# Replication of States of Fragility 2015

## April 28, 2015

## Thomas Leo Scherer

## Section 0: Github contents

StatesofFragility2015post – Word document of the original write-up

StatesofFragility2015replication.R – R script used for replication

StatesofFragility2015replication.doc – This word document describing the replication

**Replication datasets**

*replicationindices* – average score for each dimension from the scaleddata, along with the countries rank from 1 to 183 with 1 being the most vulnerable.

*replicationresultscompare* – replication results with binary indicator of vulnerability in the five dimensions merged with the original report findings.

*reportedfragility* – OECD’s original Venn diagram of fragility in dataset form.

*scaleddata* – The 15 inputs after being scaled from 0 to 100.

**Raw Indicators – see details in next section**

*Birth\_registration\_Oct2014\_135 GHE\_DALY\_2012\_country*

*healthcarecapabilities WEOOct2014all*

*wgiestimates wgiestimates2012*

*worldriskindex*

## Section 1: Data Inputs

*Note: for each dataset I use the R package* [*countrycode*](https://github.com/vincentarelbundock/countrycode) *to get consistent country names.*

**Violence: UCDP battle deaths**

*Uppsala University (2014), Dataset, v.5-2014, 1989-2013, Conflict Data Program, Uppsala University,*

*Uppsala, Sweden, available at: www.pcr.uu.se/research/ucdp/datasets/ucdp\_battle-related\_deaths\_*

*dataset (accessed 1 August 2014).*

Loaded directly from <http://www.pcr.uu.se/digitalAssets/124/124934_1ucdpbattle-relateddeathsdatasetv.5-2014dyadic.rdata>. For each state I add the best estimate of battle deaths for each conflict where the state was listed as Side A (except for Syria, where only a “low estimate” is given for 2013). I change the name of “Yemen Arab Republic” to “Yemen”.

**Violence: Interpersonal Injuries**

*WHO (2014), “Health statistics and information systems: Estimates for 2000-12, Disease burden”, in*

*Global Health Estimates, World Health Organization, Geneva, available at: www.who.int/healthinfo/*

*global\_burden\_disease/estimates/en/index2.html (accessed 1 August 2014).*

Downloaded from <http://www.who.int/entity/healthinfo/global_burden_disease/GHE_DALY_2012_country.xls>, saved as GHE\_DALY\_2012\_country. I pull out sheet 2, row 157 (becomes row 156 when loaded into R). I change “Democratic People's Republic of Korea” to “North Korea”.

**Violence: Political Instability**

*World Bank (2014a), Worldwide Governance Indicators 1996-2013, World Bank, Washington, DC, http://info.worldbank.org/governance/wgi/index.aspx (accessed 1 August 2014).*

Downloaded from <http://info.worldbank.org/governance/wgi/wgidataset.xlsx>. Columns “CC” and “CI” give the best estimates for 2012 and 2013. These columns were pulled along with the country names column and saved as the csv files wgiestimates and wgiestimates2012.

**Justice: Birth Registration**

*UNICEF (2014), “UNICEF data: Monitoring the situation of women and children”, available at: http://*

*data.unicef.org/child-protection/birth-registration.*

Downloaded from <http://data.unicef.org/download.php?file=Birth_registration_Oct2014_135.xlsx&type=topics>.

Download saved as Birth\_registration\_Oct2014\_135. The data currently available was updated November 2014, and no archive is readily available. 6 countries use data from after 2013 or 2012-2013: Madagascar, Pakistan, Senegal, Georgia, Nigeria, Vanuatu. I change “Democratic People's Republic of Korea” to “North Korea”.

**Justice: Control of Corruption**

**Justice: Rule of Law**

**Institutions: Government Effectiveness**

**Institutions: Regulatory Quality**

**Institutions: Voice and Accountability**

These are all from the same excel file as Violence: Political Instability on different sheets. The same procedure is used for each.

**Economic Foundations: Doing Business Index:**

*World Bank (2013), Doing Business 2013,World Bank,Washington, DC, available at: www.doingbusiness.org/rankings.*

Downloaded using the R package [WDI](https://github.com/vincentarelbundock/WDI) for indicator IC.BUS.EASE.XQ. To remove the regional entries, I drop rows that are unmatched after using CountryCode. Two entries for “Sub-Saharan Africa” must be dropped explicitly since CountryCode translates them to “Western Sahara”. Both 2013 and 2014 are now available. There is no data before 2013.

**Economic Foundations: Education Years**

*Barro, R. and J.-W. Lee (2010), “A new data set of educational attainment in the world, 1950-2010”,*

*Journal of Development Economics, Vol. 104, pp. 184-198, accessed at:* [*www.barrolee.com*](http://www.barrolee.com)*.*

Downloaded using the R package [WDI](https://github.com/vincentarelbundock/WDI) for indicator BAR.PRM.SCHL.15UP. Data is available every five years, including 2010.

**Economic Foundations: GDP Growth** **– ten year average**

*IMF (2014), World Economic Outlook Database, Washington, DC, available at: www.imf.org/external/pubs/ft/weo/2014/02/weodata/index.aspx.*

Downloaded from <http://www.imf.org/external/pubs/ft/weo/2014/02/weodata/WEOOct2014all.xls>, saved as WEOOct2014all. I use annual percent change in constant price GDP (WEO Subject Code NGDP\_RPCH). The dataset contains data up to 2013 for 116 countries, but some stop sooner. Djibouti’s last year for actual data is from 1999; no last year is given for Bhutan. For all cases, estimates are made up through 2019.

For each country, take the ten years before estimates begin and average the GDP Growth for those years, dropping any missing years. For Bhutan I assume the last year with data is 2012.

**Resilience: Income Inequality**- last year of gini measure ranges widely

*International Futures (IFs) model (2014), Version 7.08, Frederick S. Pardee Center for International*

*Futures, Josef Korbel School of International Studies, University of Denver, Denver, Colorado,*

*www.ifs.du.edu*

Downloaded using the R package [WDI](https://github.com/vincentarelbundock/WDI) for indicator SI.POV.GINI. I take the most recent value for each country.

**Resilience: Healthcare Capabilities** – made myself from a map

*US National Intelligence Council (2008), “Strategic implications of global health”, www.dni.gov/files/*

*documents/Special%20Report\_ICA%20Global%20Health%202008.pdf; National Center for Medical*

*Intelligence, Health-Care Capabilities, Data are available in form of a map at: www.dni.gov/files/*

*documents/Special%20Report\_ICA%20Global%20Health%202008%20foldout.pdf.*

I record the scores for each country based on their color on the map at [www.dni.gov/files/documents/Special%20Report\_ICA%20Global%20Health%202008%20foldout.pdf](http://www.dni.gov/files/documents/Special%20Report_ICA%20Global%20Health%202008%20foldout.pdf).

Saved as healthcarecapabilities.

**Resilience: Vulnerability**

*UNU-EHS (2014), World Risk Report 2014, United Nations University Institute for Environment and*

*Human Security, Bonn, Germany, available at: www.ehs.unu.edu/article/read/world-risk-report-2014.*

The WorldRiskIndex is taken from page 63 of <http://www.ehs.unu.edu/file/get/11895.pdf>. Data is saved as worldriskindex.

## Section 2: Methodology

*For each goal proposed by the OWG, the authors designed a separate index, ranked all states and economies for which data were available (2012 data or most recent year), and identified the 50 most vulnerable ones. For each goal proposed by the OWG, the authors designed a separate index. Each index is based on three datasets for each cluster (Table A.2 summarizes the indicators and goal areas linked to each ranking). Each dataset was then rescaled to a scale from 0 to 100. The index is the average of these three numbers. Using this index, country lists were sorted from worst to best performing, and the 50 lowest performing countries and economies for each index are considered “vulnerable” or “challenged” in that area. (OECD 2015, pp 103-104)*

For each of the 15 datasets, I rescale the data by subtracting the minimum value, dividing by the resulting maximum value, and multiplying by 100. The formula looks like this:

Xscaled = 100 \* (Xoriginal – Xmin) / (Xmax – Xmin)

For each of the five clusters, I average the three scaled values. I then identify the 50 lowest performing countries and check which countries are “vulnerable” in two or more clusters.

There is no discussion of missing values. I continue through the scaling and averaging simply disregarding missing values. I drop all countries that are missing more than 10 of the 15 values. Some countries end up unranked for some clusters where they are missing all three inputs, but this seems limited to small island countries.

There is also no discussion about what does and does not constitute a country. Small island countries Kiribati and São Tomé and Príncipe are included as is state de jure Palestine under the name “West Bank and Gaza Strip.” In my replication, I drop all those countries that could be considering territories or have a smaller population than Kiribati since it brings me closer to the OECD’s results. The countries dropped are listed below:

Small Countries

American Samoa Andorra Anguilla

Antigua And Barbuda Cayman Islands Dominica

Greenland Grenada Jersey

Liechtenstein Marshall Islands Micronesia, Federated States of

Monaco Nauru Netherlands Antilles

Palau Saint Kitts And Nevis Saint Lucia

San Marino Seychelles Tonga

Tuvalu Vanuatu Virgin Islands, U.S.

Territories

American Samoa Anguilla Aruba

Cayman Islands Greenland Guam

Hong Kong Jersey Macao

Martinique Netherlands Antilles Puerto Rico

Reunion Taiwan, Province Of China Virgin Islands, U.S."

The directions to use “2012 or most recent year” can be interpreted in a few ways. The most obvious is using Price-is-Right-Rules - closest to 2012 without going over. Another way is to use the most recent year available in the cited dataset. The following datasets as cited have data available after 2012:

Battle Deaths: 2012 and 2013 (2014 is cited)

WGIestimates: 2012 and 2013 (2014 is cited)

Doing Business: 2013 and 2014 (2013 is cited)

GDP Growth: 2012 or 2013 (2014 citation) (year listed is possible final year for 10-year-average)

I find that using the most recent data results in fewer edits, so I use those results for my analysis.

## Section 3: Results

Specifications:

* Most recent available year
* Dropping small countries and territories
* Dropping countries with more than 10 missing inputs

Out of 70 states shown on the fragility clusters in the OECD report, my results match perfectly for 24 countries.

|  |  |
| --- | --- |
| States Added: 6 | States Removed: 10 |
| Brazil  China  Egypt  Gabon  North Korea  Philippines | Honduras  India  Kiribati  Lao People's Democratic Republic Lesotho  Panama  Paraguay  Rwanda  Sao Tome and Principe  Sierra Leone |