



Human Rights Measurement Initiative

Methodology Handbook

6 June 2019

Document information

Key author contact details

Anne-Marie Brook

Motu Economic and Public Policy Research

anne-marie.brook@motu.org.nz

K Chad Clay

University of Georgia

kcclay@uga.edu

Susan Randolph

University of Connecticut

susan.m.randolph@gmail.com

Acknowledgements

We are enormously indebted to the hundreds of people who have helped make HRMI’s dataset come to life. We would like to particularly acknowledge: those human rights experts, academics, and other supporters who participated in our 2015, 2017 and 2018 co-design workshops; our HRMI Ambassadors, who have helped connect us with potential survey respondents in their countries of interest; the dozens of human rights experts who have contributed their expertise by helping to design and “test” our expert survey, or our data visualisations, or who were willing to give up their precious time to share with us their knowledge about human rights violations in their country – via the survey – often without any prior introduction to HRMI; the volunteer translators who have helped make our survey and website material accessible in nine languages (for the survey) and four languages (for the website); our website designers and developers, who have done a wonderful job with very limited resources; our colleagues and support staff; and our financial donors, including The Open Society Foundations who we are very grateful could also see the enormous value that good human rights data can bring. Last, but not least, we would like to thank our families who have seen far less of us than we would like.

Supported in part by a grant from the Open Society Foundations.

Disclaimer

The findings and opinions expressed in this document are those of the authors.

Copyright

This document is licensed under a Creative Commons Attribution 4.0 International copyright licence.

This means anyone may copy this document in whole or in part as long as they attribute the Human

Rights Measurement Initiative as the creator and link back to the HRMI website,
<https://humanrightsmmeasurement.org/>.

Version: 2019.06.06

Table of Contents

Document information	i
Introduction	5
1 HRMI Civil and Political Rights Metrics Methodology – Executive Summary	6
1.1 What are civil and political rights?	6
1.2 How does HRMI measure civil and political rights?	7
1.3 What questions does the expert survey ask?	7
1.4 Who can be an expert respondent?	9
1.5 How are survey responses converted into HRMI metrics?	10
1.6 How do HRMI’s metrics differ from other measures of civil and political rights?	11
1.7 How does this methodology differ from the way HRMI measures economic and social rights?	11
2 HRMI Political and Civil Rights Metrics 2018 Technical Note	13
2.1 Abstract	13
2.2 Introduction	13
2.3 What do existing measures of civil and political rights miss?	15
2.4 HRMI’s approach to civil and political rights measurement	21
2.5 Presentation of pilot data	Error! Bookmark not defined.
2.6 Conclusion	30
2.7 Appendix: pairwise comparisons	Error! Bookmark not defined.
3 HRMI Economic and Social Rights Metrics Methodology – Executive Summary	34
3.1 What are economic and social rights?	34
3.2 How does HRMI measure economic and social rights?	34
3.3 How is this different from the way HRMI measures civil and political rights?	35
3.4 How does HRMI’s economic and social rights methodology work?	35
3.5 What do HRMI’s economic and social rights scores show, exactly?	36
3.6 What are HRMI’s two different assessment standards?	37
3.7 How is HRMI’s economic and social rights metric constructed?	37
3.8 How does HRMI choose which indicators to use?	38
3.9 What is HRMI’s achievement possibilities frontier?	39
4 HRMI Economic and Social Rights Metrics 2018 Technical Note	43
4.1 Overview	44
4.2 Sources and definitions of rights and obligations	46
4.3 Measuring economic and social rights enjoyment and state resources	48
4.4 Calculating indicator scores by benchmarking a country’s obligations of progressive realisation	52
4.5 Right scores	60
4.6 References	63
4.7 Appendix	65

Table of Figures

Figure 1: Estimates of Physical Integrity Rights Performance for Countries in Pilot	Error! Bookmark not defined.
Figure 2: Estimates of Empowerment Rights Performance for Countries in Pilot	Error! Bookmark not defined.
Figure 3: Pairwise Comparisons, Torture and Ill-Treatment	Error! Bookmark not defined.
Figure 4: Range of Torture in Four Pilot Countries	Error! Bookmark not defined.
Figure 5: Attributes of Those At-Risk for Torture in Australia	Error! Bookmark not defined.
Figure 6: Comparison with V-Dem Indicators	Error! Bookmark not defined.
Figure 7: Achievement Possibilities Frontier for “Percentage of Children Not Stunted”	41
Figure 8: Rescaling the indicator scores	56
Figure 9: Oman’s resources exceed the level needed to eliminate child stunting.	58
Figure 10: Penalty for different Y/Yp values	60
Appendix Figure 1: Pairwise Comparisons, Political/Arbitrary Arrest and Imprisonment	Error! Bookmark not defined.
Appendix Figure 2: Pairwise Comparisons, Disappearance	Error! Bookmark not defined.
Appendix Figure 3: Pairwise Comparisons, Torture	Error! Bookmark not defined.
Appendix Figure 4: Pairwise Comparisons, Extrajudicial Execution	Error! Bookmark not defined.
Appendix Figure 5: Pairwise Comparisons, Death Penalty Execution	Error! Bookmark not defined.
Appendix Figure 6: Pairwise Comparisons, Assembly	Error! Bookmark not defined.
Appendix Figure 7: Pairwise Comparisons, Association	Error! Bookmark not defined.
Appendix Figure 8: Pairwise Comparisons, Assembly and Association	Error! Bookmark not defined.
Appendix Figure 9: Pairwise Comparisons, Opinion and Expression	Error! Bookmark not defined.
Appendix Figure 10: Pairwise Comparisons, Political Participation	Error! Bookmark not defined.
Table 1: Example data for BAM model	29
Table 2: Rights enjoyment indicator sets used in HRMI economic and social rights metrics	39
Table 3: Rights enjoyment indicators used to construct HRMI’s ESR metrics	52
Table 4. Sub-scores Comprising HRMI Right Scores by Assessment Standard	62
Appendix Table A: Indicator Definitions	65
Appendix Table B: Countries Defining the Frontier	Error! Bookmark not defined.
Appendix Table C: Frontier Equations, Peak Indicator Values, Income level at Peak Indicator Value, Minimum Value	72

Introduction

Human rights are those rights you have simply because you are human. Such rights are “inherent in our nature” and “allow us to fully develop and use our human qualities, our intelligence, our talents and our conscience and to satisfy our spiritual and other needs” (United Nations 1987, 4).

The Human Rights Measurement Initiative was formed to produce a comprehensive suite of metrics that cover the rights embodied in international law, particularly the collection of international treaties known as the International Bill of Human Rights. These are internationally recognised human rights acknowledged by all United Nations member states.

Why? Because we believe that for human rights to improve, they need to be measured. High-quality data will create an opportunity for tremendous advances in our knowledge and understanding about how to encourage much greater respect for human rights around the world. We encourage you to contribute to building that knowledge.

Our initial 12 metrics can be grouped into two broad categories: seven civil and political rights and five economic and social rights. Each category has its own methodology and this document details the methodology behind each measurement.

We also encourage you to use our data portal, which you can access from our website <https://humanrightsmeasurement.org/>. With the release of this interactive data tool, you can explore not only our civil and political rights metrics for 19 countries, but also our economic and social rights metrics for 120-180 countries (depending on the right). For each country you'll be able to see its relative strengths and weaknesses, and you'll also be able to explore performance on a particular right within different regions of the world. For the 19 countries we have survey data for, you will also see information on which population sub-groups are considered to be particularly at risk of abuses of each of the civil and political rights and each of the economic and social rights.

1 HRMI Civil and Political Rights Metrics Methodology – Executive Summary

This is a brief explanation of how we constructed the Human Rights Measurement Initiative (HRMI)'s civil and political rights metrics. This is a methodology developed specifically for HRMI by researchers at the University of Georgia and Motu Economic and Public Policy Research. For more in-depth information, please see Section 2.

1.1 What are civil and political rights?

The International Covenant on Civil and Political Rights (ICCPR) is a treaty adopted by the United Nations in 1966 and agreed to subsequently by 169 countries that sets out a list of civil and political rights that we are all entitled to simply by virtue of being human. Civil and political human rights ensure your ability to live, and to engage in religious, political, intellectual, or other activities free from coercion, abuse, or discrimination. HRMI's metrics cover the following seven rights, each listed together with reference to the relevant article in the ICCPR or other core UN treaties further elaborating those rights, such as the International Convention for the Protection of all Persons from Enforced Disappearance and the Convention against Torture:

- the right to be free from torture and ill-treatment (Article 7 and the Convention against Torture),
- the right to be free from execution (Article 6 and the Second Optional Protocol to the ICCPR),
- the right to be free from arbitrary or political arrest and detention (Articles 2, 9, 11, 18, 19, 21, 22, and 26),
- the right to be free from disappearance (Articles 9 and 10, and the Convention for the Protection of all Persons from Enforced Disappearance),
- the right to political participation (Article 25),
- the right to opinion and expression (Article 19), and
- the rights to assembly (Article 21) and association (Article 22).

We also produce data on two overarching civil and political rights of which each of the rights above is part: the right to safety from the state (i.e. physical integrity) and the right to empowerment (i.e. the ability to take an active role in society without fear of reprisal.)

Over time we aim to become more comprehensive by producing metrics that cover the full range of rights embodied in international law.

1.2 How does HRMI measure civil and political rights?

Obtaining reliable, unbiased, and comprehensive information is perhaps the most serious impediment to the collection of quantitative civil and political rights data. When violations by government agents are reported, states often attempt to frame the abuse as either necessary or carried out without state permission. Many violations of civil and political rights take place in secret, with the violator seeking to conceal their actions entirely, and the degree to which violators conceal their complicity only serves to exacerbate the problems.

Because objective statistics on levels of respect for these human rights are either unavailable or unreliable, HRMI collects information on the scope and intensity of abuse using an expert survey approach and converts it into metrics using Bayesian statistical techniques. HRMI also collects information from survey respondents about the people who are most at risk for violations or restrictions of their rights. Overall, the advantages of using this approach are that it allows us to:

- Directly collect previously inaccessible information from human rights researchers and practitioners (in their own language wherever possible) who are actively gathering information and monitoring human rights issues in each country.
- Produce not only central estimates of the intensity of each type of abuse in each country, but also uncertainty bands around those central estimates. This results in much more accurate and honest reporting of the level of uncertainty with regard to the intensity of abuses.
- Collect data not only on the scope and intensity of abuse, but on the range of abuse as well, i.e. information on which groups of people are particularly vulnerable to each type of abuse within each country.

So far this approach has only been used twice. In our 2017 pilot, we rolled out our expert survey to human rights experts in the following 13 countries: Angola, Australia, Brazil, Fiji, Kazakhstan, Kyrgyzstan, Liberia, Mexico, Mozambique, Nepal, New Zealand, Saudi Arabia, and the United Kingdom. For the 2019 HRMI expert survey, we added the Democratic Republic of Congo, Jordan, South Korea, the United States, and Vietnam. We expect that the survey will now be conducted annually and, over time, will expand to cover most countries in the world.

1.3 What questions does the expert survey ask?

For each of the seven civil and political human rights we measure, the expert survey includes:

- A definition of the human right, taken from international law and its interpretation by the appropriate treaty bodies at the United Nations.

- Questions about the intensity (or frequency) of violations by government agents over the last year and the year previous. For example, the intensity question about acts of torture or ill-treatment in 2018 is shown below.

In 2018, how often did **government agents**, such as soldiers, police officers, and others acting on behalf of the state, commit acts of **torture or ill-treatment**?

Never		Rarely		Sometimes		Often		Very Often		Constantly
0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

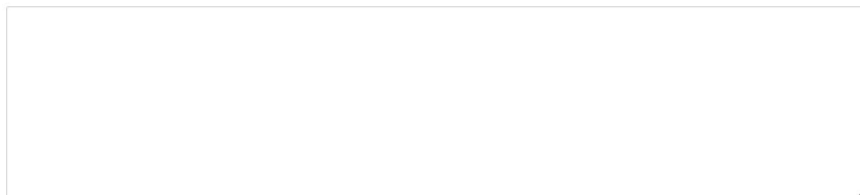
- Questions about the range of respect for the rights being discussed.
 - Respondents are asked which people were particularly at risk for violations of, or restrictions on, the right in question. Respondents could select from 32 identifiers specified in the survey or provide us with other potential identifiers, as shown below.

In 2018, which groups of people, if any, were especially vulnerable to **torture and ill-treatment** by **government agents**, such as soldiers, police officers, and other state-sanctioned actors? (**Select all that apply.**)

<input type="checkbox"/> No one was at noticeable risk	<input type="checkbox"/> People with particular political affiliations or beliefs
<input type="checkbox"/> All people were at noticeable risk	<input type="checkbox"/> Journalists
<input type="checkbox"/> Indigenous people	<input type="checkbox"/> Human rights advocates
<input type="checkbox"/> People of particular ethnicities	<input type="checkbox"/> Members of labour unions
<input type="checkbox"/> People of particular races	<input type="checkbox"/> People who protest or engage in non-violent political activity
<input type="checkbox"/> People from particular cultural backgrounds	<input type="checkbox"/> People engaged in or suspected of political violence
<input type="checkbox"/> People with particular religious beliefs or practices	<input type="checkbox"/> People engaged in or suspected of terrorism
<input type="checkbox"/> People of particular nationalities	<input type="checkbox"/> Detainees or those accused of crimes
<input type="checkbox"/> People in particular geographic locations	<input type="checkbox"/> Immigrants
<input type="checkbox"/> Foreign nationals outside of the state's territory	<input type="checkbox"/> Refugees or asylum seekers
<input type="checkbox"/> Women and/or girls	<input type="checkbox"/> People with low social or economic status
<input type="checkbox"/> LGBTQIA+ people	<input type="checkbox"/> People who are homeless
<input type="checkbox"/> Children	<input type="checkbox"/> People with less education
<input type="checkbox"/> Older people	<input type="checkbox"/> Academics
<input type="checkbox"/> People with disabilities	<input type="checkbox"/> Professionals (e.g. doctors, lawyers, teachers)
<input type="checkbox"/> People with specific medical conditions	<input type="checkbox"/> Other (Please Specify)
	<input type="text"/>

- Then, the survey provides an open field space for respondents to provide any more specific information.

In the previous question, you were asked to list whether certain types of people were especially vulnerable to **torture and ill-treatment** by **government agents** in 2018. Briefly, please give the names (if any) of the specific identities, affiliations, groups, activities, locations, or other attributes that were likely to make one especially vulnerable to **torture and ill-treatment** by **government agents**. (Please do NOT include the names of individual persons.)



Another important part of the survey included a number of *anchoring vignettes*, in which respondents were asked to score the frequency of abuses in three described hypothetical countries. Responses to these hypotheticals were used to account for differences in the interpretation of the 11-point intensity scale and contribute meaningfully to the final intensity scores produced for each country.

You can read the full expert survey questionnaire used in our pilot study [here](#). Note that this is a link to a preview of the survey only, and any responses you make will not be collected.

Looking ahead, it is likely that the survey will be modified somewhat, to take on board feedback, in future iterations. But the overall approach will most likely remain very similar.

1.4 Who can be an expert respondent?

Thus far, we have focused primarily on human rights practitioners directly monitoring the civil and political rights situation in each country. These experts are often working for an international or domestic non-governmental organisation or a civil society organisation. However, we also allowed for participation by human rights lawyers, journalists covering human rights issues, and staff working for national human rights institutions if that institution has been given A-level accreditation, showing that it is rated as fully compliant with the Paris Principles. Many of our respondents serve in several of these roles. As a result, 69% of the respondents to the 2019 HRMI Expert Survey work as a human rights practitioner or advocate, 10% work as journalists, 32% work as lawyers, and 8% only work with an NHRI or in another related field.

Wherever possible we rely on respondents who are located within the country on which they provide information. In cases of more closed and repressive countries, it has been necessary to rely on a higher proportion of respondents that are based outside of the country of interest. The 2019 survey was available to take in nine languages (Arabic, English, French, Korean, Nepali, Portuguese, Russian, Spanish, and Vietnamese) ensuring that it was accessible to as many human rights experts in our sample

as possible. This approach ensures that our expert survey is serving as a tool that gives a voice to experts located in countries around the world, to share their knowledge with the outside world in the form of quantitative metrics of civil and political rights.

This is especially valuable for human rights experts from outside of the oft-overrepresented “Western” and high income countries. Our main goal is to collect information from respondents who are first points of contact for human rights information in the country of interest and who often have access to primary sources. As such, we did not invite academics to be respondents, as academics are rarely involved in the collection of primary information and tend to rely more heavily on secondary sources. Staff at government-organised NGOs and government officials outside of A-level national human rights institutions were also excluded.

1.5 How are survey responses converted into HRMI metrics?

The statistical model we employ to convert responses to our questions about intensity of abuse into HRMI metrics is a Bayesian variant of the common factor model. Developed to study unobservable factors such as knowledge, intelligence, and personality, this approach allows us to estimate unobserved traits (in this case the level of respect for a specific human right) for individual countries, from a set of observed outcomes (in our case the responses to our survey questions) that were caused by that trait. We use this approach for three main reasons.

First, it allows us to derive sensible results from quite small sample sizes. The number of fully completed survey responses that were used to calculate the civil and political rights scores in 2019 ranged between six and 19 per country with an average number of respondents of a little more than 11. It is important to use a methodology that works with small sample sizes because the number of human rights experts in some countries is quite small, and it would be unrealistic to expect all of them to complete our survey every time we conduct it. Because our models are Bayesian, they produce a central estimate of the score for each country along with an estimate of uncertainty, around each score. A higher level of uncertainty (larger uncertainty band) results when there is more variance among survey respondents’ scores on a particular right, and/or when the number of survey respondents is smaller.

Second, this approach enables us to place each country on a common scale, even though different survey respondents may interpret the numeric values on the scale differently. For example, respondent A may give a score of 6/10, while respondent B gives the same country a score of 4/10 even if the two respondents have the same set of knowledge about what is going on in that country, simply because they interpret the scale differently from one another. Our methodology allows us to correct for that by using their responses to the questions surrounding the anchoring vignettes mentioned above.

Third, and related, it allows us to correct for any country-specific differences in interpretation of the scales. For example, if survey respondents in country X have become accustomed to a particular

intensity of abuse, it is possible they could see it as “more normal” than respondents in country Y. In this case and the one above, responses to our questions about the hypothetical countries are used as “bridging observations” to correct for any such bias and create a scale that is cross-nationally comparable.

1.6 How do HRMI’s metrics differ from other measures of civil and political rights?

There are three important differences between our measures and existing efforts. Each of these represents improvements over current practices.

First, previous efforts have either relied on reports by governments and non-governmental organisations intended for public consumption¹ (e.g. CIRI, PTS, ITT), or on surveys of academics (VDEM). By contrast, our source of information is a survey of human rights practitioners, primarily located in the country in question. This is likely to be a better source of information because it is closer to primary sources.

Second, our measures cover the following two aspects of human rights that have not previously been measured by cross-national human rights data projects: arbitrary/unlawful arrests unrelated to political activity, and the prevalence of death penalty executions.

Third, our expert survey collects information on all three of the following dimensions of rights abuse by governments (by contrast, previous efforts to measure civil and political rights have tended to focus most on intensity, with relatively limited scope):

- **Scope**, or the type of abuse the violator has engaged in. For instance, have the violators tortured political opponents, arrested them, or kept them from participating in elections? Have they done one of these things, two, or all of them?
- **Intensity**, or the frequency of the type of abuse. For example, did the violator arbitrarily imprison one or two people or hundreds?
- **Range**, or the portion of the population targeted for abuse. Did the violator focus their abuses on political opponents, on accused criminals, or on discriminated groups or classes? Or, alternatively, was the abuse indiscriminate, placing all people at risk?

1.7 How does this methodology differ from the way HRMI measures economic and social rights?

HRMI measures these two groups of rights quite differently as is consistent with state obligations under international law. Under international law, the state must immediately and completely respect, protect,

¹ Such reports include the U.S. State Department’s Country Reports on Human Rights Practices, Amnesty International’s Annual Report, and Human Rights Watch’s World Report.

and fulfil all rights listed in the International Covenant on Civil and Political Rights, while the rights listed in the International Covenant on Economic, Social, and Cultural Rights are to be progressively realised using the maximum of available resources. HRMI measures economic and social rights in two benchmarks. First, we measure how well a country is performing relative to what is reasonably feasible given its current resources, our income adjusted benchmark. Second, to capture the state's long-term obligation to fulfil the substantive rights, we also measure how well the country does relative to the best performance of countries at any income level, our global best benchmark. By contrast, our civil and political rights metrics are not adjusted to account for the resources available to a country.

A second important difference is that HRMI's civil and political rights metrics are calculated using surveys of human rights experts in each country, whereas our economic and social rights metrics are calculated from internationally comparable, publicly accessible statistical data published by national and international bodies. However, starting with the 2019 HRMI Expert Survey, we are collecting data about which people are particularly at risk to lack enjoyment of their economic and social rights, using the same kind of questions that we are using for civil and political rights. These data are available for the same 19 countries for which we currently have civil and political rights data.

2 HRMI Political and Civil Rights Metrics 2019 Technical Note

K. Chad Clay², Ryan Bakker³, Anne-Marie Brook⁴, Daniel W. Hill, Jr.⁵, and Amanda Murdie⁶

2.1 Abstract

This section details the methodology HRMI uses to measure civil and political rights violations in a sample of 19 diverse countries. In doing so, we discuss the problems present in previous attempts to measure civil and political rights cross-nationally and argue that our approach overcomes many of those problems. Using an expert survey that draws on the knowledge of human rights researchers, advocates, lawyers, journalists, and others responsible for directly monitoring the human rights situation in countries worldwide, we present measures of the intensity and distribution of respect for seven separate areas of civil and political rights, as well as two overarching categories of rights.

2.2 Introduction

Why is it difficult to obtain objective counts of the number of civil and political rights violations that occur in the world? There are several answers.

First, governments often frame and contest reporting on abuses, arguing that such acts were in fact necessary. For example, in 2015, there were reports that those suspected of terrorism and other criminal activity were being targeted and killed by Egyptian police during security raids. The Egyptian “Ministry of Interior claimed the suspects had been killed after opening fire on police officers” (USDS, 2016). However, human rights advocates argued that many of these were actually extrajudicial executions, evidenced by signs of torture on the victims’ bodies. Overall, precise numbers were obscured.

Second, government agents often attempt to engage in violations in secret, as occurred in Bangladesh in 2015, when “members of security forces in plain clothes arrested dozens of people and later denied knowledge of their whereabouts” (Amnesty International, 2016, 83). Some of the missing people were later found dead, others imprisoned, but the fates of many remain unknown.

Third, many abuses are never reported at all, or if they are reported, they never make their way into international, national, or even local media reporting. In this environment, the level of government respect

² Assistant Professor, Department of International Affairs, University of Georgia; Co-founder & CPR metrics lead, HRMI. email: kcclay@uga.edu.

³ Associate Professor, Department of Political Science, University of Georgia; CPR metrics team, HRMI. email: rbakker@uga.edu.

⁴ Co-founder & Development lead, HRMI, Motu Economic & Public Policy Research. email: anne-marie.brook@motu.org.nz.

⁵ Assistant Professor, Department of International Affairs, University of Georgia; CPR metrics team, HRMI. email: dwhill@uga.edu.

⁶ Professor, Department of International Affairs, University of Georgia; CPR metrics team, HRMI. email: murdie@uga.edu.

for civil and political rights in every country around the world is not directly observable, and producing a single, objective, unbiased count of events is impossible.

Many previous human rights data projects have attempted to mitigate these problems in human rights reporting by combining a reliance on the public documentation produced by governments and international non-governmental organisations (INGOs) with highly replicable, standards-based procedures with a great deal of success. While these approaches have helped to reduce the measurement problems caused by the weaknesses present in their information sources, those weaknesses remain. Over the years, as we have discussed existing human rights data with human rights advocates and researchers in human rights non-governmental organisations (HROs) around the world, we have heard time after time about the problems that come with relying on public reports for the purposes of measurement. While the information in the public documentation produced by such organisations is highly credible and highly unlikely to contain information on events that did not actually happen (Hill, Moore, and Mukherjee, 2013), it is also subject to political, legal, and resource constraints. This means that many known human rights violations go unreported. Further, this problem is more severe in some places than others, yielding much less information on some locations. As a result, the allegations of abuse in such reports represent a biased undercount of the level of abuse in countries worldwide (Conrad and Moore, 2011; Conrad, Haglund, and Moore, 2014). While ordered scales can serve to reduce this problem, they cannot eliminate it entirely. As a result, our conversations with human rights advocates, researchers, and others working with HROs worldwide have often ended with some variant of the same simple question: “Why not just ask us for the information directly?”

The Human Rights Measurement Initiative’s (HRMI) approach to measuring civil and political rights takes this question seriously, basing its data on information supplied by human rights experts around the world who are directly responsible for monitoring human rights practices in their particular countries or regions. In this section, we describe our methodological approach to measuring civil and political rights practices and compare it to existing efforts. Below we 1) discuss how our conceptual and operational approach differs from previous projects, 2) describe the ways we combine numeric survey responses from human rights experts into data for each country in our sample, and 3) provide some summary statistics from our survey and data.⁷

⁷ To obtain more visualizations of the data themselves or to download the data for your own use, please visit our website at <https://data.humanrightsmmeasurement.org>. For more information on the validity of our approach compared to other attempts to measure civil and political rights, please view the previous version of this paper based on the 2018 pilot data: K Chad Clay, Ryan Bakker, Anne-Marie Brook, Daniel W. Hill, Jr., and Amanda Murdie. 2018. "HRMI Civil and Political Rights Metrics: 2018 Technical Note" Motu Working Paper 18-05. Motu Economic and Public Policy Research. Wellington, New Zealand. <https://motu.nz/our-work/human-rights/hrmi-political-and-civil-rights-metrics-2018-technical-note/>.

2.3 What do existing measures of civil and political rights miss?

Why do we need more cross-national measures of civil and political rights? There are several existing data sets that, in various ways, attempt to measure at least some of these rights from different angles, e.g. Cingranelli, Richards and Clay (2014a); Conrad and Moore (2010); Gibney et al. (2015); Coppedge et al. (2017). If there are so many projects attempting to measure the same things, what could they possibly be missing?

According to Goldstein (1986), anyone that attempts to generate quantitative data on human rights will face challenges associated with definitions, data reliability, and data interpretation. With regard to definitions, most projects have decided to hew closely to the definitions of various rights found in international human rights treaties, often aided by the various treaty bodies overseeing those documents, and on this front, HRMI is no exception. However, when it comes to the problems of data reliability and interpretation, we take a significantly different tack. Over the course of this section, we discuss the approaches taken by previous attempts to measure civil and political rights cross-nationally. We then demonstrate how these different approaches to human rights information and its interpretation are likely to lead to biased, unreliable results. HRMI avoids many of the shortcomings of these existing approaches and provides more detailed, contextualised information on the distribution of abuse and those who are most affected by that abuse than any previous cross-national data project has been able to do.

2.3.1 Existing measures of civil and political rights

There are several existing measures of respect for civil and political rights, often particularly focusing on the subset of those rights known as “physical integrity rights.”⁸ Among the most widely used are the Political Terror Scale (PTS) (Gibney et al., 2015) and the indices created by the CIRI Human Rights Data Project (Cingranelli, Richards and Clay, 2014a).⁹ Each of these datasets depends on content analyses of annual reports from the US State Department, Amnesty International, and, in the case of PTS, Human Rights Watch. Academics and their students hand code these reports to produce ordinal scales that measure violations of civil, political, and personal integrity rights. These measures are grounded in international legal principles and are intended to measure violations of international human rights law. The PTS was originally created to examine “whether U.S. foreign aid was being sent to countries that violated international human rights standards, thereby being in violation of [US] federal law,”¹⁰ the law in question being the 1976 amendment to the Foreign Assistance Act which prohibits the US from providing assistance to countries which consistently engage in gross violations of internationally recognised human

⁸ Physical integrity rights are “the entitlements individuals have in international law to be free from arbitrary physical harm and coercion” (Cingranelli and Richards 1999, 407). They include the rights to be free from torture, disappearance, execution, arbitrary arrest, and political imprisonment.

⁹ The CIRI data (<http://www.humanrightsdata.com>) have not been updated since 2014, and the data only cover the period 1981-2011.

¹⁰ <http://politicalterror scale.org/About/History/>

rights. The CIRI project coding guide cites specific provisions from the International Covenant on Civil and Political Rights to ground the coding rules for each of its civil, political, and physical integrity rights scales (Cingranelli, Richards and Clay, 2014*b*). The Political Terror Scale is a single, five-point ordinal scale that measures political arrests and killings, torture, and disappearance. The CIRI dataset includes separate three-point ordinal scales for extrajudicial killings, disappearance, torture, political imprisonment, freedom of speech/press, freedom of religion, freedom of domestic movement, freedom of foreign movement, freedom of assembly/association, and electoral self-determination.

Two more recent projects have produced quantitative scales that focus specifically on torture and are also grounded in international law. One of these was created by Oona Hathaway and is described in Hathaway (2002, pp. 1969-1792). She also used US State Department annual reports to produce a five-point ordinal scale that measures the prevalence and severity of abuse that constitutes torture under international law.¹¹ There is also the Ill-Treatment and Torture Data (Conrad, Haglund and Moore, 2013, 2014), which uses Amnesty International Annual Reports, press releases, and Action Alerts to code allegations of torture. ITT's coding rules are grounded in the Convention against Torture and Other Cruel, Inhuman, Or Degrading Treatment or Punishment (henceforth, the Convention against Torture, or CAT) (Conrad and Moore, 2010), and their data include an ordinal scale measuring the prevalence of torture as well as specific information regarding each allegation, e.g. the identity of the victim and the responsible government agency.

Another recently created measure is derived from a statistical model akin to the one HRMI uses (described in section 2.4 below). Fariss (2014) uses a measurement model to combine most of the scales discussed above, as well as several indicators of genocide/mass killing created from a variety of secondary sources, into a single index of government respect for physical integrity.

Finally, the Varieties of Democracy (V-Dem) Project has, since 2014, conducted expert opinion surveys of academics to create quantitative measures of torture, political killings, freedom of association, freedom of expression, and political participation (Coppedge et al., 2017). The definition of torture provided in the V-Dem codebook (p. 221) is similar to the CAT's, though the other V-Dem scales are not explicitly grounded in international law. Academics are asked to rate countries on an ordinal scale for all of these practices, and their responses are converted into numeric scales. In the case of torture and political killing, the responses are converted to scales using a model very similar to ours. Measures of freedom of association and expression are created from a measurement model that takes as inputs several subcomponents, each of which are created in the same manner as the torture and killing scales. For example, the freedom of association index is created from sub-indices for bans on political parties, barriers to the formation and functioning of political parties, autonomy of opposition parties, multiparty

¹¹ Hathaway relies on the Convention Against Torture (CAT) and several regional treaties for her definition of torture.

elections, civil society organisation entry and exit, and civil society organisation repression. The political participation scale is created in a similar manner, except the components are aggregated by taking their average instead of using a measurement model. Of all previous efforts to measure civil and political rights, the V-Dem project is the most similar to ours as it uses expert surveys and combines the responses for individual countries using a statistical model.

Like most existing human rights measurement efforts, HRMI's civil and political rights metrics are grounded in international law and are intended to measure violations of internationally recognised human rights principles. The survey we administer explicitly defines the rights under analysis with references to relevant international treaties and conventions, including the International Covenant on Civil and Political Rights (ICCPR), the Convention against Torture (CAT), and the International Convention for the Protection of All Persons from Enforced Disappearance (henceforth, the Convention on Enforced Disappearance, or CED).

In terms of methodology and coverage, there are two main ways in which HRMI's measures represent improvements over current practices. First, and most important, our source of information is human rights advocates, researchers, lawyers, and other experts, typically located in the country in question. By contrast, previous efforts rely on NGO reports intended for public consumption, and surveys of academics in the case of V-Dem. By getting our information directly from primary sources, and by offering our survey in many different languages, we ensure that our expert survey is serving as a tool that gives a voice to human rights experts located in countries around the world, to share their knowledge with the rest of the world.

Second, our measures also cover several aspects of human rights omitted by all of the other measures discussed above. For instance:

- Our data on arrests includes arbitrary/unlawful arrests unrelated to political activity. Such arrests are prohibited by the ICCPR but are not considered by the measures discussed above.
- We provide a measure of the prevalence of death penalty executions. Use of the death penalty is a violation of the ICCPR's Second Optional Protocol but is ignored by existing measures of physical integrity rights.
- We collect and publish information on the populations who are being targeted or at highest risk for civil and political rights abuse. This may turn out to be one of the most valuable aspects of our dataset, as it helps people gain a greater understanding of abuse than can be inferred from a single number alone. We discuss all of this in greater detail in the next two sections.

2.3.2 Problems of information

The problem of obtaining reliable, unbiased, and comprehensive information is perhaps the most serious impediment to the collection of quantitative civil and political rights data. When violations are reported, states often attempt to frame the abuse as either committed out of necessity or carried out by bad actors

without the state's permission (McCoy, 2012, 52). Likewise, by their very nature, many violations of civil and political rights are clandestine, with the violator seeking to conceal their actions entirely (e.g. Conrad, Hill, and Moore, 2014; Rejali, 2009).

Further, the degree to which violators succeed in concealing their complicity in abuse only serves to exacerbate the problems surrounding any attempt to collect comparable information about different countries' human rights violations. Most previous attempts to collect cross-nationally comparable data on a full range of civil and political rights has done so by relying on public documentation, especially by the U.S. State Department and international non-governmental human rights organisations (HROs), like Amnesty International and Human Rights Watch (e.g. Cingranelli, Richards and Clay, 2014*a*; Conrad and Moore, 2010; Gibney et al., 2015). These projects have been able to produce data that are highly reliable (Fariss, 2014), but, either explicitly (Conrad and Moore, 2010) or implicitly via their construction (see the standards-based categorisation utilised by Cingranelli, Richards and Clay (2014*a*) and Gibney et al. (2015)), these projects also acknowledge severe limitations in the information on which their estimates are based. As Bollen (1986) discusses, human rights violations often go unreported in international news sources or the reports of international non-governmental organisations, even when individual journalists or organisation members have information on those violations. Human rights organisations have to be strategic in the use of their limited resources and in the maintenance of a credible international image. As such, HROs understandably focus primarily on those places and issues on which they are most likely to have an impact (Barry, et al., 2015; Hendrix and Wong, 2014). This focus on maintaining the effectiveness and credibility of the organisation means that HROs are unlikely to report on events that did not happen; however, it also means that many abuses go unreported (Hill, Moore, and Mukherjee 2013). Further, the distance between what is reported about human rights abuses and what is known about them is almost certainly larger for some countries than others. Some countries have more journalists and active members of HROs than others do; further, some countries receive a greater share of international attention than do others. As such, if we attempted to generate a count of human rights abuses based on the information sources most commonly used by previous measurement projects, we would end up with a biased undercount, in which we overestimate the degree to which human rights are enjoyed everywhere, but more in some places than others (Conrad and Moore, 2011; Conrad, Haglund, and Moore, 2014).

Many have tried to respond to the problem of the biased undercount using various means. The limited ordered scales used by PTS (Gibney et al., 2015) and CIRI (Cingranelli, Richards and Clay, 2014*a*) acknowledge the lack of precision in numbers provided by human rights reports. Nevertheless, others have argued that even these limited containers are subject to the problem of undercounting, especially if the undercount and the bias contained therein has changed over time (e.g. Clark and Sikkink, 2013; Fariss, 2014). As such, some have suggested that regression analyses utilising these potentially biased data should use some statistical method for accounting for that bias (e.g. Bagozzi, et al., 2015;

Conrad, Hill, and Moore, 2014). While this strategy may help to ensure that the inferences we draw from secondary analyses are valid, it does less in terms of providing easy to understand measurements for a wide audience. In an effort to provide something more useful in this regard, Fariss (2014) attempts to produce data that account for changing standards of accountability over time to provide an overall measure of physical integrity rights for every country in the world by utilising multiple data sets of various types of abuse. Assuming that its assumptions hold, this correction for bias could certainly serve as an improvement over previous efforts. However, one would hope to have higher quality data for each type of abuse in the first place; further, as discussed below, one would also hope to forgo the extreme data reduction process necessary to obtain these estimates, reducing several kinds of human rights practices to a single number.

The Varieties of Democracy Project (V-Dem) has attempted to sidestep these problems of information by turning to another source of information: experts on the countries being discussed (Coppedge et al., 2017). This solution is elegant, as it avoids the problems of relying strictly on the public documents produced by governments and organisations and goes directly to individuals who are hopefully (1) aware of the situation in the country about which they are being asked and (2) capable of comparing the current situation to past situations on equal footing.¹² While we believe this approach is a welcome step forward, we still have reason to doubt whether V-Dem's approach is truly the best possible option. Particularly, we question whether the experts chosen by V-Dem are truly the best possible experts to ask about the most current human rights information, particularly if we want to adequately describe our level of certainty in that information. In most cases, a V-Dem Country Expert holds a PhD degree, suggesting that most respondents are likely to be academics.¹³ While academics undoubtedly know more about the subjects at hand than the average person, they are not typically the people most responsible for collecting information on the day to day violation and enjoyment of human rights. Indeed, there is good reason to believe that academics may primarily rely on secondary sources for their human rights information. If those academics are all primarily relying on similar sources to collect their human rights information, and particularly if those sources are public media and organisational reports, then the estimates of abuse taken from those academics are likely to (1) suffer from the same bias that has arguably afflicted previous measures of human rights based on those secondary sources and (2) overestimate the certainty of those estimates because agreement between academic respondents will be inflated due to reliance on the same secondary sources. As such, while we think one may be able to gather valuable information from academics about a great many subjects (including many of subjects studied by V-Dem), we believe there is likely a better pool of respondents for studying human rights.

¹² Indeed, Fariss (2018) explicitly makes this argument, showing that V-Dem's data for certain types of human rights abuse over time closely match the pattern of change shown in his physical integrity rights data.

¹³ For more information on V-Dem Country Experts, see V-Dem's 2017 call for those experts at <https://www.v-dem.net/en/news/call-country-experts-v-dem/> (Last Accessed: March 18, 2018).

2.3.3 Problems of interpretation

Given the many problems of information laid out above, it is unsurprising that the interpretation of the limited information to which previous projects have had access has also faced huge hurdles. In particular, we focus on two overarching interpretive issues in previous data projects that we hope to improve upon: (1) the accurate representation of uncertainty and (2) the dimensionality of civil and political rights abuse.

As mentioned above, the most well-known previous attempts to measure civil and political rights are the Political Terror Scale (Gibney et al., 2015) and the CIRI Human Rights Data Project (Cingranelli, Richards and Clay, 2014a). As discussed above, each of these projects handled the problem of uncertainty in the information contained in the human rights reports by using standards-based scales, allowing for the broad categorisation of states for use in comparisons. While this is a reasonable approach to making cross-national comparisons on the basis of limited, biased information, it still has problems with regard to conveying the level of certainty we have about any single country's score. For instance, the CIRI measure for torture and ill-treatment allowed for grouping states into three categories: those with no reported abuse in the State Department and Amnesty International reports (scored a 2), those with reports that suggested that torture was practiced occasionally (1), and those with reports that suggested that torture was practiced frequently (0) (Cingranelli, Richards and Clay, 2014b). While this categorisation is reasonable given the low level of informative precision found in the human rights reports, it also leads to problems. The first problem is one of data truncation. For instance, a country with 500 documented instances of torture and another with 50,000 would fall in the same category of frequent abuse, each receiving a score of 0. While both countries are certainly engaged in high levels of abuse, they are not "equal". While many academic human rights researchers understood this, popular perception of these scores never quite caught up, with the media sometimes pointing out that unexpected countries shared a similar score with some of the world's worst human rights violators.¹⁴ Second, beyond the problem of data truncation, there was the problem of uneven information. Based on the way that CIRI and PTS scores have been constructed, it is not possible to know the degree of certainty around a country's categorical placement. Returning to CIRI's torture measure, did a state receive a 1 because it only engaged in a few instances of torture, or was it because there just was not enough information to justify placing it in the worst category? Was it close to the border line between categories or quite far away? When only one score is provided for a right, it is not possible to know the answers to these questions from the data alone.

Further, most previous attempts to collect cross-nationally comparable civil and political rights data have also ignored the dimensionality of rights abuse by governments. Stohl et al. (1986, 600-603) notes that there are three dimensions to the violation of civil and political rights: (1) scope, (2) intensity, and (3)

¹⁴ For an example of this, see Ophir Bar-Zohar's article in Haaretz from December 14, 2011, "Israel Earns Another Failing Score on Freedom of Religion Index," in which the author makes a point of mentioning that Israel received the same score as China, Iran, Saudi Arabia, and Afghanistan: <https://www.haaretz.com/1.5219143> (Last Accessed: March 18, 2018).

range. "Scope" refers to the type of abuse the violator has engaged in, i.e. the particular right being violated. For instance, have the violators tortured political opponents, arrested them, or allowed them to keep participating in elections? Have they done one of these things, two, or all of them? These are questions of scope. "Intensity" refers to the frequency of each type of abuse. For example, did the violator arbitrarily imprison one person, two people, or hundreds? Finally, "range" refers to the portion of the population that has been targeted for abuse. Did the violator focus their abuses on political opponents, on accused criminals, or on discriminated groups or classes? Or, alternatively, was the abuse indiscriminate, placing all people at risk? These are the kinds of questions one would ask regarding range.

While these dimensions of abuse have long been recognised, every previous project aimed at collecting cross-nationally comparable civil and political rights data has failed to fully capture at least one of these dimensions. For instance, while PTS captures aspects of scope, intensity, and range, it collapses all of those dimensions into a single score, essentially treating three separate dimensions if they can be captured on a single scale (Gibney et al., 2015). While CIRI does a better job of separating scope by using disaggregated measures of different types of abuse, its individual scores only measure the intensity of those particular types of abuse with no comparable measure of range. Similar to PTS, Fariss (2014) produces a single score for all physical integrity rights, and in a method similar to CIRI, V-Dem provides very little information on range (Coppedge et al., 2017).

To summarise, we are heavily indebted to the projects that have preceded HRMI. Some of us directly participated in some of these data collection efforts, while others of us have published extensively using them. All of the projects discussed here have been conducted with the best of intentions, and they have often represented the best approach possible at the time of their creation. That said, we believe that it is possible to improve on all of them. In our efforts to do this, we particularly intend to (1) use better sources of information than were previously available, (2) provide transparent indicators of uncertainty, and (3) measure the full dimensionality of civil and political rights abuse. Our approach to accomplishing these three goals is described in the next section.

2.4 HRMI's approach to civil and political rights measurement

The Human Rights Measurement Initiative (HRMI) aims to produce a comprehensive suite of measures that covers the full range of human rights listed in the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and the International Covenant on Economic, Social, and Cultural Rights, along the rights covered in other core United Nations human rights treaties (HRMI, 2018).¹⁵ Further, we seek to create measures for every country in the world in a way that ensures cross-

¹⁵ While section 2 of this methodology guide focuses on HRMI's Civil and Political Rights Measures, the HRMI pilot data also include measures of 5 Economic and Social Rights, based on the measurement strategy employed by Fukuda-Parr, Lawson-Remer, and Randolph (2015). For more information, see section 4 in this guide.

national comparability, while remaining transparent in the means by which those measures are created. Ultimately, we want to create data that are useful for human rights advocates, researchers, journalists, and anyone else seeking information on human rights worldwide. In pursuit of these goals, we have to take new approaches to the methods by which we collect and interpret human rights data.

As described above, we particularly wanted to improve on (1) the quality of information, (2) the transparency of uncertainty, and (3) the disaggregation of the dimensions of human rights abuse observed in previous civil and political rights data projects. We have attempted to answer these challenges by (1) directly collecting information from human rights researchers and practitioners that are gathering information and monitoring human rights issues in each country, (2) using statistical methods that allow us to accurately and honestly report our uncertainty with regard to the intensity of abuse, and (3) collecting data not only on the scope and intensity of abuse, but also the range of abuse (i.e. the distribution among groups at risk). In this section, we describe our approach to collecting civil and political rights data, beginning with a discussion of the 2019 version of the HRMI Civil and Political Rights expert survey, followed with a more detailed description of the model used to obtain the intensity score for each right measured.

2.4.1 The 2019 HRMI Expert Survey

In order to directly collect information on civil and political rights performance in countries around the world, we developed the HRMI Expert Survey. In 2019, the goal for civil and political rights was to collect information on state performance in 2018 and 2017 across seven areas of civil and political rights, each connected directly to language contained in the International Covenant on Civil and Political Rights (ICCPR) and other relevant international law. These are: the right to be free from torture and ill-treatment (Article 7 and the Convention against Torture), the right to be free from execution (Article 6 and the Second Optional Protocol to the ICCPR), the right to be free from arbitrary or political arrest and detention (Articles 2, 9, 11, 18, 19, 21, 22, and 26), the right to be free from disappearance (Articles 9 and 10, and the Convention on Enforced Disappearances), the right to political participation (Article 25), the right to opinion and expression (Article 19), and the rights to assembly (Article 21) and association (Article 22). As such, we designed our survey to have a section for each of these seven rights. Each section contains (1) a definition of the right under consideration, (2) questions related to the intensity of respect for that right in 2018 and 2017, and (3) questions regarding the range of respect for that right, i.e. who was targeted and/or especially at risk of abuse.¹⁶

The definition of each right was determined on the basis of international law and its interpretation by the appropriate treaty bodies at the United Nations. For instance, the definition of torture and ill-

¹⁶ A preview of our 2019 survey can be viewed in its entirety at https://ugorgia.ca1.qualtrics.com/jfe/preview/SV_1Iavg87gjvvi7Gd?Q_SurveyVersionID=current&Q_CHL=preview. This version of the survey is not “live” and responses will not be used.

treatment is broadly based on the definition found in Article 2 of the Convention against Torture and Other Cruel, Inhuman, or Degrading Treatment or Punishment (CAT). The following is taken directly from our survey:

All people have the right to be free from torture and ill-treatment. When answering the questions below, please use the following broad definition:

Torture and ill-treatment consist of “any act by which severe pain or suffering, whether physical or mental, is intentionally inflicted on a person” (CAT, Part 1, Article 1). Torture and ill-treatment may be committed for any specific purpose, including (but not limited to) attempts to obtain information or confessions, punishment for suspected or committed acts, intimidation, coercion, and discrimination.

We proceed in a similar fashion for all other rights in the survey, drawing on the ICCPR, the CAT, the International Convention for the Protection of All Persons from Enforced Disappearance (CED), the Second Optional Protocol to the ICCPR, and general comments from the Human Rights Committee.

Next, we ask our respondents about the intensity of violations by state actors. For instance, in the case of torture and ill-treatment in 2018, we ask:

In 2018, how often did government agents, such as soldiers, police officers, and others acting on behalf of the state, commit acts of torture or ill-treatment?

Respondents answered this question on the basis of an 11-point scale, ranging from a score of 0, which represented an answer of “Never”, up to a score of 10, which represented an answer of “Constantly.”¹⁷

At this point, we turned to questions about the range of respect for the rights being discussed. First, we ask our respondents to provide us with more specific information about those who were especially at-risk for abuse in 2018, asking for torture:

In 2018, which groups of people, if any, were especially vulnerable to torture and ill-treatment by government agents, such as soldiers, police officers, and other state-sanctioned actors? (Select all that apply.)

In response to this question, respondents can select from 31 identifiers pre-imported into the survey (including people of particular ethnicities, people of particular races, LGBTQIA+ people, and people with particular religious beliefs and practices, among many others), or provide us with other potential identifiers that we did not have the foresight to include. We then further follow up this question with an open-ended question asking for more specific information on why the respondent chose those particular identifiers. In the 2019 survey, we used this same format to include a section asking about which people were likely to lack enjoyment of their economic and social rights, focusing particularly on the rights to health, housing, education, food, a job, just and favourable working conditions, and, finally, the rights to strike and join and form unions. All of the data on people at risk are presented in our data and visualizations as the count and proportion of a state’s respondents that told us a particular chosen group

¹⁷ It should be noted here that the survey question is inverted from the final score presented in our results, in which higher scores represent better respect for the right in question.

was a risk for violations, restrictions, or a lack of enjoyment of the right in their country (depending on the right). Finally, our data visualization site provides summaries of the respondents' open ended responses.

Beyond each of the sections focused on particular rights, we also include sections focused on asking our respondents to score the intensity of three hypothetical countries on their respect for the rights under consideration. These hypothetical cases are included to account for differences in the interpretation of the 11-point intensity scale described above. The respondents' answers to these questions contribute meaningfully to the final intensity scores produced for each country in the manner described in the "Model Description" section below.

2.4.2 Selection of pilot countries and expert survey respondents

A significant benefit of our approach to measuring civil and political rights is the ability to avoid some of the biases that exist in the public documentation of abuses of these rights, by collecting information directly from experts on the human rights situation in each country being studied. However, this raises the question: Who qualifies to be an expert respondent to the HRMI civil and political rights survey?

Thus far, we have focused primarily on human rights practitioners who are directly monitoring the civil and political rights situation in each country. These people typically work for an international or domestic non-governmental organisation or a civil society organisation. We also allowed for participation by human rights lawyers, journalists covering human rights issues, and staff working for National Human Rights Institutions if that institution has been rated as fully compliant with the Paris Principles, i.e. those that have been given "A"-level accreditation by the International Coordinating Committee and its Sub-Committee on Accreditation (United Nations, 2010; GANHRI, 2019).¹⁸

To the extent possible, we have tried to rely on respondents who are located within the country on which they are providing information. But in cases of more closed and repressive countries, we have been, and will continue to be, forced to rely on a higher proportion of respondents who are based outside of the country of interest. Our main goal has been to collect information from respondents who are first points of contact for human rights information in the country of interest and who have often had access to primary sources. As such, we do not intend to rely on academics as respondents in most cases, as they are rarely involved in the collection of primary information and tend to rely more heavily on secondary sources. Likewise, in order to ensure that our measures are independent from government-backed sources, staff at government-organised NGOs and government officials outside of A-level NHRIs have also been excluded from being respondents.

¹⁸ The countries in our 2019 sample with an NHRI that meets this criterion are: Australia, Democratic Republic of Congo, Jordan, Liberia, Mexico, Nepal, New Zealand, South Korea, and the United Kingdom ([https://nhri.ohchr.org/EN/Documents/Status%20Accreditation%20Chart%20\(04%20March%202019\).pdf](https://nhri.ohchr.org/EN/Documents/Status%20Accreditation%20Chart%20(04%20March%202019).pdf), accessed on 5 June, 2019).

The sample of potential respondents was determined by a two-step process. First, we asked for nominations from human rights advocates worldwide for countries to include in the initial pilot in 2017. Thirteen countries were nominated, and we selected all 13 for inclusion in the pilot, as together they provided significant diversity in government type, country size, level of development, geographic location, and many other factors. The 13 countries are: Angola, Australia, Brazil, Fiji, Kazakhstan, Kyrgyzstan, Liberia, Mexico, Mozambique, Nepal, New Zealand, Saudi Arabia, and the United Kingdom. We conducted a similar search for additional states to include in the 2019 survey, with a strong focus on adding states where we could identify a “HRMI Ambassador,” i.e. a locally-focused partner that helps to connect HRMI to the local human rights community and helps us identify potential survey respondents for that country.¹⁹ As a result, we added five countries to the sample: Democratic Republic of Congo, Jordan, South Korea, the United States of America, Venezuela, and Vietnam. This diverse sample of 19 countries helps to demonstrate how well our methodology works across different contexts.

Second, relying on HRMI Ambassadors and trusted partners in non-governmental human rights organisations around the world, we engaged in a snowball sampling technique whereby potential respondents who met our criteria in each of our sample countries were referred to us. As potential respondents were added, those respondents were also asked if they could recommend potential respondents. By the end of the process, we had identified between 12 and 74 potential survey respondents per country, each of whom was sent a single-use survey link, to ensure that the survey link was not shared with unintended respondents. Table 1 summarizes the information available about the number of responses and response rates for each country in our 2019 Expert Survey. Final response rates (counting only those who filled out the survey in its entirety) ranged from 17% in Mozambique up to 60% in South Korea. The number of fully-completed surveys that were used to produce the civil and political rights data for each country ranged between 6 and 19. However, responses from partially-completed surveys were also used, to the extent possible.

2.4.3 Producing intensity scores: model description

The simplest way to combine expert survey responses on the intensity questions into a single score for each country-year would be to report the average of the survey responses for that question. While this technique is straightforward and commonly employed in many settings, there are several potential problems with this method that would bring the validity of the scores into question. Namely, simply averaging the survey responses assumes that each survey question and each expert should contribute equally to the underlying quantity being estimated. Additionally, the simple approach assumes that experts in different countries will view the scale points of the survey questions in comparable ways. In order to overcome these potential problems, we use statistical models that estimate unobserved, latent

¹⁹ For a list of HRMI Ambassadors that are willing to be publicly identified, see: <https://humanrightsmmeasurement.org/about-hrmi/the-team/>.

traits/characteristics for individual observations (in our case countries), from a set of observed outcomes (in our case survey questions).

Table 1
2019 HRMI Expert Survey Response Rates

Country	Number of people sent a survey link	Number of completed surveys	Survey completion rate
Angola	53	11	21%
Australia	62	15	24%
Brazil	74	14	19%
DRC	12	6	50%
Fiji	61	12	20%
Jordan	21	6	29%
Kazakhstan	40	17	43%
Kyrgyzstan	37	19	51%
Liberia	64	14	22%
Mexico	35	8	23%
Mozambique	48	8	17%
Nepal	51	12	24%
New Zealand	39	11	28%
Saudi Arabia	32	9	28%
South Korea	20	12	60%
United Kingdom	39	7	18%
United States	36	9	25%
Venezuela	19	8	42%
Vietnam	28	17	61%
Total	771	215	
Average	41	11	32%

The models we use are Bayesian variants of the common factor model, which were developed primarily in the fields of psychology and sociology (Bollen 1989). These models have been developed to uncover the latent dimensionality within a set of observed indicators of some concept. For example, a survey that is designed to measure an individual's political ideology, might ask a battery of questions about a respondent's position on a variety of policies/issues, such as position toward same-sex marriage, gun control, and redistribution of wealth. We would expect that a given respondent would answer these questions in similar ways, representing either more left or right-wing ideological views.

Formally, the factor model is as follows:

$$Y_{ij} = \alpha_j + \beta_j \Theta_i$$

Here Y_{ij} is individual i 's response to survey question j . Θ_i is individual i 's ideology and β_j is the factor loading that maps individual i 's response to question j to their latent position Θ . Larger values of β represent a stronger association between the survey question and the latent trait. α_j is an intercept that is often omitted by standardising both Y and Θ .

In our case, the unobserved concept of interest is the intensity of human rights respect in a given country and the observed outcomes are survey responses from experts, as defined above, in that country. In our survey, we ask experts to rate countries on their performance in the areas of the rights to:

- freedom from torture and ill-treatment,
- freedom from arbitrary or political arrest and imprisonment,
- freedom from extrajudicial execution,
- freedom from death penalty execution,
- freedom from disappearance,
- political participation,
- opinion and expression,
- assembly and association.

Respondents placed their respective countries on a 0-10 scale, where higher values correspond to worse conditions.²⁰ Questions about each country serve as the questions/items for the factor analysis, analogous to questions on a public-opinion survey, and the human rights performance of a given country is analogous to an individual's ideology in the previous example.

As in the standard setup, we treat each of our survey responses partly as a function of the "true" human rights conditions in each country. Unlike the standard approach, our model estimates a latent trait for each *item*, i.e. country-year, which is assumed to be fixed across respondents. In this setup the α and β parameters discussed above vary across respondents rather than items, so that each survey response is also a function of respondent-specific parameters that represent how each respondent translates the underlying human rights conditions in their country into a score on the numeric scale presented in the survey question. This allows for the fact that survey respondents may respond differently to the same objective conditions. That is, Respondent 1 may give a score of 6/10 in response to a particular set of objective conditions, whereas Respondent 2 could give the same country a score of 4/10. This feature of the model, combined with anchoring vignettes (described below), allows us to place each country on a common scale even when respondents treat the numeric values on the scale differently.

²⁰ As noted above, the survey question is inverted from the final score presented in our results, in which higher scores represent better respect for the right in question.

Because we are estimating a Bayesian version of the model, we must supply distributional information that is not necessary in the standard approach. As the survey responses have 11 categories, we treat them as normally distributed. We can write our model:

$$\begin{aligned} Y_{ij} &\sim \mathcal{N}(\mu_{ij}, \tau_{ij}) \\ \mu_{ij} &= \alpha_i + \beta_i \Theta_j \\ \tau_{ij} &= \tau_i \tau_j \end{aligned}$$

where Y_{ij} is respondent i 's rating of country j 's human rights conditions and Θ_j is the “true” value of human rights performance in country j .²¹ Each α_i represents respondent i 's tendency to place countries lower/higher on the scale. A respondent with a negative α tends to rank countries on the low end of the scale, while one with a positive α tends to push their rankings towards the high end. Each β_i represents how well a respondent distinguishes between poor and good human rights conditions. Respondents with β s closer to 0 place countries with different human rights performances relatively close together on the scale, while those with more positive β s place countries with different performances relatively far apart on the scale. A negative value of β would indicate that the respondent ranks countries with worse performance higher than those with better performance, which is something we allow for but which we did not observe happening in practice. Finally, we allow the variation in survey responses, τ_{ij} to be a function of both respondent and item level variation.²²

One of the advantages of our approach versus a simpler approach to aggregating survey responses to the country level (e.g. taking the simple mean of the responses) is that our approach can handle differences in how experts may view the underlying response across different countries. That is, what one expert may view as a 6 another may view as a 4. As our respondents are country-specific, we include a set of hypothetical countries, described in the survey, that all experts place regardless of their country of expertise. These “anchoring vignettes” combined with the Bayesian factor model described above, allow us to correct for any potential differences in how experts view the underlying scales in our survey. That is, we use questions about hypothetical countries as “bridging observations” in order to estimate the model and to create a scale that is cross-nationally comparable. An example data matrix for our model, with 6 respondents from 3 countries, is shown in Table 2.²³

²¹ The one exception to this across our 7 civil and political rights is our combined indicator for freedom from execution, which is simply equal to the lower of the scores between our freedom from extrajudicial execution and death penalty execution indicators.

²² This is a variation of the Bayesian Aldrich-McKelvey model. See Hare, et al (2014) for more detailed information.

²³ For a more detailed discussion of anchoring vignettes and expert surveys, see Bakker et al (2014).

Table 2: Example data for Bayesian Aldrich-McKelvey model

Respondent	Country 1	Country 2	Country 3	Vignette 1	Vignette 2	Vignette 3
1	5	–	–	1	5	7
2	3	–	–	5	7	8
3	–	1	–	3	5	10
4	–	4	–	2	3	5
5	–	–	6	6	8	9
6	–	–	9	3	6	8

We estimate our model via Markov chain Monte Carlo simulation. We adopt the following non-informative conjugate prior distributions for the parameters in our model:

$$\begin{aligned}\alpha_i &\sim U(-100, 100) \\ \beta_i &\sim U(-100, 100) \\ \Theta_{ij} &\sim \mathcal{N}(0, 1) \\ \tau_j &\sim \text{Gamma}(0.1, 0.1) \\ \tau_i &\sim \text{Gamma}(v, \omega) \\ v &\sim \text{Gamma}(0.1, 0.1) \\ \sigma &\sim \text{Gamma}(0.1, 0.1)\end{aligned}$$

We let our model run for 11,000 iterations and store the last 1,000 draws from the posterior distributions to summarise the model parameters. We assessed convergence via visual inspection of density plots and the Gelman-Rubin statistic, and all parameters show strong evidence of convergence.

This produced posterior intensity distributions with means that range from approximately -1.01 at the lowest up to 1.46 at the highest, and standard deviations that range from approximately 0.05 to 0.42. For the purposes of presentation, we rescaled these distributions to generate means that varied between around 0 and 10, with higher scores indicating better government performance with regard to that right. This was done by adding 1.02, multiplying by 3.6, and then finally, adding 1.08 to all means, while merely multiplying all standard deviations by 3.6. The resulting re-scaled distributions have means that ranged from approximately 1.10 up to around 9.99, with standard deviations ranging from 0.19 to 1.50.

2.4.4 Producing scores for safety from the state and empowerment

Using the scores produced from the survey responses, we also create two overarching indicators indicating overall safety from the state, i.e. government respect for physical integrity rights, and overall enjoyment of empowerment rights, i.e. the ability to take an active role in society without fear of reprisal.²⁴ In doing so, we use similar logic to that laid out for the common factors model above.

²⁴ These two overarching dimensions have a long history in the quantitative human rights literature. Physical integrity rights largely represent the single dimension measured by both the Political Terror Scale (Gibney et al., 2015) and Fariss (2014), while both dimensions were previously measured by the CIRI Human Rights Data Project (Cingranelli and Richards 1999; Richards, Gelleny, and Sacko 2001; Cingranelli, Richards, and Clay 2014a).

However, in this instance, we are treating each of our five disaggregated physical integrity rights indicators (i.e. freedom from torture, disappearance, extrajudicial execution, death penalty execution, and arbitrary or political arrest and imprisonment) and three disaggregated empowerment rights indicators (i.e. assembly and association, opinion and expression, and political participation), calculated via the methods described above, as caused by the unobserved variables of overall safety from the state and overall empowerment, respectively.

In order to accurately report uncertainty in the overarching indicators, we had to take into account the uncertainty calculated for each of our disaggregated indicators. As such, we took 10 draws from the posterior distributions of each of our disaggregated indicators. Then, for both safety from the state and empowerment, we ran our Bayesian measurement models ten times apiece, one for each set of draws. Each of our disaggregated rights indicators are continuous and normally distributed, so, following Bakker, Hill, and Moore (2016), we can write our model:

$$y_{i,t} \sim \mathcal{N}(\mu_{i,t}, \sigma^2)$$

$$\mu_{i,t} = \beta X_{i,t}$$

We assumed standard normal prior distributions for the latent variables, while the β parameters were assigned normal prior distributions with a mean of 1 and a variance of 2. We assigned the σ parameter a Gamma prior distribution with scale and shape parameters of 1. We let our models run for 4000 iterations and saved the last 2000 iterations to summarize the parameters. All models showed strong evidence of convergence. Finally, we produced the final means and standard deviations for each of our overarching indicators by using Rubin's (1987) rules for combining analyses following multiple imputation.

The resulting safety from the state (i.e. physical integrity rights) variable had means ranging from -.65 to 1.42, with standard deviations ranging from .36 to .50; the empowerment variable had means ranging from -1.55 to 1.39 with standard deviation ranging from .37 to .54. As with our disaggregated indicators, these variables were rescaled to have means that varied between 0 and 10, which was done by adding 1.54, multiplying by 2.5, and finally adding 1 to the means, while merely multiplying the standard deviations by 2.5. The resulting safety from the state measure has means ranging from 3.24 to 8.4 and standard deviations ranging from .89 to 1.25, while the resulting empowerment measure has means between .98 and 8.33 and standard deviations between .93 and 1.36.

2.5 Conclusion

We believe the 2019 HRMI civil and political rights data demonstrate the benefits of collecting information on the full scope, intensity, and range of government respect for civil and political rights directly from human rights experts in countries around the world. Further, the statistical methods we use to convert this information into quantitative metrics allow us to be honest about uncertainty and permit sensible cross-country comparisons. This work represents a significant advance over existing human

rights data projects, and we plan to extend coverage to a wider sample of countries as soon as possible. Indeed, the goal for HRMI going forward is to gradually expand the sample of countries to include the global population, while at the same time expanding our coverage of rights to include all of those included in the broader corpus of core international human rights treaties.

Nevertheless, much work remains to be done. How should we incorporate information on the actions of non-state actors into our metrics? How might we obtain even better disaggregated data on targeted and discriminated classes, groups, and identities? What can these data help nations learn about the importance of human rights and the best path for reforms toward greater respect for them? These questions will continue to drive our efforts as we move forward and attempt to innovate. To accomplish these goals, we will continue to need help. Indeed, as an initiative that is founded on innovation through collaboration, we sincerely hope to get feedback on our approach and move forward in a way that makes our data as useful as possible for the largest number of people we can.

References

- Amnesty International. 2016. "Amnesty International Report 2015/16: The State of the World's Human Rights." <https://www.amnesty.org/download/Documents/POL1025522016ENGLISH.PDF> (Last Accessed: March 18, 2018).
- Bagozzi, Benjamin E., Daniel W. Hill Jr, Will H. Moore, and Bumba Mukherjee. 2015. "Modeling two types of peace: The zero-inflated ordered probit (ZiOP) model in conflict research." *Journal of Conflict Resolution* 59 (4): 728-752.
- Bakker, Ryan, Seth Jolly, Jonathan Polk, and Keith Poole. 2014. "The European common space: Extending the use of anchoring vignettes." *The Journal of Politics* 76 (4): 1089-1101.
- Bakker, Ryan, Daniel W. Hill, Jr., and Will H. Moore. 2016. "How much terror? Dissidents, governments, institutions, and the cross-national study of terror attacks." *Journal of Peace Research* 53 (5): 711-726.
- Barry, Colin M., Sam R. Bell, K. Chad, Clay, Michale E. Flynn, and Amanda Murdie. 2015. "Choosing the best house in a bad neighborhood: Location strategies of human rights INGOs in the non-western world." *International Studies Quarterly* 59 (1): 86-98.
- Bollen, Kenneth A. 1986. "Political Rights and Political Liberties in Nations: An Evaluation of Human Rights Measures, 1950 to 1984." *Human Rights Quarterly* 8 (4):567-591.
- Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
- Cingranelli, David L., and David L. Richards. 1999. "Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights." *International Studies Quarterly* 43 (2): 407-417.
- Cingranelli, David L., David L. Richards and K. Chad Clay. 2014a. "The CIRI Human Rights Dataset." <http://www.humanrightsdata.com>. Version 2014.04.14. (Last accessed: March 15, 2018).
- Cingranelli, David L., David L. Richards and K. Chad Clay. 2014b. "The Cingranelli-Richards (CIRI) Human Rights Data Project Coding Manual." Manual Version 5.20.14. <http://www.humanrightsdata.com/p/data-documentation.html> (Last accessed: March 15, 2018).
- Clark, Anne Marie and Kathryn Sikkink. 2013. "Information Effects and Human Rights Data: Is the Good News about Increased Human Rights Information Bad News for Human Rights Measures?" *Human Rights Quarterly* 35(3):539-568.
- Conrad, Courtenay R., and Will H. Moore. 2011. "The Ill-Treatment and Torture (ITT) Data Collection Project Country-Year Data User's Guide." July 19. http://faculty.ucmerced.edu/cconrad2/Academic/Data_files/ITT_CY_UsersGuide19July11.pdf (Last accessed: March 14, 2018).
- Conrad, Courtenay R, Jillienne Haglund and Will H Moore. 2013. "Disaggregating Torture Allegations: Introducing the Ill-Treatment and Torture (ITT) Country-Year Data." *International Studies Perspectives* 14(2):199-220.
- Conrad, Courtenay R, Jillienne Haglund and Will H Moore. 2014. "Torture Allegations as Events Data: Introducing the Ill-Treatment and Torture (ITT) specific allegation data." *Journal of Peace Research* 51(3):429-438.
- Conrad, Courtenay R and Will H Moore. 2010. "The Ill-Treatment and Torture (ITT) Data Project Coding Rules and Norms." May 5. http://faculty.ucmerced.edu/cconrad2/Academic/Under_the_Hood_files/ITTCODEBOOKNorms6May10.pdf (Last Accessed: March 18, 2018).
- Conrad, Courtenay R., and Will H. Moore. 2011. "The Ill-Treatment and Torture (ITT) Data Collection Project Country-Year Data Users' Guide." July 19. Version 2.1. http://faculty.ucmerced.edu/cconrad2/Academic/Data_files/ITT_CY_UsersGuide19July11.pdf (Last Accessed: March 18, 2018).
- Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen,

- Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Laura Saxer, Brigitte Seim, Rachel Sigman and Jeffrey Staton. 2017. "Varieties of Democracy Codebook v7." Varieties of Democracy (V-Dem) Project. <https://www.v-dem.net/en/data/data-version-7-1/> (Last Accessed: March 18, 2018).
- Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108 (2): 297–318.
- Fariss, Christopher J. 2018. "Are Things Really Getting Better? How To Validate Latent Variable Models of Human Rights." *British Journal of Political Science* 48(1): 275–282.
- Gibney, Mark, Linda Cornett, Reed Wood, Peter Haschke and Daniel Arnon. 2015. "Political Terror Scale 1976-2015." <http://www.politicalterroryscale.org> (Last Accessed: March 18, 2018).
- Global Alliance of National Human Rights Institutions (GANHRI). 2016. "Chart of the Status of National Institutions." August 5. <http://www.ohchr.org/Documents/Countries/NHRI/ChartStatusNHRIIs.pdf> (Last accessed: March 14, 2018).
- Goldstein, Robert Jusin. 1986. "The Limitations of Using Quantitative Data in Studying Human Rights Abuses." *Human Rights Quarterly* 8(4):607–627.
- Hare, Christopher, David A. Armstrong, II, Ryan Bakker, Royce Carroll, and Keith T. Poole. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59 (3): 759-774.
- Hathaway, Oona A. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111: 1935–2042.
- Hendrix, Cullen S., and Wendy H. Wong. 2014. "Knowing your audience: How the structure of international relations and organizational choices affect amnesty international's advocacy." *The Review of International Organizations* 9 (1): 29-58.
- Hill, Daniel W., Jr., Will H. Moore and Bumba Mukherjee. 2013. "Information Politics Versus Organizational Incentives: When Are Amnesty International's 'Naming and Shaming' Reports Biased?" *International Studies Quarterly* 57: 219-232.
- Human Rights Measurement Initiative (HRMI). 2018. "Human Rights Measurement Initiative website." <https://humanrightsmasurement.org/> (Last accessed: March 18, 2018).
- McCoy, Alfred W. 2012. "In the Minotaur's Labyrinth: Psychological Torture, Public Forgetting, and Contested History." In *Speaking About Torture*, eds. Julie A. Carson and Elisabeth Weber, pages 37-58. Fordham University Press.
- Rejali, Darius. 2009. *Torture and Democracy*. Princeton, NJ: Princeton University Press.
- Richards, David L., Ronald D. Gelleny, and David H. Sacko. 2001. "Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries." *International Studies Quarterly* 45 (2): 219-239.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. Wiley and Sons.
- Stohl, Michael, David Carleton, George Lopez, and Stephen Samuels. 1986. "State Violation of Human Rights: Issues and Problems of Measurement." *Human Rights Quarterly* 8(4):592–606
- United Nations. 2010. "National Human Rights Institutions: History, Principles, Roles and Responsibilities." Professional Training Series, Number 4 (Rev. 1). http://www.ohchr.org/Documents/Publications/PTS-4Rev1-NHRI_en.pdf (Last accessed: March 14, 2018).
- United States Department of State, Bureau of Democracy, Human Rights, and Labor (USDS). 2016. "2015 Country Reports on Human Rights Practices: Egypt." April 13. <https://www.state.gov/j/drl/rls/hrrpt/2015/nea/252921.htm> (Last accessed: March 18, 2018).

3 HRMI Economic and Social Rights Metrics Methodology – Executive Summary

This is a brief explanation of how we constructed the Human Rights Measurement Initiative (HRMI)’s economic and social rights metrics (the blue ones on the bar charts).

These metrics are adopted and adapted from the Social and Economic Rights Fulfilment Index (SERF Index) developed by Susan Randolph, Sakiko Fukuda-Parr, and Terra Lawson-Remer.²⁵ The SERF Index gauges the extent to which countries meet their substantive economic and social right obligations of result. Specifically, HRMI’s *income adjusted* “Quality of Life” metric and underlying *income adjusted* economic and social right metrics are the SERF Index and its underlying Right Indices. These *income-adjusted* metrics set a benchmark consistent with each country’s immediate obligation of results under the International Covenant for Economic, Social and Cultural Rights (ICESCR). HRMI’s *global best* “Quality of Life” and its component *global best* economic and social right metrics follow a similar methodology, but set a benchmark reflecting countries’ common full obligation of results under the ICESCR. For more in-depth information on how they are constructed, please see Section 4.

3.1 What are economic and social rights?

The International Covenant on Economic, Social, and Cultural Rights (ICESCR) is a treaty adopted by the United Nations in 1966 and agreed to by 166 nations that sets out a list of economic, social, and cultural rights that we are all entitled to simply by virtue of being human. These include the rights to food, health, education, housing, work, and social security. HRMI’s metrics cover five out of six of these rights, with social security being the one that we have insufficient data on to independently measure. As relevant data covering more countries become available, we would like to incorporate cultural rights as well.

3.2 How does HRMI measure economic and social rights?

HRMI uses two benchmarks against which to assess country performance on each of the five economic and social rights as well as the summary Quality of Life metric. These are the *income adjusted* benchmark and the *global best* benchmark. HRMI’s metrics using the *income adjusted* benchmark are measures of the extent to which countries are using their resources as effectively as possible to progressively fulfil their inhabitants’ substantive economic and social rights. In other words, we look at the extent to which

²⁵Randolph, Susan, Sakiko Fukuda-Parr, Terra Lawson-Remer, Ute Reisinger and John Stewart. *SERF Index Methodology 2017 Technical Note*, Economic and Social Rights Empowerment Initiative, 2017, www.serfindex.org.

the people in a country enjoy the substantive rights they are entitled to, taking into account how rich or poor the country is and therefore how well it ought to be able to ensure that food, housing etc. are accessible for its people. HRMI's metrics using the *global best* benchmark are measures of the extent to which countries are fulfilling their inhabitants' substantive economic and social rights relative to the best performing countries at any resource level. Thus, the two benchmarks assess country performance relative to their current and full obligations, respectively.

3.3 How is this different from the way HRMI measures civil and political rights?

HRMI measures these two groups of rights quite differently as is consistent with state obligations under international law. Under international law, the state must immediately and completely respect, protect, and fulfil all rights listed in the International Covenant for Civil and Political Rights, while the substantive rights listed in the International Covenant on Economic, Social, and Cultural Rights are to be *progressively realised using the maximum of available resources*. Thus HRMI's *income adjusted* economic and social rights metrics measure economic and social rights relative to the extent to which each country ought to be able to fulfil those rights for its people, given its current resources. Because countries are obligated to progressively ensure the full realisation of the substantive rights enumerated in the ICESCR, we also track the extent to which countries have achieved full realisation of the substantive economic and social rights concerned. Because progressive realisation does not apply to civil and political rights, our civil and political rights metrics are not adjusted to account for the resources available to a country.

A second important difference is that HRMI's economic and social rights metrics are calculated from objective, internationally comparable, publicly accessible statistical data published by national and international bodies. Our civil and political rights metrics, on the other hand, are calculated using surveys of human rights experts in each country. This is because objective statistical data that meets our standards, are not available for most civil and political rights. For more details on how we measure civil and political rights please see Section 2.

The methodology used to identify those population subgroups more likely to have their rights violated are the same for civil and political rights as well as economic and social rights and is described in section 2.4.1.

3.4 How does HRMI's economic and social rights methodology work?

Under international law, as noted above, countries are obligated to use "the maximum of [their] available resources" to progressively achieve "the full realization of the rights" specified in the Covenant

(International Covenant of Economic, Social, and Cultural Rights, Article 2.1). This means that at any given time each country has a different level of obligation and a given country's obligation increases over time as its resource capacity expands. The methodology used to construct our *income adjusted* metrics aims to assess the level of rights enjoyment achieved relative to the country's current level of obligation; that is, what the country could feasibly achieve in terms of fulfilling its people's rights given the level of resources it has. We do this by mapping an evidence-based achievement possibilities frontier to benchmark each country's obligation at any given time. The estimated achievement possibilities frontiers plateau at the resource where it is first becomes possible to ensure the highest level of rights enjoyment observed at any income level. This plateau value is the benchmark used to construct our *global best* metrics.

This methodology is unique in:

- Considering the perspective of both the rights-holder (i.e. the individual people) and the duty-bearer (i.e. the government) both currently and eventually;
- Making possible objective assessment of whether the overall situation in a country is improving or deteriorating both vis-a-vis the country's immediate and eventual obligations;
- Allowing cross-country comparisons of countries' fulfilment of their current and eventual economic and social rights obligations; and
- Providing a methodology to examine disparity in rights fulfilment between regions, or between racial and ethnic or other population sub-groups within a given country.

3.5 What do HRMI's economic and social rights scores show, exactly?

HRMI's economic and social rights scores show the **percentage of the benchmark achievement obtained**. In the case of the *income adjusted* benchmark, this is the percentage of the currently feasible achievement obtainable, **given the country's per capita income level**. A low score means a country is not fulfilling the rights concerned nearly to the extent that should be possible at its per-capita income level. A score of 100% on our *income adjusted* metrics does not mean everyone in the country enjoys the right; it means the country is doing as well at ensuring that right as the best performing country has at that per-capita income level. Thus, in the case of a very poor country, the economic and social rights metric score can be quite high, even though a lot of people in that country do not have proper access to food, housing, education, etc. In contrast, in the case of the *global best* metrics, HRMI's economic and social rights scores show the country's percentage achievement relative to the highest enjoyment level observed in any country at any per capita income level. A score of 100% on our *global best* metrics means the country is doing as well at ensuring the right as the best performing country has at any per-capita income

level. In the case of a very poor country, its scores on our *global best* metrics will be substantially lower than its scores on our *income adjusted* metrics. In contrast, for wealthy countries there will be little if any difference between their scores on the *income adjusted* and *global best* metrics.

3.6 What are HRMI’s two different assessment standards?

HRMI’s economic and social rights metrics use two separate assessment standards: our “low- and middle-income” assessment standard and our “high-income” assessment standard. The low-and middle- income assessment standard holds countries to a basic standard that reflects the challenges that low- and middle-income countries face. The high-income standard holds countries to a higher standard more reflective of the economic and social rights challenges that high-income countries face.

We have these two different assessment standards because richer countries, having more resources, are typically further advanced in making sure that their people are well fed, housed, educated, etc. So we need to use indicators that can capture the different challenges these countries face. For example, richer countries have often already achieved high education participation and their focus is on raising the *quality* of education. Although education quality is also critically important for less developed countries, the indicator for education quality is not available for most low- and middle-income countries. Scores using both standards are calculated for all countries where the data are available, enabling researchers to evaluate countries with the available data on either standard.

3.7 How are HRMI’s economic and social rights metrics constructed?

We construct HRMI’s economic and social rights metrics by following the steps below:

- Step 1: Identify indicators that broadly summarise the extent to which people enjoy each economic and social right, and which are available on an internationally comparable basis for a large number of countries in the world.
- Step 2: Specify how much a country ought to be able to fulfil its people’s rights given the country’s per capita income in the case of the *income adjusted metrics*, and once resource constraints no longer are binding in the case of the *global best* metrics, and compute indicator performance scores for each indicator reflecting the extent to which a country meets its obligations.
- Step 3: Combine indicator performance scores (using the benchmark concerned) into aggregate metrics for each of the five economic and social rights and average the result to obtain the Quality of Life score.

3.8 How does HRMI choose which indicators to use?

We use a number of criteria when selecting which data will be the best indicators of economic and social rights fulfilment, including:

- How well the indicator reflects enjoyment of the right (concept validity);
- Reliability of the data;
- Objectivity of measurement methods;
- Comparability across countries and over time;
- Public accessibility;
- Data availability vis-a-vis country coverage and frequency of collection, and;
- The extent of variation among countries.

Indicator sets are selected to:

- Reflect the challenges most relevant to fulfilling a given right, rather than to encompass all aspects of a given right;
- Prefer those specifying the percentage of the population enjoying the right over those indicating the average level of enjoyment of the right across the population. This is because high levels of enjoyment on the part of some people can hide the denial of the right to many;
- Prefer indicators of flow variables to indicators of stock variables, since they give us a more up-to-date picture of the human rights situation; and
- Prefer bell weather indicators sensitive to a variety of factors related to rights fulfilment.

We attempt to keep the number of indicators of a given right to three, because our goal is to provide a summary measure of performance that is comparable across countries and can show trends over time. Our selection of indicators is practically constrained by:

- **Availability:** Because the surveys providing many of the indicators on enjoyment of rights are not conducted annually, the data used for each year are not always unique. For example, in the case of the Right to Education metric for Benin, the 2014, 2015, and 2016 series use data on the adjusted net primary school enrolment rate in 2014.
- **Relevance:** Ensuring all students complete primary school is not an issue for OECD countries, so although this is an indicator we use in our low- and middle- income assessment standard, it is not an indicator used in our high-income assessment standard.

Table 1: Rights enjoyment indicator sets used in HRMI economic and social rights metrics

Economic and social right	Indicator
<ul style="list-style-type: none"> Assessment standard 	
Food	
<ul style="list-style-type: none"> Low- and middle- income High-income 	<ul style="list-style-type: none"> % children (under 5) not stunted % population food secure (based on the food insecurity experience scale, FIES)
Education	
<ul style="list-style-type: none"> Low- and middle-income Both High-income 	<ul style="list-style-type: none"> Adjusted net primary school enrolment rate Net secondary school enrolment rate % students scoring level 3 or better on the PISA test (by topic-- math, science and reading)
Health	
<ul style="list-style-type: none"> Low- and middle- income Both Both High-income 	<ul style="list-style-type: none"> Modern Contraceptive use rate Child (under 5) survival rate Age 65 survival rate % live births not low birth weight
Housing/Water/Sanitation	
<ul style="list-style-type: none"> Low-and middle-income Low-and middle income High-income 	<ul style="list-style-type: none"> % population with "basic" access to water on premises % population with access to "basic" sanitation % population with "safely managed" sanitation
Decent Work/Social Security	
<ul style="list-style-type: none"> Low- and middle-income High-income High-income 	<ul style="list-style-type: none"> % with income >\$3.20 (2011 PPP\$**) per day % with income > 50% median income % unemployed not long-term unemployed

*PISA is the Program for International Student Assessment that implements the surveys of student learning outcomes that we use.

**PPP\$ means purchasing power parity dollars. This means that currency conversions between countries have been made using conversion factors that equate to the actual purchasing power of the currencies rather than using official exchange rates. The prices used are those prevailing in 2011

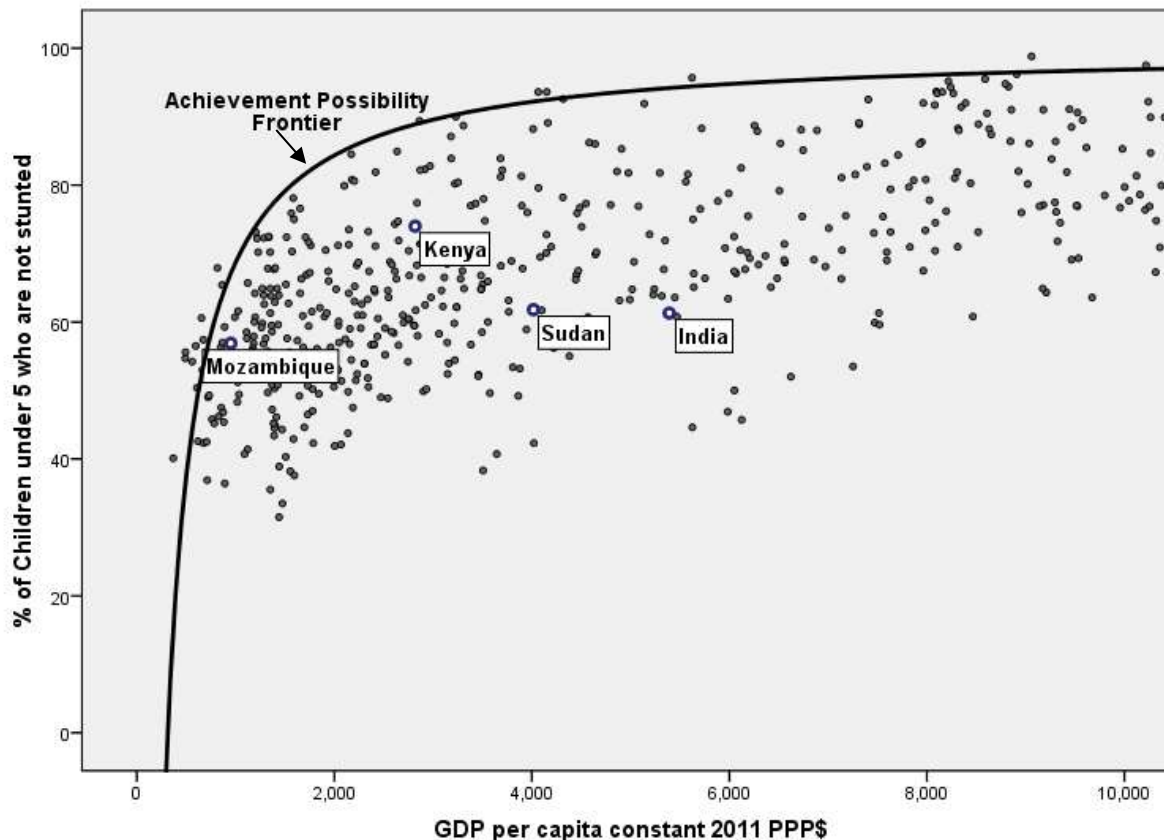
3.9 What is HRMI's achievement possibilities frontier?

This is a tool HRMI uses to assess what is feasible for countries to achieve, in terms of their ability to deliver on economic and social rights for their people. This is done by seeing what has been achieved by other countries over history and at different levels of available resources.

The achievement possibilities frontier (APF) for a given indicator is constructed by plotting the observed value of the indicator against per capita GDP (2011 PPP\$) for all countries over the 1995 to

2016 period. The frontier is defined as the outer envelope of the scatter plot, and the equation specifying the frontier is estimated by fitting a curve to the observations that define the outer boundary of the scatter plot. See Section 4 for detailed interpretation.

Figure 1: Achievement Possibilities Frontier for “Percentage of Children Not Stunted”



The approach to assessing a country’s performance using the *income adjusted* benchmark is to compare the country’s actual performance to the feasible performance as benchmarked by the achievement possibilities frontier. For example, India’s child stunting rate in 2014 was 38%, implying the percentage of children not stunted was 62%. However, at its per-capita GDP of \$5,391 (2011 PPP\$), it should be possible to ensure that 94% of Indian children under 5 are not stunted. So our first cut at assessing India’s performance on the right to food using the *income-adjusted* benchmark takes the ratio of the observed percentage of children that are not stunted (62%) to the benchmark percentage of children not stunted (94%) and then multiplies by 100 to yield the percentage of the feasible level achieved.

After that some final steps in our calculations are still needed. Since the plausible range of indicators varies, we also need to standardise scores by taking into account how close the lowest observed value is to zero. In the case of our right to food indicator, the lowest value observed is 31% (the percentage of children not stunted in Bangladesh in 1995). We therefore standardise the scores by computing the percentage of the feasible level achieved with reference to the minimum observed score. So, looking again at India, its achievement relative to this minimum observed score is $62\% - 31\% = 31\%$ of children not stunted. Relative to the minimum, it is feasible for India to achieve $94\% - 31\% = 63\%$ of children not stunted. Thus, India’s score on the Right to Food is calculated as $(31\%/63\%) \times 100 = 49.2\%$.

The *global best* benchmark sets the benchmark for all countries at the maximum height of the estimated APF or at the income level at which all people enjoy the right. Since 2.3 percent of children that are growing normally are expected to be more than two standard deviations below the mean height for their age the *global best* benchmark is set at \$13608 (2011 PPP\$), the income level where the APF reaches 97.7% of children are not stunted. India's *global best* right to food score is calculated as before, only substituting the *global best* benchmark for the *income adjusted* benchmark. $(62\% - 31\%)/(97.7\% - 31\%) = 46.5\%$. So long as resources are binding, country scores using the *global best* benchmark will be lower than those using the *income adjusted* benchmark.

In the case of some richer countries, HRMI's economic and social rights metrics also take into account the fact that some countries have many times the resources needed to ensure that all people enjoy a given right, yet still fail to make sure that everyone enjoys the rights to which they are entitled. For example, Oman and Mexico have nearly an identical percentage of children that are not stunted (86.4% for Mexico and 85.9% for Oman), yet Oman's per-capita income is nearly 2.5 times higher than Mexico's. For countries like Oman with per-capita income levels multiple times what is needed to reach the frontier, but who still fail to do so, we impose a penalty on their *income adjusted* and *global best* scores.

4 HRMI Economic and Social Rights Metrics 2018 Technical Note²⁶

Susan Randolph, Sakiko Fukuda-Parr, Terra Lawson-Remer, Ute Reisinger, and John Stewart

This technical note provides a detailed explanation of the methodology used to construct the Human Rights Measurement Initiative's (HRMI's) 2019 economic and social rights metrics (and future updates that use the same methodology). HRMI's economic and social right metrics are adopted from the International Social and Economic Rights Fulfilment Index (SERF Index) and were developed by Susan Randolph, Sakiko Fukuda-Parr, and Terra Lawson-Remer. As with most measurement initiatives, the SERF Index methodology has evolved to take account of emerging conceptual and data issues. The International SERF Index has been refined four times since it was initially published in 2009. HRMI's 2019 economic and social rights metrics are the underlying Right Indices that comprise the 2019 Update of the International SERF Index scores and cover the years 2006 to 2016.

The book, *Fulfilling Social and Economic Rights* by Sakiko Fukuda-Parr, Terra Lawson-Remer and Susan Randolph (Oxford: Oxford University Press, 2015) provides a detailed account of the basic SERF Index methodology and insights gained from its application that is accessible to practitioners. The conceptual and methodological underpinnings of the SERF Index are also fully elaborated in two peer reviewed publications:

- Fukuda-Parr, Sakiko, Terra Lawson-Remer and Susan Randolph (2009) 'An Index of Economic and Social Rights Fulfillment: Concept and Methodology.' *Journal of Human Rights*. 8: 195-221. (<http://www.informaworld.com/smpp/title~db=all~content=g914018350>)
- Randolph, Susan, Sakiko Fukuda-Parr, Terra Lawson-Remer (2010) 'Economic and Social Rights Fulfillment Index: Country Scores and Rankings.' *Journal of Human Rights*, 9.3, 230-261. (<http://www.informaworld.com/smpp/title~db=all~content=g926038290>)

²⁶ This technical note is adapted from Randolph, S., S. Fukuda-Parr, T. Lawson-Remer, U. Reisinger and J. Stewart, "SERF Index Methodology: 2017 Technical Note (Economic and Social Rights Empowerment Initiative, 2017), www.serfindex.org.data with permission from the Economic and Social Rights Empowerment Initiative. Refinement of the SERF methodology was supported in part by the National Science Foundation under grant number 1061457. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

4.1 Overview

HRMI's economic and social rights (ESR) metrics (or scores) measure the performance of countries on the fulfilment of key economic and social rights obligations. HRMI's metrics use objective, internationally comparable, publicly accessible statistical data published by national and international bodies. HRMI's ESR metrics provide summary scores for human rights that are grounded in international law. The International Covenant for Economic, Social, and Cultural Rights (ICESCR) articulates a list of essential substantive economic and social rights that the 166 nations, representing a wide range of cultural traditions, who have ratified it concur are essential. These are the rights to food, health, education, housing, work, and social security. HRMI's ESR metrics cover five out of six of these rights. We don't yet have sufficient internationally comparable data to independently include social security. However, the indicators used to measure the right to work also capture key elements of the right to social security; available data just do not enable a full separation between the right to work and the right to social security.

A fundamental principal of international law is that countries have a duty to progressively realise economic and social rights to the maximum of their available resources. Statistics like school enrolment and infant mortality tell us only the extent to which individuals enjoy economic and social rights, but not whether a state is complying with its obligations to progressively respect, protect, and fulfil human rights. Measuring economic and social rights fulfilment requires considering the perspectives of both the rights-holding individual and the duty-bearing government. While many widely available socio-economic indicators and other metrics, such as the Human Development Index (HDI) assess the level of rights enjoyment, they ignore the obligation level of the duty bearing state. HRMI's ESR methodology estimates obligations for progressive realisation by using an innovative approach that maps an evidence based 'achievement possibilities frontier' (APF) to benchmark each country's obligation at any given time. This methodology is the only ESR metrics methodology that:

- Considers the perspective of both the rights-holder and the duty-bearer measuring state compliance with obligations of progressive realisation;
- Makes possible objective assessment of whether the overall situation in a country is improving or deteriorating;
- Allows cross-country comparisons of rights fulfilment; and
- Provides a methodology to examine disparity in rights fulfilment between regions, or between racial and ethnic or other population sub-groups.

The HRMI *income-adjusted* metrics measure a country's achievement relative to what it is feasible to achieve at the country's per capita income level. That is, they look at the enjoyment level of a right relative to the best practice in countries with approximately the same per capita income, the *income-*

adjusted benchmark level of rights enjoyment. More specifically, the HRMI *income-adjusted* scores show the percentage of the feasible achievement obtained, given the country's per capita income level. A low score means a country is not fulfilling the right concerned to the extent possible at its per capita income level. An *income-adjusted* score of 100% on a right or right aspect **does not mean** everyone in the country enjoys the right; it means the country is doing as well at ensuring the right as the best performing countries at that per capita income level. Thus, in the case of a very poor country, the score on the right can be quite high even though the enjoyment level of the right is quite limited. A country achieving an *income-adjusted* score of 100% cannot rest on its laurels. All countries are obligated to progressively fully realise the rights enumerated in the ICESCR. As a result, HRMI also assesses country performance relative to what evidence shows is feasible once income is no longer a constraint, what we call the *global best* benchmark. A country with a high *income-adjusted* score may simultaneously have a low *global-best* score on the same right or right aspect indicating that although it is meeting its immediate obligation under Article 2.1 of the ICESCR to fulfil that right to the "*maximum of its available resources*", it still has a long way to go fully meet its obligation to fulfil the right. Some countries have more than enough income to reach the *global best* benchmark but fail to do so. For these countries a penalty is imposed on the countries' scores; the penalty is larger the lower the enjoyment level of the right and the greater the country's income beyond that necessary to fully fulfil the right. As a result, in the case of a country with a high per capita income, the country's score on a right or right aspect could well be lower than the raw indicator score reflecting the enjoyment level of the right or right aspect.

Data constraints coupled with the different rights challenges in high income countries versus other countries have led to our creation of two separate assessment standards:

- The "low and middle income" assessment standard holds countries to a basic level of rights fulfilment, and
- The "high-income" assessment standard holds countries to a higher standard more relevant to the right challenges facing high-income countries.

Scores using both standards are calculated for all countries with available data, enabling researchers to evaluate countries with the available data on either standard. HRMI's ESR metrics are comparable across time for each country, as well as between countries. When computing a country's score on a right, the most recently available data on a given right enjoyment indicator (and the per capita income data for the corresponding year) is used. However, because the surveys providing many of the indicators on enjoyment of rights are not conducted annually, the data used for each year are not always unique. For example, in the case of the Right to Education metric for Benin, the 2014, 2015, and 2016 series use data

on the adjusted net primary school enrolment rate in 2014. If the most recently available data on an indicator is more than 10 years prior, the score for that right is recorded as “missing”.²⁷

The construction of HRMI’s ESR metrics is further elaborated below. In addition to the ESR metrics described below, HRMI collects information on which population subgroups are most likely to be unable to claim the different economic and social rights for the 19 countries that were part of our CPR survey. The methodology used to identify these “groups at risk” is described in section 2.4.1. above.

4.2 Sources and definitions of rights and obligations

*The International Covenant of Economic, Social, and Cultural Rights (ICESCR)*²⁸ commits governments to achieve realisation of economic, social and cultural rights *progressively*. As stated in Article 2.1:

“Each State Party to the present Covenant undertakes to take steps, individually and through international assistance and co-operation, especially economic and technical, to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant, by all appropriate means including particularly the adoption of legislative measures.”

The ‘*progressive realisation*’ provision recognises that states have very different starting points in their ability to achieve full enjoyment of economic and social rights, as noted by Fukuda-Parr, et al. (2015)

“Countries around the world face hugely different levels of deprivation and capacity. Inherent in the idea of progressive realization is that a government’s ability to fulfill rights commitments depends on the level of resources (financial and other) available in the country.”²⁹

The enjoyment of the right to the highest attainable standard of health, for example, cannot be achieved overnight, as facilities need to be built, personnel trained, and policy incentives for businesses and households put in place and so on, for people to have access to healthcare. These arrangements require financial resources which may be beyond what governments and households can currently mobilise. Consequently, the performance of states with regard to progressively realising economic and social rights cannot be judged on the basis of outcomes – enjoyment of rights by people – alone. For example, the performance of the United States and Malawi cannot be compared on the basis of their

²⁷ Downloadable excel files with information on the “most recent data year” for each indicator used in the construction of each right index for each year are available at www.serfindex.org/data. Researchers who prefer a less generous look back period can use the files from the 2017 Update of the International SERF Index to recode observations they consider too old as missing.

²⁸United Nations (1966). International Covenant on Economic, Social and Cultural Rights (ICESCR). Adopted 16 December 1966, General Assembly Resolution 2200 (XXI), U.N. GAOR, 21st Session, Supp. No. 16, U.N. Document A/6316 (1966), 993 U.N.T.S. 3 (*entered into force* 3 January 1976).

²⁹ Fukuda-Parr, Sakiko, Terra Lawson-Remer, and Susan Randolph, *Fulfilling Social and Economic Rights* (Oxford: Oxford University Press, 2015, p. 11).

respective levels of child survival rates considering the hugely different levels of capacity in these two countries.

Thus, a country's performance in fulfilling obligations for economic and social rights depends on:

- the actual economic and social rights (ESR) outcomes people enjoy, as indicated by socio-economic statistics that proxy for particular rights; and
- a society's capacity for fulfilment, as determined by the amount of economic resources available overall to the duty-bearing state.

The provision of progressive realisation has complicated and frustrated efforts to monitor countries' fulfilment of their economic and social rights obligations, since, as Human Rights measurement scholar Chapman notes:

“it necessitates the development of a multiplicity of performance standards for each right in relationship to the varied... contexts of specific countries”.³⁰

That is, measures of ESR outcomes must reflect variable local specificities. The monitoring procedure adopted by the Committee on Economic, Social and Cultural Rights assesses performance relative to 'benchmarks'. But this leaves the problem of setting the benchmark. In the absence of a conceptual and evidence-based model for setting benchmarks, States enjoy considerable discretion over where their benchmark is set, thus effectively leaving open an 'escape hatch' for States to avoid meeting their ESR obligations.

HRMI's ESR metrics overcome this problem. The innovation of the methodology lies in the construction of Achievement Possibilities Frontiers (APFs) that use an evidence-based approach to specify each country's level of obligation *for progressive realisation* with regard to various aspects of each economic and social right. The basic construction of HRMI's Right metrics involves the following steps:

- Identify indicators that broadly summarise: i) the enjoyment level of the substantive rights articulated in international law and ii) country resource capacity.
- Specify country obligations with regard to each of the selected indicators and compute indicator scores reflecting the extent to which a country meets its obligations on each aspect of the right.
- For each substantive right, aggregate the indicator scores for the different right aspects or the right into a right score by averaging the indicator scores.

³⁰Chapman, Audrey. 'The Status of Efforts to Monitor Economic, Social, and Cultural Rights,' in *Economic Rights: Conceptual, Measurement and Policy Issues*, eds. Shareen Hertel and Lanse Minkler (Cambridge: Cambridge University Press, 2007). Chapter 7, p 150.

For most ESRs resource constraints do not bind throughout the income range observed globally. Once resources no longer impose a constraint, countries are obligated to fully fulfil the substantive rights guaranteed under the ICESCR.³¹ For this reason, HRMI uses two different benchmarks against which to assess country performance, the *income-adjusted* benchmark and the *global best* benchmark. HRMI's metrics using the *income adjusted* benchmark are measures of the extent to which countries are using their resources as effectively as possible to progressively fulfil their inhabitants' substantive economic and social rights. In other words, we look at the extent to which the people in a country enjoy the substantive rights they are entitled to, taking into account how rich or poor the country is and therefore how well it ought to be able to ensure that food, housing etc. are accessible for its people. HRMI's metrics using the *global best* benchmark are measures of the extent to which countries are fulfilling their inhabitants' substantive economic and social rights relative to the best performing countries at any resource level. Thus, the two benchmarks assess country performance relative to their current and full obligations, respectively.

4.3 Measuring economic and social rights enjoyment and state resources

4.3.1 Sources and definitions of rights and obligations

HRMI ESR metrics draw on international law – the Universal Declaration of Human Rights³² (UDHR), ICESCR³³ and numerous other international human rights legal instruments³⁴ – to define the substantive rights of individuals and the obligations of states. The substance of these rights is detailed in General Comments of the Committee on Economic, Social, and Cultural Rights (CESCR).³⁵

³¹ United Nations (1966).

³²United Nations (1948). *Universal Declaration of Human Rights (UDHR)*, Adopted 10 Dec. 1948, United Nations General Assembly Res. 217 A (III), (1948).

³³ United Nations (1966).

³⁴ These international legal instruments include the General Comments of the relevant treaty body committees, reports of Special Rapporteurs, and other documents such as reports of seminars, task forces and working groups.

³⁵Committee on Economic Social and Cultural Rights.(1991)'General Comment 4:The Right to Adequate Housing',6thSession, 13 December;(1997) 'General Comment 7: The Right to Adequate Housing—Forced Evictions', 16thSession, 20 May; (1999a) 'General Comment 11:Plans of Action for Primary Education', 20thSession, Geneva, 26 April – 14 May 1999, Document E/C.12/1999/4; (1999b) 'General Comment 12: The Right to Adequate Food', 20thSession, Geneva, 26 Apr – 14 May, Doc. E/C.12/1999/5; (1999c) 'General Comment 13: The Right to Education',21stSess. 15 November – 3 December 1999, Document E/C.12/1999/10; (2000) 'General Comment 14: The Right to the Highest Attainable Standard of Health', 22nd Session, 25 April – 12 May 2000, Document E/C.12/2000/4 ; (2005) 'General Comment18:The Right to Work', 35th Session, 7-25 November 2005, Document E/C.12/GC/18, 6 February 2006;(2008) 'General Comment 19:The Right to Social Security", 39th Session, 5-23 November. Document E/C.12/GC/19, 4 February 2008.

The General Comments identify seven substantive economic and social rights; the right to:

- adequate food,
- education,
- highest attainable standards of physical and mental health,
- adequate housing,
- water and sanitation,
- decent work, and
- social security.

Following the Office of the High Commissioner for Human Rights 2012 guidelines on using indicators to monitor human rights, we collapse these into six rights, in view of the fact that access to water and sanitation are key components of the right to housing.³⁶

States bear the primary responsibility for the realisation of the rights of citizens and individuals residing within their borders. Their obligations are threefold: *to respect, to protect, and to fulfil* rights. These obligations also include the cross-cutting *procedural rights of non-discrimination, participation, and accountability*. General Comments 3³⁷ and 9³⁸ along with the *Limburg Principles*³⁹ and *Maastricht Guidelines*⁴⁰ elaborate the nature and extent of the obligations accepted by State parties to the Covenant.

HRMI's ESR metrics measure State parties' compliance with their obligations for progressive realisation of economic and social rights, focusing on outcomes reflected in enjoyment of the substantive rights by people and in the case of HRMI's *income*-adjusted scores, adjusted for state capacity. They do not attempt to assess the extent to which States ensure the *procedural rights* of non-discrimination, participation, and accountability. HRMI's ESR metrics complement other measurement tools such as those suggested by the Office of the High Commissioner for Human Rights.⁴¹ These and other recent initiatives, such as the Right to Education Index⁴² focus on different aspects of obligations, such as

³⁶United Nations Office of the High Commissioner for Human Rights (2012). *Human Rights Indicators: A Guide to measurement and implementation*. HR/PUB/12/5. New York: Office of the High Commissioner for Human Rights, United Nations.

³⁷ Committee on Economic, Social and Cultural Rights (1990) 'General Comment 3: The Nature of States Parties' Obligations', 5th Sess., December 14.

³⁸ Committee on Economic, Social and Cultural Rights (1998) 'General Comment 9: The Domestic Application of the Covenant' 19th Session, 16 November – 4 December, Document E/C.12/1998/24, 3 December 1998.

³⁹United Nations (1987). *The Limburg Principles on the Implementation of the International Covenant on Economic, Social and Cultural Rights*. Guidelines adopted at a workshop sponsored by the International Commission of Jurists, the Faculty of Law of the University of Limburg, and the Urban Morgan Institute for Human Rights, University of Cincinnati, Maastricht, Netherlands, 22-26 January 1997, Document E/CN.4/1987/17.

⁴⁰United Nations (2000). *The Maastricht Guidelines on Violations of Economic, Social and Cultural Rights*. Guidelines adopted at a workshop sponsored by the International Commission of Jurists, the Urban Morgan Institute for Human Rights and the Center for Human Rights of the Faculty of Law of Maastricht University, Maastricht, Netherlands, 22-26 January, 1997. Document E/C.12/2000/13.

⁴¹United Nations Office of the High Commissioner for Human Rights (2012). For comparison of SERF with other proposals, see Randolph et al, *Journal of Human Rights* 2010, and Fukuda-Parr, Sakiko, 'The Metrics of Human Rights: Complementarities of Human Rights and Capabilities Approach', *Journal of Human Development and Capabilities*, Vol. 12:1 pp73-89.

⁴² See http://www.results.org/issues/global_poverty_campaigns/right_to_education_index/.

process (or policy efforts made by government), structure (institutionalised provisions), and outcomes (level of rights enjoyment in the population), while assessing performance on 50 to 100 aspects of each right. However, none attempts to provide a broad summary of performance and benchmark outcomes according to the obligation of progressive realisation as HRMI’s ESR metrics do.

4.3.2 Selecting the indicators of rights enjoyment and resource capacity

A number of criteria govern the selection of the indicators. Beyond making sure selected indicators appropriately reflect enjoyment of the right concerned and resource capacity, selected indicators must be:

- based on reliable data;
- measured with objective methods;
- legitimately comparable across countries and over time; and
- publicly accessible.

To satisfy these criteria, all data sets used to construct HRMI’s ESR metrics are international series that are maintained by international organisations. Further considerations for indicator selection include:

- data availability and country coverage;
- frequency of data collection;
- the extent of variation among countries;
- ability to reflect the challenges most relevant to fulfilling a given right, rather than to encompass all aspects of a given right;
- indicators specifying the percentage of the population enjoying the right are preferred to those indicating the average level of enjoyment of the right across the population;
- indicators of flow variables are preferred to indicators of stock variables; and
- preference is given to bellwether indicators sensitive to a variety of factors related to rights fulfilment.

In general, we have sought to keep the number of indicators reflecting different key aspects of a given right down to three.

Our selection of indicators is also practically constrained by current data availability. This, plus different rights challenges in high income OECD countries versus most other countries led to our creation of two separate sets of scores using two different assessment standards: one standard relevant to the majority of countries, our “low- and middle- income” assessment standard, and the other most relevant to high income countries, our “high-income” assessment standard. For example, the high-income assessment standard includes a measure of the quality of schooling, performance on the Program for International Student Assessment (PISA) math, science and reading tests, among the education indicators. The quality of education is no less a concern for all other countries, it’s just that there is no measure with broad coverage available at this time for low- and middle- income countries. Regarding relevance, ensuring all

students complete primary school is not an issue for high-income OECD countries, so although this is an indicator we use in our low- and middle-income assessment standard, it is not included in our high-income assessment standard.

Data limitations currently preclude defining separate metrics for all six rights. HRMI's ESR metrics include separate scores for five rights—the rights to food, health, education, housing, and work—with key elements of the right to social security captured by the right to work and cover the years 2006 through 2016. Available data do not enable us to fully separate the right to work from the right to social security at this time. In the case of HRMI's ESR metrics for the right to food using the high-income assessment standard data are only available for 2015 and 2016. We have found it necessary to use two different assessment standards given the differences in data availability and current rights challenges between the two groups of countries. However, right scores using both standards are calculated for all countries (low- and middle-income as well as high-income countries) with available data, enabling researchers to evaluate countries with the available data on either standard. Table 2 below shows the indicators currently used to measure enjoyment of key aspects of each right for each of the two assessment standards.⁴³

Appendix Table A gives details of sources and definitions for each indicator. A detailed discussion of why particular indicators were selected is provided in Fukuda-Parr, Lawson-Remer, and Randolph (2015). As noted at the outset, States are required to fulfil economic and social rights *progressively*, and to commit the *maximum of available resources* to meet this obligation. HRMI ESR metrics use per capita GDP as the indicator of State resource capacity measured in 2011 purchasing power parity (PPP) dollars. While it might be argued that States with larger budgets or better institutions have a greater capacity to fulfil economic and social rights than those with the same per capita income but smaller budgets or poorer institutions, a State's capacity depends on the choices it makes with regard to its taxing policies and institutional structure. Since the obligation to progressively realise economic and social rights requires States to collect and expend resources at the level necessary to meet their rights obligations, it is appropriate to measure resource capacity as reflected by the total resources available to the State, not the portion of those resources the State chooses to tap. We measure GDP per capita data in 2011 international purchasing power parity dollars (2011 PPP\$) to standardise for inflation and purchasing power across countries and thus enable comparison over time and across countries.⁴⁴

⁴³ In response to feedback from a wide range of scholars and practitioners, some of the indicators used to construct the SERF Index—and accordingly HRMI's ESR metrics—have been refined in the current version of the SERF Index and differ from those reported in Randolph, Fukuda-Parr and Lawson-Remer (2010) and Fukuda-Parr, Lawson-Remer, and Randolph (2015) and HRMI's 2017 data set.

⁴⁴ Purchasing power parities (PPPs) are the rates of currency conversion that equalise the purchasing power of different currencies by eliminating the differences in price levels between countries. The year 2011 is the most recent survey year of the International Comparison Project that estimates PPP\$ and accordingly the PPP\$ prices are the prices prevailing in 2011. See for example http://siteresources.worldbank.org/ICPEXT/Resources/ICP_2011.html for more information.

Table 2: Rights enjoyment indicators used to construct HRMI's ESR metrics

Human Right/Indicator	Assessment Standard	
	Low- and Middle income	High-income
Right to food		
% children (under 5) not stunted	√	
% people food secure		√
Right to education		
Net secondary school enrolment rate	√	√
Adjusted net primary school enrolment rate	√	
% students achieving level 3 or better on math, science and reading PISA test		√
Right to health		
% Children (under 5) surviving to age 5	√	√
% people surviving to age	√	√
% couples (15-49) using modern contraceptives	√	
% new-borns not low birth weight (<2500 grams)		
Right to housing		
% population with basic sanitation	√	
% population with basic water on premises	√	
% population with safely managed sanitation		√
Right to work		
% population with income > \$3.20 (2011 PPP\$) per day	√	
% population with income > 50% median income		√
% unemployed not long-term (>12 months) unemployed		√

4.4 Calculating indicator scores by benchmarking a country's obligations of progressive realisation

Achievement Possibility Frontiers (APFs) use an evidence based approach to benchmark each country's obligation with regard to each indicator reflecting the different aspects of each right. The APFs reflect what is feasible to achieve when a country allocates the maximum of available resources to fulfilling economic and social rights and uses those resources effectively as is evidenced by the experience of the best performing countries at different per capita GDP levels. The frontiers are constructed so as to be stable over the medium term thus enabling inter-temporal comparison.⁴⁵ Specifically, the APF for a given indicator is constructed by plotting the observed value of the indicator against per capita GDP (2011

⁴⁵ Although knowledge of how to transform resources into rights enjoyment will change over time, rapid and abrupt changes in best practice technology are unlikely.

PPP\$) for each country over the 1995 to 2016 period.⁴⁶ The frontier itself is defined as the outer envelope of the scatter plot, and the equation specifying the frontier is estimated by fitting a curve to the observations that define the outer envelope of the scatter. While earlier versions of the SERF Index measured GDP per capita in 2005 PPP\$,⁴⁷ as is the case for the 2017 and 2019 updates of the SERF Index, HRMI's ESR metrics utilise the re-estimated frontiers based on the more recently available 2011 PPP\$ exchange rates which is based on a broader survey coverage than the 2005 PPP\$ series and has an improved methodology. Appendix Table B shows the equations specifying the frontier for each indicator.⁴⁸

To better understand the process, consider the construction of HRMI's Right to Food Score using the low-and middle-income assessment standard. The first step, as discussed above, is to figure out the best statistical indicators to monitor. Some of HRMI's metrics use multiple indicators, but only a single right enjoyment indicator is used in constructing HRMI's right to food score for the low- and middle-income assessment standard—a measure of child malnutrition prevalence. Specifically, as shown in Table 2, we use the percentage of children under 5 years of age who are not stunted, that is, whose height is not unusually low relative to the median (precisely, not more than 2 standard deviations below the median). These data come from the World Health Organization's Global Database on Child Growth and Malnutrition. The stunting rate is a bellwether indicator of family malnutrition. It has been found to be more sensitive to both chronic caloric insufficiency and a diet chronically lacking in adequate protein and micronutrients and is less likely to be influenced by temporary illness than other measures of child under-nutrition. Also, because parents tend to protect the nutritional wellbeing of their children over their own, the child stunting rate also reflects the inability of parents to adequately ensure their own nutritional wellbeing. Because our focus is on rights enjoyment, we subtract the child stunting percentage from 100%. We then construct a scatter plot of the percentage of children under 5 who are not stunted against GDP per capita (2011PPP\$) using all available country observations from 1995 to 2016.

These data are shown in Figure 7 below, where each black dot is a single country observation for a particular year. The most recent observations available for Mozambique, Kenya, Sudan, and India are highlighted. As can be seen there is a substantial spread between the best and worst performing countries at each per capita GDP level. We use econometric techniques to fit a curve to the outer-boundary of the

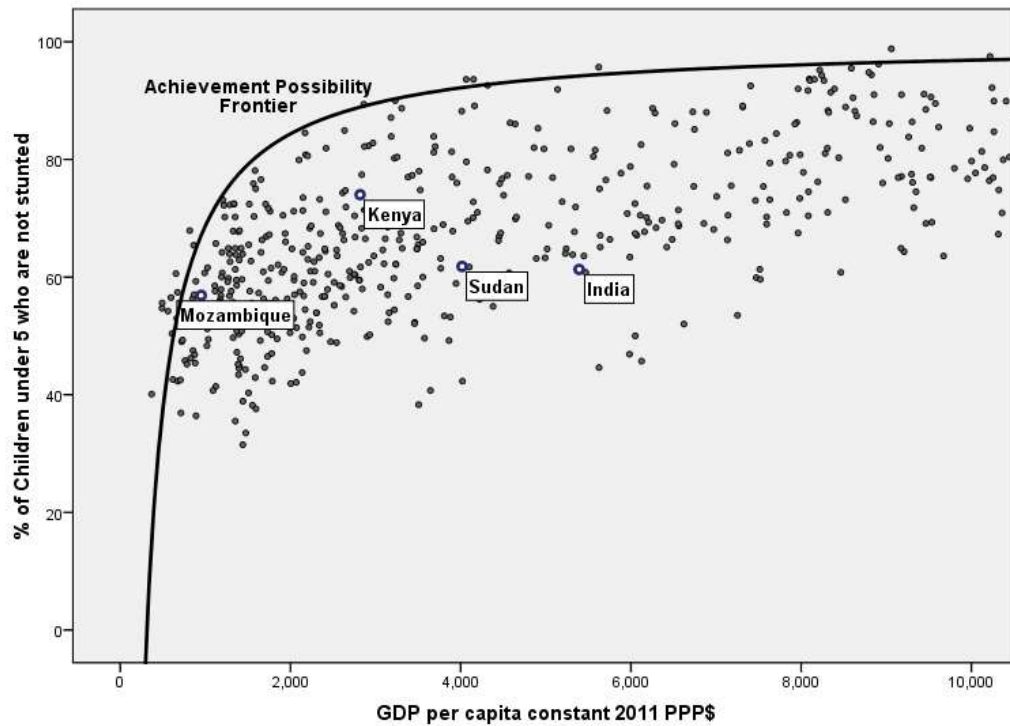
⁴⁶ The APFs for HRMI's ESR metrics were constructed in 2018 using all data available at that time since 1995.

⁴⁷The book, Fukuda-Parr, Lawson-Remer, and Randolph (2015) and two papers, Fukuda-Parr, Lawson-Remer, and Randolph (2009), and Randolph, Fukuda-Parr and Lawson-Remer (2010) further detail the basic methodology, although the 2019 version of the International SERF Index, the version upon which HRMI's ESR metrics are based, incorporates some additional refinements as indicated in this technical note.

⁴⁸ To guard against measurement error and ensure that the frontiers reflect what is reasonably achievable, observations from countries engaged in civil war at the time of the observation were eliminated, and for purposes of estimating the frontier, the per capita income corresponding to observations occurring in the wake of the Post USSR transition when per capita income levels in many of the former Soviet Republics and Eastern European countries briefly and temporarily plummeted were reset to the per capita income level just prior to the start of the transition until per capita income levels recovered. See Fukuda-Parr, Lawson-Remer, and Randolph (2015, 2009), and Randolph, Fukuda-Parr and Lawson-Remer (2010) for further details.

scatter plot (the solid black curve in Figure 7). This fitted curve is the Achievement Possibilities Frontier (APF). Based on country experience, it provides a benchmark for each per capita income level of the percentage of children it is feasible to ensure are not stunted. The APF defines the level of a State's immediate obligation for any given per capita GDP level (2011 PPP\$).

Figure 7: Achievement Possibilities Frontier for “Percentage of Children Not Stunted”



4.4.1 Assessing state performance: the *income-adjusted* indicator score

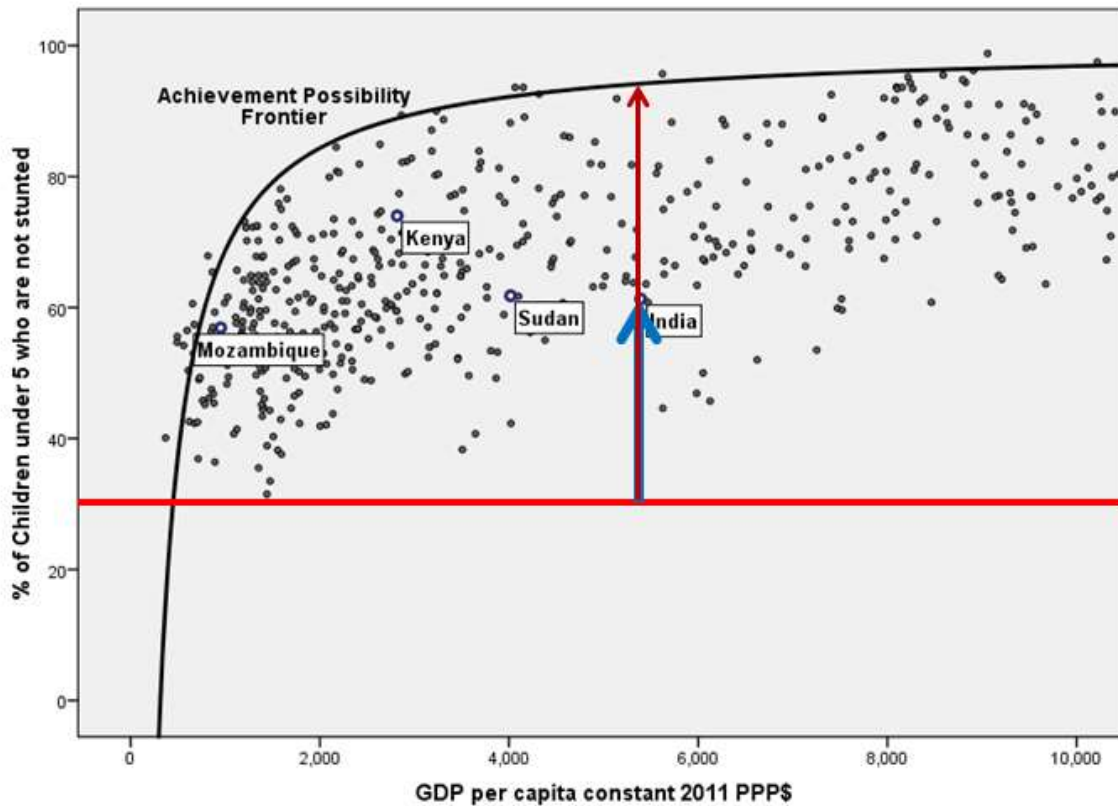
Ignoring, for the moment, some critical refinements, the approach to assessing State performance with regard to its immediate obligation is to compare the State's actual performance to the feasible performance as benchmarked by the APF at the country's current GDP per capita. So again, looking at Figure 7, India's child stunting rate in 2014 (the most recent year data were available for India when the frontier was estimated) was 38%, implying the percentage of children not stunted was 62%. However, at its then per capita GDP of \$5,391 (2011 PPP\$), it should be possible as shown by the APF to ensure 94% of children under 5 are not stunted. Thus our first cut at assessing India's performance using the *income-adjusted* benchmark is to take the ratio of the observed percentage of children that are not stunted (62%) to the benchmark percentage of children not stunted (94%) and then multiply by 100 to yield the percentage of the feasible level achieved.

Two things should be noted about Figure 7 above. First, the observed percentage of children that are not stunted never reaches a value approaching zero. In fact, the lowest value observed is 31%, the

percentage of children not stunted in Bangladesh in 1995. The observed minimum score differs widely across indicators. For example, the minimum observed score for the child survival rate (100% - % child mortality rate) is 68% (Niger in 1990) and that for the percentage of the rural population with access to improved rural water is 0% (Cambodia and Mozambique in 1990). Given that we are comparing multiple indicators in the construction of HRMI's ESR metrics, we need to standardise these indicators for two reasons. First, if we fail to do so our scores will not be comparable across rights and indicators with a larger actual range will drive right scores comprised of more than one aspect. Second, we need to take into account the fact that even in the absence of any focus on rights, certain indicators, such as the child survival rate, would have positive values while positive scores on other indicators, such as access to an improved water source, or primary school completion rates, substantially depend on public provision of goods and services and could be zero or close to zero.

We standardise the scores by computing the percentage of the feasible level achieved with reference to the minimum observed score on the indicator in the case of those indicators that do not substantially depend on public provision of goods and services. In Figure 2 below, the red horizontal line shows the minimum observed score of 31% on the child not stunted rate. So, for India, its achievement relative to this minimum observed score is $62\% - 31\% = 31\%$ of children not stunted—the height of the blue arrow. Relative to the minimum, it is feasible for India to achieve $94\% - 31\% = 63\%$ of children not stunted—the height of the red arrow. Thus, India's *income-adjusted* score on the Right to Food is calculated as $(31\%/63\%) \times 100 = 49.2\%$.

Figure 2: Rescaling the indicator scores



More generally, the rescaling formula is:

$$S = 100 [(actual\ value - minimum\ value) / (frontier\ value - minimum\ value)]$$

Here, formally we refer to S as the **rescaled indicator score**. The numerator of the ratio in brackets reflects the extent to which a given right aspect is enjoyed, while the denominator of the ratio reflects the level of the State's immediate obligation to ensure that right aspect. After multiplying by 100, the rescaled indicator scores can be interpreted as the percentage of the immediate obligation met. The minimum values are set to approximate the indicator value one would expect to observe in a country with a subsistence per capita income level that places no priority on ensuring economic and social rights. This is approximated as zero for those indicators for which the score significantly depends on state provision of goods and services (e.g. the net primary school enrolment rate); otherwise, as discussed, above it is

approximated as the minimum value observed in any country in any year since 1990.⁴⁹ The minimum scores for each indicator are shown in Appendix Table B.

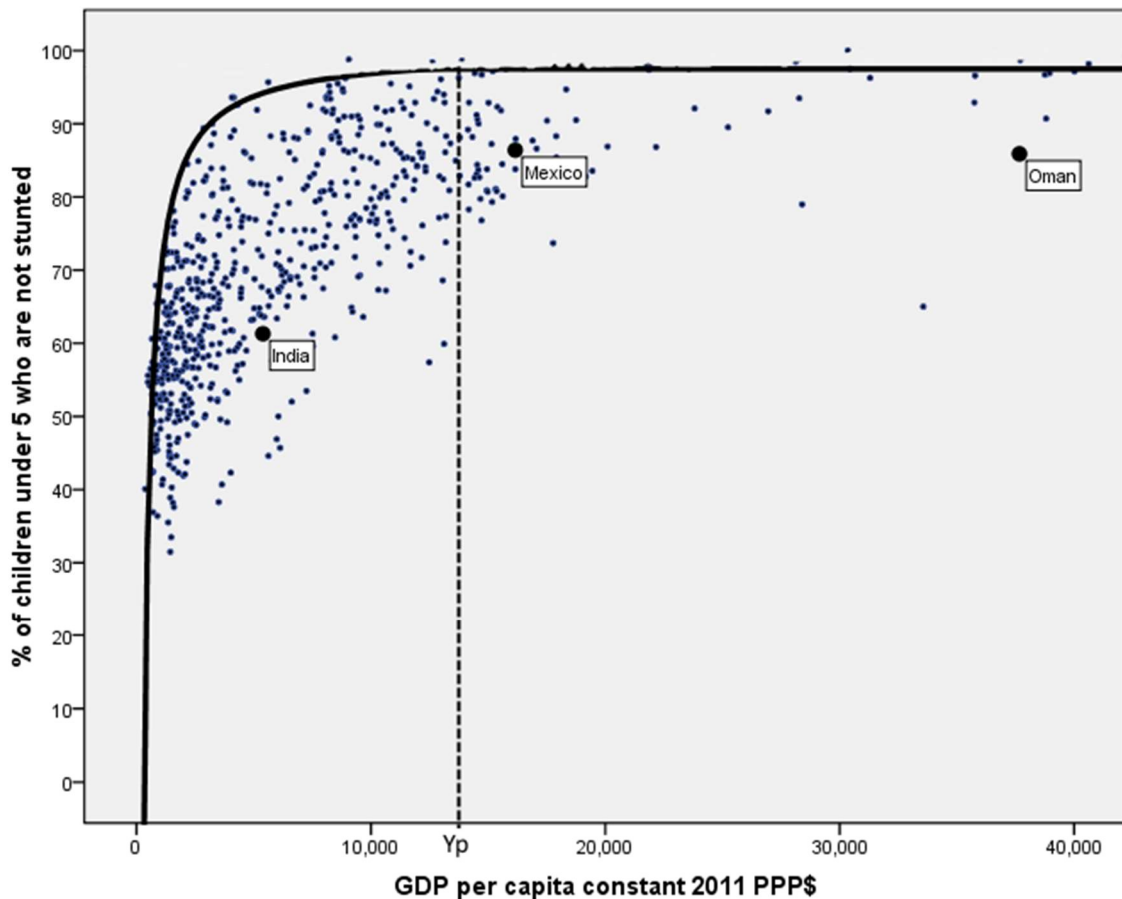
There is one more issue that needs to be taken into account: some countries have many times the resources needed to ensure all people enjoy a given right but fail to ensure that all people in fact enjoy that right. Figure 9 below fills out the scatter plot and APF for the percentage of children that are not stunted to include higher per capita income levels. Notice that the APF peaks and then becomes horizontal. The indicator value where the APF peaks (that we call X_p), is the highest value of the indicator observed at any income level. It is the current global best performance and in most cases implies the right aspect concerned is enjoyed by everyone in the country. In the case of the % of children that are not stunted, this occurs at 97.7%, since the height of 2.3% of children is expected to be more than 2 standard deviations below the median height for a well-nourished population. Appendix Table B specifies the global best (X_p) values for all the indicators. It should also be noted that in many cases, the frontier reaches a peak and then plateaus at a per capita GDP level well below the highest observed per capita income level.

We call the per capita income level where the frontier first reaches its peak Y_p . It can be interpreted as the minimum per capita GDP required to ensure enjoyment of the right aspect concerned by everyone in the population given current knowledge of the structures and measures (legislation, policies, programs, etc.) that promote that goal. In the case of the percentage of children that are not stunted, this occurs at \$13,608 (2011 PPP\$) as seen in Figure 3 below.

In general, countries with income levels exceeding Y_p have more than sufficient income to ensure everyone enjoys the aspect of the right concerned. The Y_p values differ substantially across indicators and are also shown in Appendix Table B. The rate at which resources can be transformed into enjoyment of the right aspect concerned is shown by the shape of the frontier as it rises to its peak value and is implicit in the estimated frontier equations. Those rising more steeply imply greater ease in transforming income into enjoyment of the right aspect concerned.

⁴⁹ With regard to the minimum values used to rescale indicators, the distinction between those indicator scores that substantially depend on public provision of goods and services (with a consequent 0 minimum) and those that do not is a refinement incorporated into the 2011 and later updates of the SERF Index as well as HRMI's ESR metrics.

Figure 3: Oman's resources exceed the level needed to eliminate child stunting.



The frontier value of the indicator will be the same for countries with per capita income levels above Y_p whether their per capita income level is exactly Y_p or two times Y_p , or even 10 times Y_p , and thus their rescaled performance indicator score will be the same. However, it makes little sense to evaluate two countries with the same indicator score as performing equally well if one has twice as much income as another. Looking again at Figure 3, notice that Oman and Mexico have nearly the identical percentage of children that are not stunted (86.4% for Mexico and 85.9% for Oman), yet Oman's per capita income is nearly 2.5 times higher than Mexico's (\$37,667 vs. \$16,158 measured in 2011 PPP\$). Also notice that for per capita income levels higher than \$13,608 (2011 PPP\$), the value of Y_p for the percentage of children not stunted, which is a bit less than Mexico's per capita income, the frontier reaches its peak value (97.7%), so resources no longer constrain countries' ability to eliminate child stunting. For countries like Oman with per capita income levels multiple times what is needed to reach the frontier but who still fail to do so, we impose a penalty on their rescaled indicator score. In Oman's case, based on the formula discussed below this is about 10 percentage points. A penalty is also imposed on

Mexico's rescaled indicator score, but the penalty is much smaller since its per capita income is only slightly higher than Y_p .

More generally, the final step in calculating the indicator score is to deduct a penalty from the rescaled indicator score when a country has income that is more than sufficient to ensure everyone in the country enjoys the right aspect concerned but fails to ensure that everyone does so. Thus, the final indicator score, what we formally call the **adjusted indicator score, A**, is:

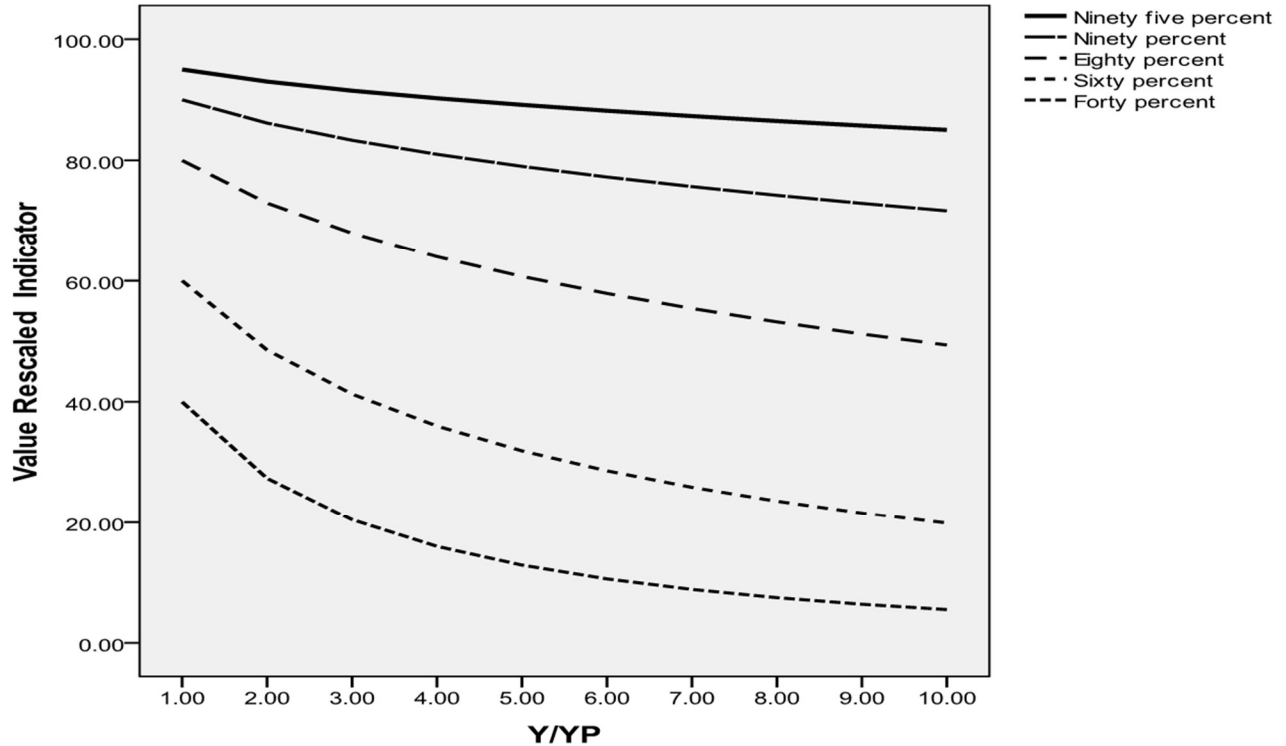
$$A = S \text{ if } Y \leq Y_p$$

$$A = S - \text{penalty if } Y > Y_p$$

A number of alternative penalty formulas were considered in Fukuda-Parr, Lawson-Remer, and Randolph (2009) along with a set of axioms defining the characteristics one would like such a penalty formula to have. On the basis of the axioms, penalty formula F was identified as meeting all but the flexibility criterion. A refinement of penalty formula F offered in Randolph, Fukuda-Parr, Lawson-Remer (2010) ensures it meets the flexibility criterion as well. The resultant **adjusted indicator score, A**, when $Y > Y_p$ is:

$$A = 100 \left[\left(\frac{S}{100} \right) \left(\frac{Y}{Y_p} \right)^\beta \right]$$

The value of β determines the severity of the penalty and for purposes of calculating HRMI's indicator scores, β is set equal to 0.5. Figure 4 plots the adjusted indicator score against the ratio of a country's per capita GDP to the Y_p value for rescaled indicator scores, S scores, of 95%, 90%, 80%, 60%, and 40%. For example, the figure indicates that if a country has an S score of 95%, the penalty reduces the adjusted indicator score to 85% as its income rises to ten times the minimum amount necessary to fulfil the right aspect concerned.

Figure 4: Penalty for different Y/Yp values

4.4.2 Assessing state performance: the *global best* indicator score

The procedure used to calculate HRMI's *global best* indicator scores are identical to those used to calculate HRMI's *income-adjusted* indicator scores except that the benchmark against which country indicator values are compared is the peak value of the APF (Y_p), our *global best* benchmark. More specifically, the rescaling formula becomes: the rescaling formula is:

$$S_{gb} = 100 [(actual\ value - minimum\ value) / (global\ best\ value - minimum\ value)]$$

The *global best* benchmark is the same for all countries and reflects a country's progress toward meeting its full obligations of result. Looking again at the example of India, the numerator is India's achievement relative to the minimum observed score is 62%-31%=31% of children not stunted—the height of the blue arrow, the same as before. However, the denominator changes and is calculated relative to the constant global best score, or 97.7% - 31% = 66.7%. Thus, India's *global best* score on the Right to Food is calculated as $(31\%/66.7\%) \times 100 = 46.5\%$.

As is the case for the *income-adjusted* scores, countries with income in excess to the minimum required to achieve the *global best* benchmark but fail to do so have a penalty applied to their indicator score. Accordingly, the *global best* adjusted indicator score, A_{gb} , is calculated as:

$$A_{gb} = S_{gb} \text{ if } Y \leq Y_p$$

$$A_{gb} = S_{gb} - \text{penalty if } Y > Y_p$$

The penalty formula for the *global best* adjusted indicator score is the same as for the *income-adjusted* adjusted indicator score. In the case of countries with incomes less than Y_p , their adjusted indicator scores will be lower using the *global best* benchmark than the *income adjusted* benchmark reflecting the fact that, given the principle of progressive realisation, their immediate obligation is less than their full obligation. In the case of countries with incomes greater than Y_p , the adjusted indicator scores will be identical regardless of the benchmark used.

4.5 Right scores & Quality of Life score

Each substantive right score is computed as the simple average of the underlying rescaled and adjusted indicator scores for the different aspects of the right assessed. For simplicity sake, we will refer to the rescaled and adjusted indicator scores simply as the indicator scores from here on out. So for example, using the low- and middle-income assessment standard, the right to education score is the average of the indicator scores for the adjusted net primary school enrolment rate and the net secondary school enrolment rate. In the event a single bellwether indicator is used to assess the enjoyment of a right, the substantive right score is simply the relevant indicator score. So for example, using the low-middle income assessment standard, the right to food score (*income-adjusted* or *global best* depending on the benchmark concerned) is the indicator score for the percentage of children that are not stunted. Thus, differentiating between the different indicator scores with i , and denoting n as the number of indicator scores relevant to right k , the formula for a given **substantive right score**, R_k , is:

$$R_k = \sum A_i / n$$

The exception is for the high-income right to education score. Here the three education quality indicator scores (percent of students achieving level 3 or better on the PISA math, PISA science and PISA reading scores) are first averaged and then that result is averaged with the education access score (the net secondary school enrolment rate).

Finally, the substantive right scores are averaged to yield HRMI's "*Quality of Life*" score⁵⁰. That is, given that we have 5 substantive rights:

$$\text{Quality of Life} = \Sigma R_k / m$$

Table 3 below shows the indicator scores that are averaged for each right for both assessment standards.

Table 3. Sub-scores Comprising HRMI Right Scores by Assessment Standard

Right and Sub-Rights	Assessment Standard	
	Low- and Middle-Income	High-income
Right to food score		
Children non-stunted score	√	
Food security score		√
Right to education score		
Secondary school enrolment score	√	√
Primary school enrolment score	√	
Education quality score (average math, science & reading)		√
Right to health score		
Children surviving to age 5 score	√	√
People surviving to age 65 score	√	√
Contraceptive use score	√	
Birthweight score		√
Right to housing score		
Basic sanitation score	√	
Water on premises score	√	
Safely managed sanitation score		√
Right to work score		
Absolute poverty score	√	
Relative poverty score		√
Long term unemployment score		√

⁵⁰ The Quality of Life score is HRMI's name for the SERF Index.

4.6 References

- Chapman, Audrey. (2007) The Status of Efforts to Monitor Economic, Social, and Cultural Rights. Chapter 7, pp. 143-164 in Shareen Hertel and Lanse Minkler (eds.), *Economic Rights: Conceptual, Measurement and Policy Issues*. Cambridge: Cambridge University Press.
- Chapman, Audrey. (1996) “A Violations Approach for Monitoring the International Covenant on Economic, Social, and Cultural Rights,” *Human Rights Quarterly*, 18(1), 23-66.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1990) “General Comment 3: The Nature of States Parties’ Obligations”, Fifth Session, December 14.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1991) “General Comment 4: The Right to Adequate Housing”, Sixth Session, 13 December.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1997) “General Comment 7: The Right to Adequate Housing—Forced Evictions”, Sixteenth session, 20 May.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1998) “General Comment 9: The Domestic Application of the Covenant” Nineteenth session, 16 November – 4 December, Doc. E/C.12/1998/24, 3 December 1998.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1999a) “General Comment 11: Plans of Action for Primary Education”, Twentieth session, Geneva, 26 April – 14 May 1999, Document E/C.12/1999/4.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1999b) “General Comment 12: The right to adequate food”, Twentieth Session of the Economic and Social Council, Geneva, 26 April – 14 May. Document E/C.12/1999/5.
- (Economic and Social Council), Committee on Economic Social and Cultural Rights. (1999c) “General Comment 13: The Right to Education” Twenty-first session 15 November – 3 December 1999. Document E/C.12/1999/10.
- (Economic and Social Council), Committee on Economic, Social and Cultural Rights. (2000) “General Comment 14: The Right to the Highest Attainable Standard of Health”, 22nd Session, 25 April – 12 May 2000. Document E/C.12/2000/4.
- (Economic and Social Council), Committee on Economic, Social and Cultural Rights. (2005b) “General Comment No. 18: The Right to Work”, 35th Session, 7-25 November 2005. Document E/C.12/GC/18, 6 February 2006.
- (Economic and Social Council), Committee on Economic, Social and Cultural Rights. (2007) “General Comment No. 19: The Right to Social Security” 39th session, 5-23 November. Document E/C.12/GC/19, 4 February 2008.
- Fukuda-Parr, Sakiko, Terra Lawson-Remer and Susan Randolph. (2009) “An Index of Economic and Social Rights Fulfillment: Concept and Methodology”, *Journal of Human Rights*, 8: 195-221.
- Sakiko Fukuda-Parr, Terra Lawson-Remer and Susan Randolph, *Fulfilling Social and Economic Rights* (Oxford: Oxford University Press, 2015)
- Randolph, Susan, Sakiko Fukuda-Parr, and Terra Lawson-Remer. (2010). “Economic and Social Rights Fulfillment Index: Country Scores and Rankings”, *Journal of Human Rights*. 9: 230-261, 2010.
- United Nations. (1945) Charter of the United Nations (San Francisco, 26 June 1945) (*entered into force* 24 Oct. 1945).
- United Nations. (1948) Universal Declaration of Human Rights (UDHR), *adopted* 10 Dec. 1948, United Nations General Assembly Res. 217 A (III) (1948) New York.
- United Nations. (1966a) International Covenant on Economic, Social and Cultural Rights (ICESCR), 1966. *Adopted* 16 Dec. 1966, General Assembly Res. 2200 (XXI), U.N. GAOR, 21st Sess., Supp. No. 16, U.N. Doc. A/6316 (1966), 993 U.N.T.S. 3 (*entered into force* 3 Jan. 1976).
- United Nations. (1987) The Limburg Principles on the Implementation of the International Covenant on Economic, Social and Cultural Rights. Guidelines adopted at a workshop sponsored by the International Commission of Jurists, the Faculty of Law of the University of Limburg, and the Urban Morgan Institute for Human Rights, University of Cincinnati, Maastricht, Netherlands, 22-26 January 1997, UN doc. E/CN.4/1987/17.

United Nations (2000) The Maastricht Guidelines on Violations of Economic, Social and Cultural Rights.

Guidelines adopted at a workshop sponsored by the International Commission of Jurists, the Urban Morgan Institute for Human Rights and the Center for Human Rights of the Faculty of Law of Maastricht University, Maastricht, Netherlands, 22-26 January, 1997. UN doc. E/C.12/2000/13.

United Nations Office of the High Commissioner for Human Rights (2012). *Human Rights Indicators: A Guide to Measurement and Implementation*. HR/PUB/12/5. New York: Office of the High Commissioner for Human Rights, United Nations.

4.7 Appendix

Appendix Table A: Indicator Definitions

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
Resources				
GDP per capita (2011 PPP\$)	Both	World Bank International Comparison Project.	Nov. 18, 2018 from WB WDI at https://databank.worldbank.org/data/source/world-development-indicators#	GDP per capita based on purchasing power parity (PPP), PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2011 international dollars.
Education				
#Net Secondary School Enrolment (UIS_UNESCO)	Both & By Sex	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics (UIS).	January 1, 2019 from WB EdStats https://datacatalog.worldbank.org/dataset/education-statistics	Total number of students in the theoretical secondary school age group enrolled in secondary school, expressed as a percentage of the total of secondary school age. Divide the number of students enrolled who are of the official age group for secondary education by the population for the same age group and multiply the result by 100. NER at each level of education should be based on enrolment of the relevant age group in all types of schools and education institutions, including public, private and all other institutions that provide organized educational programmes.

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
#Adjusted Net Primary Enrolment	Low-and Middle Income & by Sex	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics (UIS).	December 31, 2018 from WB EdStats https://datacatalog.worldbank.org/dataset/education-statistics	<p>Total number of students of the official primary school age group who are enrolled at primary or secondary education, expressed as a percentage of the corresponding population. Divide the total number of students in the official primary school age range who are enrolled in primary or secondary education by the population of the same age group and multiply the result by 100. NERA should be based on total enrolment of the official primary school participation age group in primary and secondary education for all types of schools and education institutions, including public, private and all other institutions that provide organized educational programmes (pre-primary education must be excluded). NERA gives more precise measure of the participation of the official primary school age population to the education system (excluding pre-primary education). It reflects the actual level of achievement of the Universal Primary Education (UPE) goal. In fact, while the Net enrolment rate (NER) shows the coverage of pupils in the official primary school age group in the primary education level only, the NERA extends the measure to those of the official primary school age range who have reached secondary education because they might access primary education earlier than the official entrance or they might skip some grades due to their performance. Increasing NERA might mirror improving participation of children in the official primary school age, the decrease of the target population or both. A value of 100% indicates theoretically that the country has accomplished the UPE goal. However, this condition is not sufficient for UPE due to, for example, a high repetition rate, which might lead pupils to dropout after primary school age without completing primary education. The difference between NERA and NER provides a measure of the proportion of children in the official primary age group who are enrolled in secondary education.</p>

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
# % Scoring Level 3 or greater on PISA science test	High-Income & by Sex	Program for International Student Assessment (PISA)	January 5, 2019 from http://pisadataexplorer.oecd.org/ide/idepisa/	<p>Achievement levels run from 1 to 6. At Level 3, students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 students are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim. Level definition from figure 4.10 in OECD (2018), "PISA for Development Reading Framework", in <i>PISA for Development Assessment and Analytical Framework: Reading, Mathematics, and Science</i>, OECD publishing, Paris, https://doi.org/10.1787/9789264305274-4-en</p>
# % Scoring Level 3 or greater on PISA math test	High-Income & By Sex	Program for International Student Assessment (PISA)	January 5, 2019 from http://pisadataexplorer.oecd.org/ide/idepisa/	<p>Achievement levels run from 1 to 6. At Level 3, students can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning. Level definition from figure 3.3 in OECD (2018), "PISA for Development Reading Framework", in <i>PISA for Development Assessment and Analytical Framework: Reading, Mathematics, and Science</i>, OECD publishing, Paris, https://doi.org/10.1787/9789264305274-4-en</p>

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
# % Scoring Level 3 or greater on PISA Reading Test	High-Income & By Sex	Program for International Student Assessment	January 5, 2019 from http://pisadataexplorer.oecd.org/ide/idepisa/	Achievement levels run from 1 to 6. Tasks at this level [level 3] require the reader to locate, and in some cases recognize the relationship between, several pieces of information that must meet multiple conditions. Interpretative tasks at this level require the reader to integrate several parts of a text in order to identify a main idea, understand a relationship or construe the meaning of a word or phrase. They need to take into account many features in comparing, contrasting, or categorizing. Often the required information is not prominent or there is much competing information; or there are other text obstacles, such as ideas that are contrary to expectation or negatively worded. Reflective tasks at this level may require connections, comparisons, and explanations, or they may require the reader to evaluate a feature of the text. Some reflective tasks require readers to demonstrate a fine understanding of the text in relation to familiar, everyday knowledge. Other tasks do not require detailed text comprehension but require the reader to draw on less common knowledge. Level definition from figure 2.2 in OECD (2018), "PISA for Development Reading Framework", in <i>PISA for Development Assessment and Analytical Framework: Reading, Mathematics, and Science</i> , OECD publishing, Paris, https://doi.org/10.1787/9789264305274-4-en
Food				
100% - Malnutrition Prevalence - height for Age (% children under 5)	Low- and Middle-Income & By Sex	UNICEF, WHO, World Bank: Joint child malnutrition estimates (JME)	December 28, 2018 from WB WDI from https://databank.worldbank.org/data/reports.aspx?source=world-development-indicators	Prevalence of stunting is the percentage of children under age 5 whose height for age is more than two standard deviations below the median for the international reference population ages 0-59 months. For children up to two years old height is measured by recumbent length. For older children height is measured by stature while standing. The data are based on the WHO's new child growth standards released in 2006. UNICEF, WHO, World Bank: Joint child malnutrition estimates (JME). Aggregation is based on UNICEF, WHO, and the World Bank harmonized dataset (adjusted, comparable data) and methodology.
#100% - % Moderately or Severely Food Insecure	High-Income	Food and Agricultural Organization (FAO)	January 20, 2019 from the Sustainable Development Goals Database from https://unstats.un.org/sdgs/indicators/database/	The indicator measures the percentage of individuals in the population who have experienced food insecurity at moderate or severe levels during the reference period. The severity of food insecurity, defined as a latent trait, is measured on the Food Insecurity Experience Scale global reference scale, a measurement standard established by FAO through the application of the Food Insecurity Experience Scale

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
Health				
Total Age65 survival rate (% cohort)	Both & b7y Sex	United Nations Population Division (UNPD) World Population Prospects: 2017 Revision	December 25, 2018 from WB HNP https://datacatalog.worldbank.org/dataset/health-nutrition-and-population-statistics	Survival to age 65 refers to the percentage of a cohort of newborn infants that would survive to age 65, if subject to age specific mortality rates of the specified year. Total Age65 survival rate estimated using sex specific Age 0 population data. Age population, refers to population at the specified age level from WB HNP data set.
100% - child (under 5) mortality rate/10	Both & by Sex	UN Inter-agency Group for Child Mortality Estimation (UNICEF, WHO, World Bank, UN DESA Population Division)	December 24, 2018 from https://databank.worldbank.org/data/source/world-development-indicators#	Under-five mortality rate is the probability per 1,000 that a newborn baby will die before reaching age five, if subject to age-specific mortality rates of the specified year.
Modern contraceptive Use rate (% couples 15-49)	Low- and Middle-Income	Compiled by United Nations Population Division from household surveys, including Demographic and Health Surveys and Multiple Indicator Cluster Surveys.	December 27, 2018 from https://databank.worldbank.org/data/source/world-development-indicators#	Contraceptive prevalence rate is the percentage of women who are practicing, or whose sexual partners are practicing, at least one modern method of contraception. It is usually measured for women ages 15-49 who are married or in union. Modern methods of contraception include female and male sterilization, oral hormonal pills, the intra-uterine device (IUD), the male condom, injectables, the implant (including Norplant), vaginal barrier methods, the female condom and emergency contraception. Retrieved from WB WDI 12_27_18 at https://databank.worldbank.org/data/source/world-development-indicators#
100% - % Low Birth Weight Babies	HiYOECD	Primary source OECD statistics, secondary source WB WDI.	December 27, 2018 from https://databank.worldbank.org/data/source/world-development-indicators#	Number of live births weighing less than 2500 grams as a percentage of total number of live births.

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
Housing				
#% population with basic access to water on premises	Low- and Middle-Income	WHO UNICEF Joint Monitoring Project (JMP)	January 12, 2019 from JMP data https://washdata.org/data/household	Improved drinking water sources are those which, by nature of their design and construction, have the potential to deliver safe water. The JMP subdivides the population using improved sources into three groups according to the level of service provided. In order to meet the criteria for a safely managed drinking water service, people must use an improved source meeting three criteria: a)it should be accessible on premises, b)water should be available when needed, and c)the water supplied should be free from contamination. If the improved source does not meet any one of these criteria but a round trip to collect water takes 30 minutes or less, then it will be classified as a basic drinking water service.
#% population with basic sanitation.	Low- and Middle-Income	WHO UNICEF Joint Monitoring Project (JMP)	January 12, 2019 from WB WDI https://databank.worldbank.org/data/source/world-development-indicators#	Basic sanitation is defined as use of improved facilities which are not shared with other households. Improved sanitation facilities are those designed to hygienically separate excreta from human contact. Data from WHO UNICEF Joint Monitoring Project (JMP) https://washdata.org/data/household Retrieved from WB WDI 1_12_2019.
#% households with safely managed sanitation	High-Income	WHO UNICEF Joint Monitoring Project (JMP)	January 12, 2019 from WB WDI https://databank.worldbank.org/data/source/world-development-indicators#	Safely managed sanitation is defined as use of improved facilities which are not shared with other households and where excreta are safely disposed in situ or transported and treated off-site. Improved sanitation facilities are those designed to hygienically separate excreta from human contact.
Right to Work				
100% - % Relative Poverty Rate	High-Income	Primary Luxembourg Income Study Secondary OECD	January 8 2019 from https://www.lisdatacenter.org/data-access/key-figures/ and https://data.oecd.org/inequality/poverty-rate.htm	Indicator of poverty status of the household to which the individual belongs to, based on the equalised disposable household income concept and with respect to the 50% of the median. Retrieved 1_8_2019 from Luxembourg Income Study Key Figures https://www.lisdatacenter.org/data-access/key-figures/ Secondary data source OECD (Definition: The poverty rate is the ratio of the number of people whose income falls below the poverty line; taken as half the median household income of the total population.) Retrieved 1_8_2019 from https://data.oecd.org/inequality/poverty-rate.htm

Indicator	Assessment Standard	Primary Source	Date Accessed	Definition
100% - Poverty Headcount ratio (at \$3.20 2011 PPP\$ per day)	Low- and Middle-Income	World Bank	January 11, 2019 from https://databank.worldbank.org/data/source/world-development-indicators#	Poverty headcount ratio at \$3.20 a day is the percentage of the population living on less than \$3.20 a day at 2011 international prices. As a result of revisions in PPP exchange rates, poverty rates for individual countries cannot be compared with poverty rates reported in earlier editions. Data from World Bank, Development Research Group and are based on primary household survey data obtained from government statistical agencies and World Bank country departments. Data for high-income economies are from the Luxembourg Income Study database. See PovcalNet (http://iresearch.worldbank.org/PovcalNet/index.htm). For more information.
100% - % Long term Unemployed (%unemployed)	High-Income	OECD	January 11, 2019 from https://data.oecd.org/unemp/long-term-unemployment-rate.htm	Long-term unemployment refers to the number of people with continuous periods of unemployment extending for 12 months or longer, expressed as a percentage of the total unemployed.

Indicates new indicator and frontier estimate for 2019 update.

Appendix Table B: Frontier Equations, Peak Indicator Values, Income level at Peak Indicator Value, Minimum Value

Right/Indicator	Frontier Equation*	Peak Value (Xp)	Income Level at Xp (Yp)	Minimum Value
Right to Food				
% Children (under 5) Not Stunted (all, male, female)	$Y = 100 - 31300/x$ for $X < 13608$; else 97.7	97.7 (based on WHO definition that 2.3% population will be > 2 s.d below mean in healthy population)	\$13608 (2011 PPP)	31% (Bangladesh in 1995)
% Food Secure	$Y = 99 - 50000/(x-500)$	99%	Asymptotic	6% (South Sudan 2015)
Right to Education				
PISAscience	$Y = 87 - 280000/(x-2500)$	87	Asymptotic	0 (Actual minimum observed 3% Dominican Republic, 2015)
PISAmath	$Y = 83 - 180000/(x-3000)$	83	Asymptotic	0% (Actual minimum observed 2% in Dominican Republic, 2015)
PISAreading	$Y = 85 - 180000/(x-2000)$	85	Asymptotic	0% (Actual minimum observed 4% in Algeria, 2015)
Net secondary school enrolment	$Y = 102 - 58000/x$ for $x \leq 29000$; else=100	100	29000	0 (Actual minimum observed 2.7% in Mozambique, 1999)
Adjusted net primary enrolment	$Y = 102 - 3400/(x-600)$ for $x \leq 2300$; else = 100	100	2300	0 (Actual minimum observed, 21% for females in Niger, 1999)

Right/Indicator	Frontier Equation*	Peak Value (Xp)	Income Level at Xp (Yp)	Minimum Value
Right to Health				
Age 65 Survival Rate (total population)	$Y = 93 - 50000/(x + 700)$	93	Asymptotic	16% (Zimbabwe in 2002)
Male Age 65 Survival Rate	$Y = 91 - 62000/(x + 1000)$	91	Asymptotic	16% (Zimbabwe in 2002)
Female Age 65 Survival Rate	$Y = 96 - 55000/(x+800)$	96	Asymptotic	16% (Zimbabwe in 2002)
% Child (Under 5) Survival Rate (both sexes)	$Y = 100 - 6000/x$	100%	Asymptotic	68% (Niger in 1990)
% Male Child (Under 5) Survival Rate	$Y = 100 - 6000/x$	100	Asymptotic	68% (Niger in 1990)
% Female Child (Under 5) Survival Rate	$Y = 100 - 5000/x$	100	Asymptotic	68% (Niger in 1990)
Modern Contraceptive Use Rate	$Y = 85 - 30000/x$	85%	Asymptotic	0% (South Sudan 1% in 2006)
% New-borns Not Low Birth Weight	$Y = 97 - 5600/x$	97%	Asymptotic	40% (Lao PDR 1991, 1994)
Right to Housing				
Access to Water on Premises (% Population)	$Y = 113 - 75000/x$ for $x < 5739$; else 100%	100%	5739	0% (Actual minimum 1.7% in South Sudan, 2011)
Access to Basic Sanitation (% Population)	$Y = 111 - 40000/(x-200)$ for $x < 3436$; else 100%	100%	3436	0% (Actual minimum 4.6% in Madagascar, 2000)
Access to safely managed sanitation (% population)	$Y = 105 - 210000/x$ for $x < 42,000$; else = 100%	100%	42000	0% (Actual minimum 4% in Niger, 2000)

Right/Indicator	Frontier Equation*	Peak Value (Xp)	Income Level at Xp (Yp)	Minimum Value
Right to Work				
Not Long-term Unemployed (% unemployed)	$Y = 100 - 22000/x$	100%	Asymptotic	0% (Actual minimum 99.8% in South Korea, 2014))
Not Relatively Poor (> 50% Median Income)	$Y = 97 - 35000/(x-4000)$	97%	Asymptotic	70% (Peru in 2004)
Not Absolutely Poor (> 3.20 2011 PPP\$ per day)	$Y = 108 - 60000/x$ for $x < 5454$; else = 100%	100%	\$5454	0% (Actual minimum 1.5% Congo, Dem. Rep.)

* In the equations above, Y refers to the indicator concerned while x refers to GDP per capita measured in 2011 PPP\$.

