980-2013	ILO
Population (F)	25.98-544,386.9	11,053.89 (41,981.62)	184	1980-2013	ILO

Sources: World Bank; Barro and Lee (2014); UNESCO; ILO; and IMF staff estimates. F refers to female.

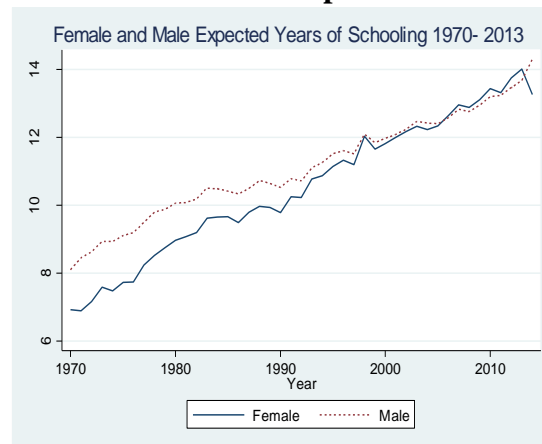
The indicator expected years of schooling has limited availability across time.² For many countries, there might only be one observation available between the years 1970-2013. To include as many countries possible, the UNDP uses observations from as far back as 2002 to

¹ The number of countries reflects the sample size for the latest listed year of data.

² Referred to by UNESCO as "School Life Expectancy: Primary to Tertiary."

calculate the 2013 GDI. This is problematic when measuring gender equality because while expected years of schooling is an indicator that may not change drastically year by year, we see that the cumulative effect over a decade shows significant change in a country. Using data for all countries with at least two observations we calculate the average growth rate in expected years of education. We find the average yearly growth rate for expected years of schooling for females is 2.3 percent and for males 1.7 percent, while the average ten year growth rate is 16.8 percent for females and 11.4 percent for males. Figure D1 depicts the gender gap in expected years of schooling. From 1970-2013 we can see the gap narrowing; the two lines converge around 2000. Therefore using data from 1990 to represent a country's gap in 2000, for example, may not be an accurate representation. This raises the question of whether it is useful to include this indicator in the index at all.

Figure D1. Female and Male Expected Years of Schooling



Sources: UNESCO; and IMF staff estimates.

The UNDP uses two measures to create the “knowledge” sub-index: expected years of schooling and mean years of schooling. UNDP argues that it is necessary to include both measures of educational attainment because they measure educational attainment in two different age groups. Expected years of schooling refers to children and their chances of receiving education, while mean years of schooling refers to the adult population who have completed formal schooling. Yet there are other gender indices in which the educational attainment sub-index was limited to data on the adult population, namely the UNDP’s Gender Inequality Index. Moreover, expected years of education and mean years of education are highly correlated (with a correlation coefficient of about 0.9); thus, the benefit of including both indicators is minimal.

Of all the indicators in the GDI, ratio of female to male wage has the poorest data coverage, with data on this indicator, in the year with highest data coverage, available for only 68 countries. Forty-four percent of the observations cover Europe and Central Asia, compared to only 0.6

percent of the observations for sub-Saharan Africa. Table D2 shows the regional distribution of data.

Table D2. Wage Data Coverage by Region

Region	Percent of Observations	Share of Population
East Asia and Pacific	15.9	33.7
Europe and Central Asia	43.9	14.1
Latin America and Caribbean	24.0	8.5
Middle East and North Africa	10.4	5.3
North America	1.8	5.4
South Asia	3.2	22.6
Sub-Saharan Africa	0.6	10.1

Sources: International Labor Organization; and IMF staff estimates.

To deal with the missing wage data, UNDP uses the global weighted wage ratio average of 0.8. This is a poor substitute for a number of reasons. First, Europe and Central Asia is over-represented in the sample based on population. Using the global average assumes that the wage ratio in countries with missing data is the same as the average in countries outside the same region and/or income group. Second, using the global average wage ratio penalizes countries with data where the reported wage ratio is below the global average.

The wage data are only available from 1995-2011. To cover the period 1990-2013 and fill in the series for countries with missing data, we impute the missing data using interpolation for those countries that have a reasonable amount of data (at least 5 years) and for those that do not, we use a regional average in place of the UN's global average ratio.

After the wage interpolation, our data cover 146 countries from 1990-2013. To construct the GDI, TC version, we followed the steps of computing the GDI, as described in the HDR 2014 technical notes.³ Our calculations vary slightly from the UNDP's because data have been updated since the UNDP's original calculations. Given the problem of missing wage data, we also try substituting the labor force participation rate as an indicator in the "standard of living" sub-index. The most significant deviation from the UNDP's GDI is that the index value for all countries is reduced. In fact, while the UNDP's GDI shows a number of countries with gender disparity in favor of women, we see that all countries now have an index value below one.

Table D3 provides a correlation matrix using Spearman's rank correlation between the UNDP's GDI; the GDI, TC version, with wages; and the TC version with the labor force participation rate

³ HDR technical notes available at http://hdr.undp.org/sites/default/files/hdr14_technical_notes.pdf

in the standard of living sub-index. It is apparent that while the two ranks are still strongly correlated they do differ slightly.

Table D3: Spearman Rank Correlation Between Different Calculations of the GDI

GDI	Original	TC Version	TC version with LFPR
Original	1		
TC version	0.96***	1	
TC version with LFPR	0.75***	0.81***	1

***, **, and * denote significance at the 1, 5, and 10 percent level

Sources: UNDP Human Development Reports; World Bank; UNESCO; ILO; and IMF staff estimates.

Construction of the GII, TC version

Our data cover 141 countries from 1990-2013. Five indicators are included in our calculation of the GII: maternal mortality ratio, adolescent fertility rate, share of female seats in national parliaments, educational attainment at secondary and tertiary levels, and labor force participation rate.

To have a complete time series we needed to interpolate and extrapolate data for years of missing data. The methods used are as follows. Data for maternal mortality ratio (MMR) are available beginning from 1990 in five-year intervals. Data for in-between years are interpolated using linear interpolation.

Data for adolescent fertility rate are from United Nations Department of Economic and Social Affairs and are available through the World Development Indicators database for most years in our time period (1990-2013). In general, data coverage for this indicator is good with a consistent time series. Countries with missing data tend to be smaller states.

The indicator, female seats in parliament, provided by the Inter-Parliamentary Union, is available beginning in 1990. However, for most countries, data from 1991- mid/late 1990s are missing. Although the gap in data for some countries is large (around eight years), the share of women in parliaments does not change drastically over a short period of time and thus we use linear interpolation to fill in years of missing data.

Educational attainment for the purposes of the GII is defined as attainment at secondary and higher education levels, or as referred to by UNESCO “population over the age of 25 with at least secondary education.” To create this variable, we use data from two sources: Barro and Lee (2014) and UNESCO Institute for Statistics.

Using the Barro and Lee data, we add the “percent of secondary schooling attained in population” to the “percent of tertiary schooling attained in population” to estimate the

population with at least secondary education. To supplement missing data, we use UNESCO's indicator "population with at least secondary education (+25)." We use linear interpolation when the missing data are between two points of available data. However, for recent years where linear interpolation is not possible, we use the most recent year of data as a substitute. For some countries data from as far back as 2010 is used in the calculation of the most recent year of the index.

Finally, data for labor force participation rate are provided by the ILO and are available through the World Bank Indicators database. In general, countries with data have a complete time series from 1990-2013. For countries where data are missing in the most recent years we use the most recent available year of data to fill in the missing data points.

To reconstruct the GII, we followed the steps as described in UNDP's Human Development Report 2014 technical notes.⁴ To compare our calculations of the GII to UNDP's calculations we calculated the Spearman's rank correlation, the most recent year of the index. We found that our calculations come very close to the UNDP's calculations with Spearman's rho of 0.99 and significant at the 1 percent level. The slight variation is due to updates in the indicators used.

⁴ See http://hdr.undp.org/sites/default/files/hdr14_technical_notes.pdf.