ECINEQ
Society for the Study of Economic Inequality

Working Paper Series

Measuring the statistical capacity of nations

Grant J. Cameron
Hai-Anh H. Dang
Mustafa Dinc
James Foster
Michael M. Lokshin

ECINEQ WP 2019 - 491

ECINEQ 2019 - 491
March 2019

www.ecineq.org

Measuring the statistical capacity of nations*

Grant J. Cameron†
Hai-Anh H. Dang
Mustafa Dinc
World Bank, USA

James Foster
George Washington University, USA

Michael M. Lokshin
World Bank, USA

Abstract

The international development community has used the World Bank's Statistical Capacity Index since its inception in 2004. The Sustainable Development Goals create new challenges for national statistical systems to produce high-quality and internationally comparable data. This paper reviews measurement methodologies, posits desired attributes, and presents theoretical and empirical frameworks for the new, improved index to monitor progress in the statistical capacity of nations. The paper illustrates the properties of the updated index with global data from 2016.

Keywords: statistical capacity, statistical indicators, statistical index, national statistical system, data, measurement.

JEL Classification: C8, H00, I00, O1.

---

*We would like to thank Juderica Dias for helpful support with the data, and Junhe Yang for capable research assistance. We are grateful to the UK Department of International Development for funding assistance through its Knowledge for Change (KCP) program and Strategic Research Program (SRP).

†**Contact details**: Cameron (gcameron1@worldbank.org) is manager with the Advisory Services and Economic Statistics Unit, Development Data Group, World Bank; Dang (hdang@worldbank.org; corresponding author) is economist with the Survey Unit, Development Data Group, World Bank; Dinc (mdinc@worldbank.org) is senior economist with the Advisory Services and Economic Statistics Unit, Development Data Group, World Bank; Foster (fosterje@gwu.edu; corresponding author) is the Oliver T Carr Professor of International Affairs and Professor of Economics, George Washington University; Lokshin (mlokshin@worldbank.org) is manager of the Survey Unit, Development Data Group, World Bank.

1. Introduction

The national statistical system, or NSS, plays a crucial role in modern economies. It provides stake-holders, ranging from policy makers to stock market analysts and the general public, with the latest data on the country's socio-economic developments. At the international level, monitoring progress on global undertakings such as the recently established Sustainable Development Goals (SDGs) requires high-quality data that must be produced consistently across different national statistical systems. Assessing and improving the capacity of a country's NSS has long been a part of the global agenda. The international Partnership in Statistics for Development in the 21st Century (PARIS21) Task Team was established in 2002 to help measure country statistical capacity. Over the subsequent years, a few capacity assessment tools have been developed to identify the weaknesses and strengths of national statistical systems.[1]

The World Bank's Statistical Capacity Index (SCI) is one such tool that has been widely employed.[2] Several international and national agencies have adopted the SCI for measuring progress in statistical capacity building and related investments. The United Nations, for example, uses the SCI to measure trends in the development of national statistical capacity (United Nations, 2016). The SCI is used to evaluate the efficiency of statistical support provided to a country as well as the need to further develop its statistical capacity (PARIS21, 2002). Some regional organizations use the SCI to identify areas of improvement in their member countries (OIC, 2012), while researchers use the SCI as a benchmark to validate their new statistical indexes (Sanga et

[1] Statistical capacity is usually interpreted as the ability of an NSS to meet user needs for relevant and good quality statistics in a timely manner. An NSS often consists of a number of different data-producing agencies and departments (such as the national statistical office, the central bank, and statistical departments within other line ministries), which renders the task of directly measuring statistical capacity a difficult one.

[2] For brevity, we refer to both the Statistical Capacity Indicators and the Statistical Capacity Index as the SCI in the rest of the paper. We will make it clear where we refer to either the indicators or the index. We similarly refer to both the Statistical Performance Indicators and the Statistical Performance Index as the SPI. We also refer to the old SCI as the SCI, and the newly revised SCI presented here as the SPI.

al., 2011). The World Bank mainstreamed the SCI in its monitoring and assessment framework and has adopted it as a baseline indicator in various projects at the country level.[3] The SCI is based on publicly available data, and this has various advantages over other indexes of statistical capacity. A key advantage of the SCI is that it can provide assessment of a country's statistical capacity in an internationally comparable and cost-effective manner.

Existing efforts in building indexes to assess statistical capacity have focused on the practical aspects such as data collection, organization, and legal issues, paying little attention to the underlying theoretical principles that are indispensable for the construction of a reliable, transparent, and consistent statistical capacity index. For example, the UNECE, in a recent Global Assessment report, discusses only the legal basis, description of the statistical system, data source, and processing of the target country (UNECE, 2014). The FAO, in its guidelines for assessing country capacity in producing agricultural statistics, provides instructions on completing the questionnaires and on compiling the assessment indicator (FAO, 2014), but pays no attention to the axiomatic principles of these indicators. The U.S. Census Bureau developed and recently updated (2017) the Tool for Assessing Statistical Capacity (TASC) with a primary objective of measuring the overall capacity of an NSS by providing a breakdown of the areas of strength and weakness. However, the focus of this instrument is on measuring the capacity of an NSS to conduct household-based surveys and censuses.[4] To our knowledge, only Sanga, Dosso, and Gui-Diby (2011) discuss the technical framework behind the African Statistical Development Index (ASDI).

[3] For other recent examples that use the SCI, see: Beegle et al. (2016) for an analysis of the relationship between good governance and statistical capacity in African countries; Tapsoba, Noumon, and York (2017) for the impacts of statistical capacity on reducing procyclical fiscal policy; and UNICEF (2018) for the role of statistical capacity in tracking the SDG for child development.

[4] We return to provide more discussion on the SCI and these other methods in Section 2 below.

In this paper, we aim at laying out the conceptual foundation behind statistical capacity indexes, and construct a new index based on practical and theoretical considerations. We review existing measurement methodologies, posit desired attributes, and propose updated indicators, and an updated Statistical Capacity Index (hereafter referred to as the Statistical Performance Index, or SPI). On the empirical front, we expand the number of indicators in the old SCI by almost twice, and we extend the sample of covered countries by one-half to all countries in the world.

This paper consists of six sections. We provide a brief overview on the SCI in the next section before presenting the theory in Section 3. We offer in this section detailed discussion on the proposed index ranging from the conceptual framework (including its policy relevance, its concept of statistical capacity and desirable characteristics) to the technical framework (including its dimensions, aggregation methods, and axioms). We then describe the SPI and its new features before comparing it with the SCI in Section 4 and providing some empirical illustrations in Section 5. Section 6 concludes.

2. Background on the SCI

The SCI is a tool for assessing country-level statistical capacity, developed in 2004 to assess the effectiveness of the World Bank's lending projects related to improvements in countries' statistical capacity. Table 1 compares the SCI with the statistical capacity measurement indexes used by other organizations including PARIS21, the Food and Agriculture Organization of the United Nations (FAO), the United Nations Economic Commission for Europe (UNECE), the United Nations Economic Commission for Africa (UNECA), and the U.S. Census Bureau. Table 1 shows that the SCI is the only tool that provides comparable data across different countries over time. Further, the SCI covers the largest number of countries—146 countries starting from 2004 and has by far the widest coverage of country-level indicators.

There is also a key methodological difference between the SCI and other assessment tools. The latter methods collect data directly from national statistical offices' staffs or local experts. While this procedure may provide a more in-depth analysis and uncover finer details in the organization of an NSS, it incurs high costs and is (far) more time-consuming. Furthermore, even the best evaluator can bring personal biases, non-uniform conceptions of capacity, or other subjective elements. Direct interviews of government officials might bias responses and complicate comparability across countries.[5]

The SCI follows a different approach by using publicly available data and focusing on a set of easily observable and verifiable indicators. It provides internationally comparable, objective, country-level assessments across the globe. The World Bank has mainstreamed these indicators in its monitoring and assessment frameworks such as the Corporate Score Card (CSC) and the IDA Results Measurement System (RMS). The SCI is also used in various World Bank projects as a baseline indicator.

Yet, there are several areas in which the existing SCI can be improved. First, it comprises of a limited number of indicators and includes no indicators of some important surveys such as the labor force surveys and establishment surveys. Second, it ignores the data dissemination practices of an NSS, which is one of the key features of data usage. Third, the SCI has been criticized for placing too much weight on statistical output and activities, while neglecting the infrastructure and resource components of statistical systems. It was also criticized for its weighting scheme (i.e., equal weighting of each dimension and individual indicator) and aggregation method (i.e., simple arithmetic average).

[5] Government officials may have an incentive to overestimate their country's statistical capacity when there are concerns about their ability to deliver; the opposite can happen when they are requesting additional international aid or technical assistance.

Finally, since its launch in 2004, the SCI's methodology and coverage have basically remained the same, while the global data landscape has changed significantly. NSSs have made significant advancements with data collection and dissemination practices. At the same time, the adoption of the Sustainable Development Goals (SDGs) set an ambitious development agenda for the next 15 years on ending poverty, protecting the planet, and ensuring prosperity for all by 2030. This, in turn, increased the demand for data and raised the bar for national statistical systems regarding their capacity to produce high-quality data. We thus propose to improve the current SCI to better suit the changing global data landscape.

We next present the theoretical framework that can be used to construct an index that measures statistical capacity.

3. Theory

3.1. Outline of Policy Relevant Measurement

In order to construct a measure that is policy relevant it is helpful to follow a series of basic steps.[6]

The first step asks the question: what phenomenon is being measured? A clear conception helps orient the process by which the measure is assembled and will prove valuable in communicating its underlying meaning.

The second step asks: for what purpose or purposes is the index being sought? Knowing how the index will be used can greatly affect subsequent choices in its construction, and its eventual

[6] This process is similar for many types of measurement exercises. See for example Alkire et al. (2015) in the context of multidimensional poverty measurement.

suitability. In particular, it will help define the unit of analysis both for data gathering and reporting purposes.

The third step identifies a list of essential characteristics, or desiderata, that the methodology should exhibit. This list of "pre-axioms" helps orient the construction process and define what success means.

A fourth step identifies the conceptual space in which measurement is to take place. If there are multiple conceptual dimensions, consideration must also be given to the relative importance of each.

The fifth step selects the form of the variables to be used and the aggregation method to be employed – how the variables are to be combined into an overall measure.

The sixth step identifies a set of axioms that the resulting index should satisfy to have the greatest practical utility. Axioms are not sterile mathematical requirements, but rather contain the salient nuggets of policy required of the index: which aspects of the data should be ignored, which should be reflected, and helpful consistency requirements over subsets of data. Together, these six steps comprise the core theoretical elements of our proposed measurement technology.

3.2. Process to Construct an Index of Statistical Capacity

We now turn to a detailed discussion of these six steps and the elements involved in the context of an index of statistical capacity. We will also consider the practical details of our implementation of the general measurement approach.

3.2.1. What Is Being Measured? The Concept of Statistical Capacity

Statistical capacity is the ability of a National Statistical System (NSS) to sustainably collect, compile, and disseminate data to inform national and international policy decisions. Capacity is difficult to measure since it is only partially revealed by evaluating achievements or other observable characteristics of the NSS. A system may have the capacity to produce quality data products but has not yet done so; or it may no longer have the capacity to produce despite having succeeded in the past. Our approach is to focus entirely on the observables for the NSS, but then to expand consideration to a broader range of performance indicators, to better reflect its underlying capacity. The challenge then becomes how to identify the important dimensions of products and processes that should be covered by such a measure, as well as the specific variables within each dimension.

But before discussing dimensions and variables, it is important to spell out the purpose of the index and its desirable characteristics.

3.2.2. The Purpose of the Index

The purpose of a statistical capacity index is to assess a country's statistical capacity to monitor the progress of reforms and projects in the area of statistical capacity building and identify the areas for improvements in the National Statistical Systems. An index can also facilitate inter-country comparisons and help interpretation of the trends both within a country over time and across countries.

3.2.3. Desiderata

The desiderata for an index (or more generally a measurement method) are the general characteristics or criteria that we would like it to have. Consider the following list that modifies a standard set found in the literature on measurement.[7]

i. *Simple*. It must be understandable and easy to describe.

ii. *Coherent*. It must conform to a common-sense notion of what is being measured.

iii. *Motivated*. It must fit the purpose for which it is being developed.

iv. *Rigorous*. It must be technically solid.

v. *Implementable*. It must be operationally viable.

vi. *Replicable*. It must be easily replicable.

vii. *Incentive Compatible*. It must respect country incentives.

The first characteristic speaks to the simplicity of a measurement approach, and to the ease with which it might be communicated. Methods that measure an underlying concept in an unclear or unduly complicated fashion discourage participation by stakeholders in policy discussions. When there are significant barriers to understanding a measure, this by itself can lead stakeholders to question its authenticity as a policy relevant tool. In contrast, an index or method that is more easily understood allows stakeholders to evaluate for themselves its qualities. This desideratum also extends to the data upon which the measure is based by requiring a clear and understandable link between the indicators used in an index and the underlying data. If indicators are vaguely defined or sourced, then the interpretation of an index can be superficially clear, but actually obscure. This aspect of the requirement is facilitated by having online data sources and adequate

[7] See, for example, Alkire et al. (2015).

documentation to offer a clear provenance for each indicator. Transparency of the measure and its components is a key characteristic of an index of statistical capacity, which will facilitate its take-up by stakeholders, promote independent confirmation of its findings, and support its credibility.

The second characteristic of coherence speaks to the measure's authenticity in capturing the phenomenon it aims to measure. Simply put, one should be able to describe in plain language the underlying concept of what is being measured by the index, and then argue that this concept is indeed a credible version of what needs to be measured. In the present case, we will measure a nation's statistical capacity by examining its products and processes.

The third desideratum considers the extent to which the index fulfills the purpose for which it has been constructed. As noted above, the purpose of a statistical capacity index is to accurately and timely measure a country's statistical capacity. In creating the index, it is important to keep the motivating objectives in mind, and work towards their achievement.

The fourth characteristic addresses the rigor of the methodology embodied in the measure. Much of this concerns the aggregation method employed and the axioms that might be satisfied by an index of statistical capacity. For example, one would not want to have a multidimensional index for which unambiguous improvements in individual components would be associated with a decrease in overall capacity as measured by the index. To give substance to this general characteristic, we outline several axioms that the index might fulfill. There are also other measurement problems that a technically solid index should avoid. For example, if the data used in the index are not cardinally relevant, then the index must respect this constraint when aggregating; or if there are several dimensions, but no meaningful basis has been established for comparisons, then the dimensions should not be combined. These latter two examples are closely linked to the broader concern of robustness, which requires results derived from the index to be robust to allowable changes in its components.

The fifth characteristic brings real world data collection and choice into the picture, requiring that the index can actually be implemented using available data. This carries with it restrictions on the qualities of the data to be used, such as the following considerations. Clearly the purpose and the coherence desiderata require the coverage of the data to be sufficient across countries and dimensions. Data should be homogenized across countries to ensure that the individual variables are comparable and be suitable for evaluating countries at all levels of statistical capacity. This criterion also suggests that the index should be based on publicly verifiable data compiled, collected, and disseminated by the NSS. As such, the actual NSS products—such as the census, key SDG indicators, and other variables—could provide a reliable and cost-effective basis for an index designed to measure its statistical capacity.[8] Implementability requires a verification that this is true.

The sixth characteristic also places constraints on data but focuses on the additional dimension of time. Not only should the index be able to be calculated and compared across countries at a given point in time but, consistent with its purpose, it should also be possible to construct the index across many points in time. Consequently, the data requirements should also be maintained through time to allow reasonably consistent estimates of progress and this should translate to a time path that is coherent and predictable. In contrast, if an index were to exhibit undue volatility through time this could suggest replicability problems stemming from the data, the aggregation method, or from both. It should be noted that replicability is also supported by the simplicity

[8] Again, other sources, such as expert opinions or intensive studies carry with them a number of problems, including high monetary cost and increased measurement error due to country-specific orientation and idiosyncratic or irregular collection. The choice of data will be further discussed below.

desideratum, which can lower the cost of replication by others and hence ensure that, through many eyes, there are fewer errors.

The seventh and final criterion concerns the interaction between actions by the NSS and the behavior of the index through time. In essence, a country should not see its index of statistical capacity fall when the only movement between time periods is an improvement in individual variables. Some global indices are constructed in such a way that the index value itself depends on data from *other* countries. A given country can achieve absolute improvements in all facets of the measure and still have its index value drop.[9] This possibility can disrupt the clear pathway between a country's policies and the index level, thereby diminishing the government accountability that the index is meant to support.[10] Alternatively, suppose the index adopts a goal-oriented or targeted approach, by which targets are set and progress is measured as the extent of movement towards reaching them. A target may be fixed through time – like an absolute poverty line – when conditions change, or it might vary with the outside conditions – like a relative poverty line would. In the latter case, it can be unclear whether the progress (or lack thereof) is due to changes in the underlying variables of interest or changes in the target. Incentive compatibility suggests using fixed or absolute targets that link the measure to actual progress.

In sum, constructing a measure of statistical capacity entails many distinct choices that can appear to be arbitrary and unrelated to one another if no context is provided. A set of desired characteristics, or desiderata, can provide the guiding principles that help organize these choices

[9] This was a problem with the traditional Human Development Index, which relied on empirical goalposts (or its highest and lowest values for each dimension) to fix its minimum and maximum values. See Alkire and Foster (2010).
[10] Undue volatility of an index, where the changes in index level appear to be spurious or unrelated to actual policies, can likewise disrupt the policy relevance of the index as an accountability tool. We return to more discussion in the next section.

to obtain a relevant and useful measurement tool. The desiderata described above will be invoked several times in what follows.

3.2.4. Dimensions

The concept adopted here views the statistical capacity of an NSS in terms of its range of products, and the processes it uses to produce and disseminate them. The production process for statistical outputs has certain similarities to the traditional production model from economics, and this analogy might be useful for identifying the salient dimensions of a statistical capacity measure. The production process begins with a technology that is used in generating the statistical products, and the level of this technology is clearly a relevant component of statistical capacity. The resulting statistical outputs might be divided into two general categories. First are the intermediate products, which have direct use for specialists but require additional processing to create products suitable for general use. For example, a census can be helpful for policy analysts but must be processed to obtain useful statistics. Second are the final products, which are available in a form that can be understood by the public. The key macro statistics of a country would naturally be viewed as final products. Even after the products have been created, their existence does not imply that potential users will actually have access to them. Statistical products may be available to only a few users, or available to all. The final dimension then covers the extent to which statistical products are disseminated.

This simple framework helps to identify four coherent dimensions for a measure of statistical capacity, namely: (i) Methodology, Standards and Classifications (MSC), which provides information on the technology being used by the NSS; (ii) Census and Surveys (CS), which describes the intermediate products of the NSS; (iii) Availability of Key Indicators (AKI), which focuses on key final products needed for policy; and (iv) Dissemination Practices and Openness (DPO), which evaluates the extent to which products are publicly disseminated. It is easy to see that each of these dimensions is centrally related to the statistical capacity of an NSS.[11] We return to more discussion on these four dimensions in Section 4.1.

As with any measurement exercise, there will likely be dimensions missing from the list. One dimension that could be particularly relevant for measuring statistical capacity is *flexibility*, or the ability of an NSS to adapt to changing circumstances.[12] For example, in response to a change in national priorities, the NSS may need to alter the schedule or scope of data production. Alternatively, as new technologies for data gathering are being developed, international guidelines may change, requiring the NSS to alter its methods. However, to fully capture the missing dimension of flexibility, one needs data about circumstances that have not yet have happened, and this directly conflicts with the implementability desideratum.[13] Likewise, the absence of such data in the present context leads to the use of more observable variables.[14] A second missing dimension might be *statistical infrastructure* as represented by the underlying array of inputs that the NSS

[11] It may be useful to briefly discuss whether one dimension is inherently more important than another. For instance, since earlier dimensions are needed for the latter dimensions, they can be arguably viewed as being more important. Yet, from the point of view of stakeholders, the opposite could also be seen as true. Notice that both the first and last dimensions could be considered to be *process* dimensions, while the middle two are *product* dimensions. We could posit that process and product are equally important to evaluating statistical capacity, while making a judgment that the individual components are also of equal importance. If so, then the four dimensions would be consistent with Atkinson et al (2002) who observe that "the interpretation of a set of indicators is greatly eased where the individual components have degrees of importance that, while not necessarily exactly equal, are not grossly different."

[12] For fundamental discussions of flexibility in the context of production, see, for example, Kreps (1979).

[13] Parts of this aspect might be captured in the technology or product dimensions: if the country is able to produce a best practice census, then it is plausible that it could produce comparably sophisticated products if and when needed. An analogy can be seen with Sen's notion of capability, which evaluates a person's well-being via sets of available options, or capability sets, which require information on alternatives that have not yet been chosen. The difficulty of obtaining counterfactual information has led many researchers to measure capability using actual achievements or functionings.

[14] One related critique is that since the variables are based on past data, the resulting measure cannot be "forward looking" enough to convey the future possibilities for a given NSS. This inevitable consequence is analogous to the challenge of measuring "vulnerability to poverty" using household survey data. See Dang and Lanjouw (2017) for a recent study that addresses this challenge.

has at its disposal for producing and disseminating data. This could include the physical plant, the quantity and quality of computing facilities, and the number and skill level of workers, among other inputs. All of these are clearly relevant information for gauging the capacity of an NSS. However, it is also clear that implementability would once again be a problem, since gathering accurate, homogenous data on inputs would be a costly and drawn out task.

3.2.5. Variables and Aggregation

Once the dimensions have been specified, attention turns to identifying the variables and selecting an appropriate aggregation method. The desiderata suggest that the variables in an index should be coherent with the concept being measured; they should be publicly available (i.e., drawn from NSSs, rather than from experts or other sources). The aggregation method should be selected with rigor in mind, including the axioms or properties that the method satisfies. At the same time, it should aim for simplicity to maximize general understanding and impact.

We turn next to discussing our proposed aggregation method. To help focus discussion on policy relevance and conceptual issues, we leave the more technical details to Appendix 1, Part A. We also offer a simple numerical example that further illustrates the ideas in this subsection in Appendix 1, Part B.

Basic Setup

In the domain of statistical capacity, a variable is typically derived from a simple "yes-no" question concerning a normative guideline that a country's NSS should meet. For example, in the MSC dimension, a relevant question is whether a modern system of national accounts is being used. In the CS dimension, a question might ask whether there has been a population census within the last ten years. Such goals and targets can be found in discussions of standards and best practices or international guidelines for an NSS.[15] The resulting data on attainments are reported by the NSS to the World Bank, the IMF, the UN, the ILO and other international organizations, or can be gathered directly from the NSS website.

Each of these "yes-no" questions generates a dichotomous variable having a 0-1 representation, where 1 means that the underlying test or target has been successfully achieved, while 0 indicates it has not. When there are $V \geq 2$ dichotomous variables a_v for $v = 1, \dots, V$, then, the basic data are given by an achievement vector or profile $a = (a_1, a_2, \dots, a_V)$ summarizing the test or targets achieved by the NSS. As noted by Atkinson (2003), when variables are dichotomous (or can be dichotomized), a measurement approach called a "counting method" is applicable and, indeed, has become standard for many types of measurement exercises.[16] We propose to apply this method here to construct our index.

A counting method begins with a vector $w = (w_1, w_2, \dots, w_V)$ containing the values or weights $w_v > 0$ that will be used to assess the various achievements. The resulting counting index C is defined by

$$C = C(a; w) = \frac{w \cdot a}{w \cdot u} = \frac{w_1 a_1 + \cdots + w_V a_V}{w_1 + \cdots + w_V} \tag{1}$$

where $w \cdot a$ denotes the inner product of w and a, while $u = (1, \dots, 1)$. Intuitively, the counting index calculates the sum of the values of the achievements in a as a share of the maximum total value that could be achieved. Equivalently, C is the weighted mean $\mu(a; w)$ where

[15] Indeed, all the indicators under the dimension of "Availability of Key Indicators" are motivated by the SDGs and can be mapped to the SDG indicators.
[16] See, for example, the adjusted headcount ratio of Alkire and Foster (2011), the social exclusion index of Chakravarty and D'Ambrosio (2006), the Women's Empowerment in Agriculture Index (http://ophi.org.uk/policy/national-policy/the-womens-empowerment-in-agriculture-index/) and the Better Jobs Index (https://mejorestrabajos.iadb.org/en/indice).

$$\mu(a;w) = \frac{w_1}{w_1+\cdots+w_V}\,a_1 + \cdots + \frac{w_V}{w_1+\cdots+w_V}\,a_V \quad (2)$$

so that the weight on each a_v is $w_v/(w_1 + \cdots + w_V)$.[17] Clearly, C takes on values between 0 and 1.

A key challenge is how to discern the relative importance of the different attainments in order to aggregate up the counts. One overarching structure for doing so is found in the nested method of Alkire and Foster (2011), which extends Atkinson's (2003) "equal importance" construction to subdimensions and variables. For example, in the case of statistical capacity, it could be argued that the CS dimension should be divided into two equally important subdimensions, one for censuses and a second for surveys. Likewise, variables that reside within a given subdimension (or dimension if it has no subdimensions) should be selected to have roughly equal contributions within the group. A nested approach helps to account for the relative values of variables, subdimensions and dimensions in a fashion that is coherent with the concept being measured.[18]

Useful Properties

Once w has been set and the counting index C has been defined, it offers four useful properties we examine below. First, C is additively decomposable by subsets of variables; second, C is additively decomposable by subsets of countries (or regions). Third, C can be constructed using

[17] Although the entries of w will typically sum to 1, this general definition allows the sum to exceed or fall short in line with different approaches to counting indices. This generality is also convenient in defining sub-indices that use only a subset of the values in w.

[18] Other empirical methods can be employed to infer the importance of individual variables, such as applying principal component techniques to obtain weights from the principal component combination that explains the most variation. This approach has the benefit of letting the data speak for themselves. However, estimation results are completely dependent on the timing and coverage of the data set used to calibrate the weights, which raises hard-to-answer questions: How far back in time should the data extend? What should be done when new data come online? Can an outside indicator of what is being measured actually be found? As such, empirical methods may yield data-oriented weights that totally ignore variables conceptually considered to be important. See also Decancq and Lugo (2013) for a discussion of empirical methods for finding weights.

either dichotomous or multi-valued variables. Finally, C allows for different emphasis on progress for countries at different outcome levels. We offer a formal statement of the first property below.

Proposition 1. Decomposability into Subsets of Variables
Let S and S' be two nonempty sets of variables with empty intersection whose union is $\{a_1, \dots, a_V\}$. *Define the two associated sub-indices of* C *as*

$$C^S = \mu(a^S; w^S) = \frac{w^S \cdot a^S}{w^S \cdot u^S}\,; \qquad C^{S\prime} = \mu(a^{S\prime}; w^{S\prime}) = \frac{w^{S\prime} \cdot a^{S\prime}}{w^{S\prime} \cdot u^{S\prime}}$$

where w^S, a^S *and* u^S *(or* $w^{S\prime}$, $a^{S\prime}$ *and* $u^{S\prime}$*) are obtained from* w, a *and* u *by removing the variables outside of S (respectively, S').*

The index C can be decomposed into these two sub-indices as follows

$$C = \frac{w^S \cdot u^S}{w \cdot u} C^S + \frac{w^{S\prime} \cdot u^{S\prime}}{w \cdot u} C^{S\prime} \tag{3}$$

Proof: Appendix 1, Part A.

A couple remarks are in order for Proposition 1. First, notice that $w^S \cdot u^S$ gives the sum of values for variables in *S*, while $w \cdot u$ is the sum of all values, and hence the coefficient on C^S is the share of the total value from variables in *S*; an analogous interpretation holds for *S'*. As such, the relative contribution of each group of variables to the overall index can be obtained by dividing the two terms on the right-hand side of Equation (3) by the overall index value *C*.

Second, the above decomposition for *C* applies to any choice of *S* and *S'*. Consequently, it can be generalized to more than two choices, such as the four dimensions of the index discussed earlier. For example, define the associated dimensional index $C^{(d)}$ as

$$C^{(d)} = \mu(a^{(d)}; w^{(d)}) = \frac{w^{(d)} \cdot a^{(d)}}{w^{(d)} \cdot u^{(d)}} \tag{4}$$

where $w^{(d)}$, $a^{(d)}$ and $u^{(d)}$ contain only the data associated with variables in dimension *d*, for any $d = 1, 2, \dots, D$. We can then obtain the overall index *C* as the average of all the dimensional indices

$$C = \sum_{d=1}^{D} \frac{1}{D} C^{(d)} \tag{5}$$

Finally, in the process of defining weights for the index, we also need to define a tree structure of dimensions (and subdimensions) for the variables, where all the variables in the same dimension (or subdimension) have the same weight. This useful feature allows us to update the index over time by, say, adding variables and changing the composition of a relevant dimension while keeping the other dimensions fixed.[19]

The second feature broadens the decomposability of the index to cover subgroups of countries or regions.

Proposition 2. Decomposability into Subsets of Countries
Let N be the number of countries considered and let A denote the $N \times V$ achievement matrix whose n^{th} row a^n is the achievement vector for country n = 1,…,N.
Define the index applied to the collection {1,…,N} of countries as
$C = C(A; w) = \mu(A; w)$, *where the weighted mean $\mu(A; w)$ of matrix A is given by*

$$\mu(A; w) = \sum_{n=1}^{N} \sum_{v=1}^{V} \left(\frac{1}{N} \frac{w_v}{w_1 + \cdots + w_V}\right) a_v = \sum_{n=1}^{N} \frac{1}{N}\, \mu(a_v; w) \qquad (6)$$

The index C can be decomposed into the country-level (or region-level) sub-indices as follows

$$C(A; w) = \frac{1}{N} \sum_{n=1}^{N} C(a^n; w) \qquad (7)$$

Proof: Appendix 1, Part A.

Intuitively, $C(A; w)$ provides a way of monitoring regional or global progress that is based upon the progress in individual countries. Thus, for the $N \times V$ achievement matrix for all countries, Proposition 2 offers a decomposition by row (or country), while Proposition 1 (Equation (5)) provides a decomposition by column (or dimension). Note that we can derive from Equation (7) the complementary index

$$M(A; w) = 1 - C(A; w) \qquad (8)$$

that evaluates the extent to which countries in the region are falling short of targets.

[19] But note that the main decomposition axiom will be defined with respect to the basic groups.

The nested counting index is well suited for dichotomous variables and its simple aggregation method – a weighted mean – is easy to interpret.[20] Yet, while each dichotomous variable represents a separate test or target, some tests may be related, giving rise to additional possibilities for variables and aggregation. For example, two distinct questions might refer to the system of national accounts being used by the NSS, where there are two levels of standards in common use. A first variable might ascertain whether the NSS is at least using the lower standard; the second might determine whether the higher standard is being employed. The three feasible combinations for the two variables are: (0,0), which indicates that neither system is being employed; (1,0), which indicates that the lower and not the higher standard is being employed; and (1,1), which indicates that the higher standard system is being employed.[21]

It is possible to combine the two related variables into a single, non-dichotomous variable that would assign unique values to the three levels, say 1 to the higher outcome (1,1), some numerical value λ between 0 and 1 to the lower outcome (1,0), and 0 to the outcome where neither standard is in place (0,0). Constructed variables like this are commonly used in measurement exercises. Yet, assigning an arbitrary numerical value λ to a variable does not by itself make the variable cardinally meaningful or comparable to other variables. Consequently, the resulting index and country rankings will simply not be robust to admissible changes in values.[22]

[20] Notably, other options exist for aggregating dichotomous variables, but these have limitations. For example, one option is the class of general means (or means of order α), which includes the geometric mean as a special case. Since variables here have 0 values, only general means with positive parameter α are applicable (the geometric mean is also problematic in that it takes on a value of 0 whenever any variable is 0 and thus violates monotonicity). Furthermore, a general means would rank countries the same way but lose the useful decompositions and simple interpretation of the index value. Another option is a dashboard approach, where disaggregated variables are presented in their natural state. Yet, this approach does not offer easy interpretation, especially in our case where the number of variables is large.

[21] The outcome (0,1) would not be observed in this case of related variables, since achieving the higher level entails achieving the lower level.

[22] Both cardinality and comparability can be defined in terms of admissible changes in the scales used to represent variables.

Our index provides a systematic method to avoid such mistakes, which is stated in Proposition 3 below.

Proposition 3. Equivalence between Dichotomous and Multi-Valued Indicators
Suppose that there are T many related dichotomous variables providing information on ascending levels of a given aspect of statistical capacity. Let S' be the set of subscripts for these T variables so that $a^{S'}$ denotes the relevant vector of achievements. Let $C(a^{S'}; w^{S'})$ be the counting index for this subset of variables; it assigns the T+1 options for $a^{S'}$ the numerical values of 0, 1/T, …, (T-1)/T, and 1, respectively. Now replace the T dichotomous variables with the single variable x having (one of) these T+1 values and assign it the total weight inherited from the original T variables. The weighted mean across x and the remaining dichotomous variables generates identical index values as the original counting measure and can be used equivalently.

Proof: Appendix 1, Part A.

Since the constructed single variable has a linear scale (i.e., with equal increments between adjacent levels), the resulting index is neutral in gauging improvements at different starting values. In certain circumstances, though, it might be helpful to have a range of indices that emphasizes improvements at different outcome levels (i.e., the incentive compatible desideratum). For example, for higher capacity nations that have covered all the basics, it could be useful to have an index that emphasizes improvements at the upper level. Likewise, for lower capacity nations that have not yet achieved basic levels, an index that emphasizes improvements at lower levels might have merit. Our index can be modified to accommodate these cases, following the intuition found in the poverty measurement literature.[23]

Let $C_\alpha(a'; w')$ be defined by

$$C_\alpha(a'; w') = \frac{w_1}{w_1 + \cdots + w_V}(a'_1)^\alpha + \cdots + \frac{w_V}{w_1 + \cdots + w_V}(a'_V)^\alpha \qquad (9)$$

for positive α. Intuitively, the index first transforms each variable by a positive parameter, then takes the weighted average of the transformed levels. As noted above, this does not change the

[23] See for example Foster, Greer, and Thorbecke (1984).

values of dichotomous variables; only the constructed variables will be affected by the transformation. For α < 1, the transformation is concave and places greater emphasis on improvements at the lower end of the outcome range; for α > 1, the transformation is convex and places greater emphasis on improvements at the higher end; and for α = 1, the index becomes $C_\alpha(a';w') = C(a';w')$, the usual counting index. While C is the central index that will be used, other indices from the parametric class can help focus on high-level or low-level improvements when needed.

3.2.6. Axiomatics

Axioms are rigorous properties for an index to satisfy, and they are more formalized and generalized than the properties discussed earlier in Subsection 3.2.5. Axioms help in understanding what an index is actually measuring and in deciding which index to use. Knowing which properties an index satisfies can help in interpreting the empirical results obtained using that index; certain forms of policy analysis become possible only when the index satisfies a given property. Some axioms can be interpreted as "nuggets of policy" that specify the kinds of changes that should leave the index value unchanged and those that should alter it. Others break down the index value to help understand how dimensions contribute to that value. Our proposed index satisfies three axioms, which include symmetry, monotonicity, and subgroup decomposability.

As noted in Foster et al (2013), axioms can be usefully grouped into three categories: invariance axioms, which indicate what not to measure; dominance axioms, which indicate what the index should measure; and subgroup axioms, which break down or build up indices by variables or units of analysis. The three axioms our proposed index of statistical capacity satisfies are closely related with these three groups of axioms. In what follows, a generic index of statistical capacity over profiles $a = (a_1, \dots, a_V)$ will be denoted by F.

Symmetry: in other measurement environments where the number of people, dimensions, or other factors may differ across comparisons, invariance axioms are often used to ensure consistency. In the present context, the index F is being applied to one country's data with a fixed number of dimensions and dichotomous variables, so properties of this sort are not needed. A second common form of invariance axiom is anonymity or symmetry whereby the index value is unaffected when variable levels are switched. In the present context, where the variables have a structure as represented by hierarchical tree T and partition P, universal symmetry is not appropriate. Motivated by Basu and Foster (1998), one might consider a weaker form of symmetry that is contingent on variables being "similarly placed" in the variable structure. We say that profile b is obtained from profile a by a *basic switch* if $b_v = a_{v'}$ and $b_{v'} = a_v$ for some $v \neq v'$ in the same basic group, while $b_{v''} = a_{v''}$ for all other v''. In other words, the only difference between b and a is that two variable values in the same basic group have been switched. A statistical capacity measure F satisfies *basic symmetry* if $F(a) = F(b)$ whenever b is obtained from a by a basic switch. Notice that any nested counting index C satisfies basic symmetry because it has the same weight on every variable in the same basic group.

Monotonicity: the main axiom for F is an intuitive dominance axiom requiring the index value to reflect improvements in variables. We say that profile b is obtained from profile a by an *improvement* if $a_v \geq b_v$ for all v, and $a \neq b$ or, in other words, if profile a vector dominates profile b. A statistical capacity index F satisfies *monotonicity* if $F(a) > F(b)$ whenever a is obtained from b by an improvement. This simple but significant requirement ensures that the index value rises whenever one variable rises from 0 to 1 and the rest of the variables do not fall in value. The index $C(a; w)$ satisfies this property since each w_v is strictly positive. Notice that monotonicity supports the incentive compatibility criterion, since it ensures that a country is not penalized when it successfully raises its profile.

*Subgroup decomposability: s*ubgroup axioms allow the index to be divided into salient sub-indices and linked back to the original index for policy analysis. In the present case, the main decomposition is over the basic groups given in partition $P = (p_1, \dots, p_K)$. A statistical capacity index *F* satisfies *basic decomposability* if there exist weights $\rho_k \geq 0$ summing to 1 and sub-indices $c_k(p_k)$ such that

$$C(a) = \Sigma_{k=1}^{K} \rho_k c_k(p_k) \tag{10}$$

In other words, there is a collection of indices, one for each basic group of variables, such that *C* can be expressed as a weighted average of these basic indices. This is clearly the case for the nested counting index $C(a; w)$, as it is based on a weighted mean. Likewise, Equation (5) (after Proposition 1) follows from Equation (10) by aggregating across basic groups within each dimension, so that the overall index value is just the average of the dimensional index values. These decompositions can help inform why one country is doing better than another or help describe how a single country is progressing over time.

As noted above, the single country index $C(a; w)$ can be expanded into an index $C(A; w)$ that covers all countries in a region or even the universe of covered countries. The formula used to do this – Equation (7) – doubles as another form of decomposition that expresses the aggregate index and an average of the country indices. Since $C(a; w)$ is the index of primary interest here, the equation will not be expressed as a formal property here. However, the fact that Equation (7) and

$C(A; w)$ are available allows users to have a better understanding of regional levels and trends in statistical capacity.

4. Description of the SPI

We apply the theoretical framework discussed above to develop a new Statistical Performance Indicator (and Index, or SPI). In this section, we discuss the features of the SPI, and compare it with the SCI. We also briefly discuss some data challenges with the construction of the SPI.

4.1. SPI and Its New Features

As discussed earlier in Subsection 3.2.4, the SPI is built around four main dimensions: i) Methodology, Standards and Classifications, ii) Censuses and Surveys, iii) Availability of Key Indicators , and iv) Dissemination Practices and Openness. We further discuss each dimension, its indicators, and other practical implementation details in this section. More details on each indicator, its score, and its data sources are provided in Appendix 2.

Internationally accepted and recommended methodology, classifications, and standards provide the basis for NSSs to generate internationally relevant statistical indicators and facilitate data exchange and data integration. The first dimension (Methodology, Standards and Classifications) assesses whether NSSs have the necessary capacity to adopt and comply with international statistical standards. Twelve indicators are proposed for the first dimension to ensure that data from different sources can be compared, and to inform data users that the data are reliable and meet established technical standards. The setting and applying standards are usually regulated by law and typically the responsibility of the main national statistical agency.

The second dimension (Censuses and Surveys) reflects the availability and frequency of major censuses and surveys, which are designed to collect information mandated by the National Statistical Acts. This dimension comprises of eight indicators on population and housing census, agricultural census, business census, income and expenditure surveys, and other surveys on agriculture, health, labor force, and establishments.

The third dimension (Availability of Key Indicators) evaluates NSSs by reviewing the availability of the country data for the most recent year in international databases. In addition, the selected indicators produced by NSS should address the development concerns of countries, especially with the SDGs. The 12 indicators included in this category range from the headcount poverty rate to the under-five mortality rate, the primary school completion rate, the manufacturing value added in GDP, and net trade in goods and services. These indicators overlap to large extent with the Tier 1 SDGs, and include other standard macro-economic indicators as well, which are conceptually clear and regularly produced using established methodology and standards.

Providing data to the relevant international agencies demonstrates that an NSS's practices meet the quality standards and data production timeliness. The last dimension (Dissemination Practices and Openness) is built on the principle that quality statistics should be delivered to the public in a timely, easily accessible manner and free of charge. It includes 10 indicators grouped under two sub-sections: Dissemination Capacity of NSO and Openness of Data.

These four dimensions are closely linked and capture the production cycle of NSSs in collecting, producing, and disseminating high-quality statistics. By following internationally recommended standards and classifications, an NSS can produce data of good quality that are both comparable within the country over time and across different countries. By combining administrative sources and timely censuses and surveys, an NSS can collect, process, and generate data products covering different aspects of households and establishments. Finally, an NSS can disseminate these data products through its official websites, regular publications, and by submitting them to relevant international organizations, which further strengthen data transparency and quality.

Table 2 provides a summary of the data sources of the different indicators that comprise the SPI. Since we use data from more than one source to construct an indicator, the numbers shown in this table are the maximum numbers of indicators each data source contributes to. Table 2 shows that the indicators produced by the World Bank account for more than half of all the indicators, followed by the indicators produced by the IMF and the NSOs; other UN agencies contribute the remaining indicators.

4.2. Comparing the SPI and the SCI

The SPI has several advantages over the SCI, particularly in terms of data coverage. Table 3 shows that SPI has:

i) Richer and more comprehensive dimensions covering different data aspects ranging from data generation, curation, and dissemination to data analysis.

ii) More indicators: the SPI has 42 indicators (of which 39 are used for scoring), versus 25 indicators in the SCIs.

iii) More countries: the SPI covers more than 200 countries, especially including high-income countries, while the SCI covers fewer than 150 countries and includes no high-income countries.

More importantly, the SPI is built on the conceptual and theoretical framework laid out above (Section 3), while the theoretical principles of the SCI are not clearly formulated. The SPI satisfies all the desiderata of a statistical capacity index: simple, coherent, motivated, rigorous, implementable, replicable, and incentive compatible. Furthermore, from an institutional viewpoint, the SPI is designed to be better aligned with other capacity assessment tools. For example, it is consistent with the Systematic Country Diagnostics (SCD), a new assessment tool that the World Bank recently developed to identify a country's priority development areas.

On the other hand, the increased numbers of indicators and countries covered for the SPI require more data compared to the SCI. The data for the SPI come from different sources, ranging from established databases with international organizations to national statistical agencies' websites. Given this heterogeneity of data sources, some data inconsistency may be expected. The complexities with the former are not just simply linearly related to the increase in quantities, but also concern data quality challenges such as missing data. Missing indicators for less-developed countries can represent a quite different data situation from missing indicators for richer countries. While the former may suggest weak statistical capacity, the latter might be explained by the little need of tracking certain indicators. A specific example is child stunting. In the SPI database, we assign maximum scores to richer countries (including the OECD and the EU) for this indicator assuming that richer countries have achieved the desired target of reducing child stunting and provide the benchmark against which progress for developing countries can be measured.

5. Some Empirical Illustrations

5.1. Data

We have constructed the database for the SPI for 2016. Efforts are underway to construct indicators for the SPI after 2016, as well as going backward several years for better comparison. We provide below some empirical illustrations using the available SPI data and various other data sources including the World Bank's World Development Indicators database (World Bank, 2018b), MIT's Economic Complexity Index (see, e.g., Hidalgo and Hausmann (2009)), and UNICEF's SDG indicators for child welfare (UNICEF, 2018).

5.2. Empirical Illustrations

We discuss in this subsection the relationship between the SPI scores and some major characteristics of a country such as its GDP, population size, and economic complexity. We subsequently offer a decomposition of the SPI by region, before providing some further empirical comparison between the SPI and the SCI.

SPI and Country Characteristics

We start first with examining the relationship between the SPI and a country's log GDP per capita (Figure 1). We expect richer countries to have better statistical systems because they have more resources to allocate for statistical activities, and also because they tend to have more complex and diversified economies that require more data. Indeed, there is a statistically significant and positive correlation between a country's SPI and its income level. Figure 1 suggests that a 10 percent increase in a country's GDP per capita is associated with approximately a 0.7 percentage point increase in its SPI score (see the box inside the figure).[24]

As discussed earlier, a country's statistical system produces statistics that reflect the socio-economic conditions of the nation. As such, the more complex (advanced) a country's economic level is, the more likely that its NSS is more developed. Indeed, besides an inherently stronger demand for the NSS to keep track of its various economic activities, a richer country likely has

[24] Alternatively, moving up an income category (as defined by the World Bank (2018a)) can see a country improve its SPI score by 4.9 percentage points (Figure 1.2, Appendix 1).

more resources to invest in its NSS. To examine this hypothesis, we plot in Figure 2 a country's SPI against its Economic Complexity Index (ECI). This figure indicates a strongly positive relationship between the two indexes, with a 0.1 increase in a country's ECI being associated with a 1.5 percentage point increase in its SPI score.[25]

A Regional Decomposition of SPI Scores

We turn next to decomposing the SPI score by region in Figure 3. This figure demonstrates that among the six regions, Europe and Central Asia (ECA) and South Asia (SAS) perform above the global average, with the remaining four regions falling behind in the following decreasing order: Middle East and North Africa (MENA), East Asia and Pacific (EAP), Latin American (LAC), and Sub-Saharan Africa (SSA). Notably, ECA stands out from the other regions, which have a more similar score: the difference between its score and that of the second-strongest region SAS is 15 percentage points, which is more than the difference between the latter region's score and that of the weakest-performing region SSA.

It can be useful to further examine whether the patterns seen in Figure 3 apply to scores in all four dimensions. Disaggregating the regional scores further by the four dimensions can help us disentangle which dimensions drive these differences and can be improved. Estimation results are shown in Figure 4, which offers several interesting observations. First, a country's score on a certain dimension can be quite different from that of its overall SPI score. In particular, all dimensions—except for the DPO dimension—have a ranking order for country performance that is different from that with the overall score. An interesting example is the ECA region, which is consistently the best performer on all dimensions but AKI, where the SAS region is now the best

[25] For comparison, the US and the UK both have an ECI of 1.6, while the corresponding figure for China and Mexico is 0.9.

performer. On the other hand, SSA performs slightly better than the global average on the AKI dimension, but is the weakest performer on the rest. Second, countries generally perform best in the AKI dimension, where their average dimension score is 77 percent, which is almost the corresponding figure for the other dimensions.

Further Comparison of the SPI versus the SCI

We turn next to comparing the SPI and the SCI for 2016, when data for both indexes are available. We restrict our comparison sample to the 146 countries covered by both indexes. Estimation results suggest that richer countries, or countries with a more complex economy tend to have higher SPI scores. Indeed, for this sample of countries, a 10 percent increase in a country's GDP per capita is associated with approximately a 0.6 percentage point increase in its SPI score, twice the corresponding increase for the SCI. Similarly, a 0.1 increase in a country's ECI being associated with a 1.5 percentage point increase in its SPI score but only a 1.1 percentage point increase in its SCI score. For further visual illustration, we plot in Figure 5 the SPI and SCI scores against log of GDP per capita (Panel A) and economic complexity index (Panel B). While we leave out the regression results shown in previous figures for lack of space, the slopes of the fitted lines for the SPI scores are clearly steeper than those for the SCI scores in both these graphs. Furthermore, when the standardized distributions of both indexes are plotted on the same graph (Figure 1.3, Appendix 1), the SCI has a shorter tail on the right, suggesting that it has a less full range than the SPI.[26]

[26] Since the two indexes have different scales, we standardize the two indexes so that they both have mean 0 and standard deviation 1 for better comparison in Figure 1.3. In addition, this figure also indicates that the SPI appears more similar to the normal distribution than the SCI, which is confirmed by formal statistical tests. One possible reason

Unsurprisingly, the rankings of countries are different for the two indexes. A particularly interesting case is El Salvador, a lower-middle-income country in Latin America.[27] This country ranks as number 5 of 145 countries based on the SCI scores. This would place this country in a surprisingly higher position than many richer countries, including its regional upper-middle-income neighbor—Colombia—which is ranked at number 10. Yet, according to a recent evaluation for 44 indicators concerning children in the 2030 SDG agenda by UNICEF, El Salvador has inadequate data for tracking progress on more than half (i.e., 27) of these indicators. The same figure for Colombia is much lower at just 14 indicators (UNICEF, 2018). The SPI ranks El Salvador as 48 while the ranking of Colombia changes to 12.

Another example is Botswana, an upper-middle-income country in Africa. This country has an SCI rank of 122, near the bottom of the SCI scores. This places Botswana lower than Madagascar (number 89) and Zimbabwe (number 105), both low-income countries in Sub-Saharan Africa. Indeed, Botswana's GDP per capita was $15,807 in 2011 PPP, which was respectively 11 times and 8 times higher than that of Madagascar and Zimbabwe (World Bank, 2018b). Furthermore, a recent study suggests that Botswana has accomplished both upward consumption mobility and pro-poor growth in the past decade, which reduced its headcount poverty rate to 18 percent (using the international poverty line of $1.9/ day in 2011 PPP dollars). The opposite story held Madagascar, where the country has undergone much downward mobility, which drove its poverty rate up to as high as 82 percent (Dang and Dabalen, in press). Zimbabwe, on the other hand, has a higher income level than Madagascar, but is observed to have implemented at most one household consumption survey during the period 1990-2012; the corresponding figure for Botswana is two

is that, since the SPI is the average of a larger number of indicators than the SCI, it is more likely to have a normal distribution by the central limit theorem (see, e.g., Casella and Berger (2002)).

[27] Unless otherwise noted, we use the World Bank's (2018) income classification for all countries.

comparable surveys (Beegle et al., 2016). Perhaps these studies can provide supportive evidence to help understand why Botswana sees its SPI ranking go up significantly to number 52, while the corresponding figure for Madagascar falls to number 133, and that of Zimbabwe remains almost unchanged.

6. Conclusion

In this paper we offer a systematic discussion of the construction of the Statistical Performance Index. Our theoretical and empirical methodology builds on the existing strengths of the widely-used Statistical Capacity Indicator of the World Bank. We present the conceptual and mathematical foundation behind this index, and significantly expand the number of indicators and countries covered by this index. We also provide some brief empirical illustrations of the updated index with global data in 2016.

This new index could be seen as the first step before more resource-intensive country-specific assessments to inform multi-year improvement plans. Our proposed framework is also flexible enough to allow for future revisions as the global data landscape evolves. For example, we can incorporate new indicators such as whether an NSO uses cloud computing to store their data or implements household panel surveys in the relevant dimensions without creating major changes to the total scores. Our framework may also be relevant to the construction of other indexes in related areas, such as tracking the global SDGs or child development. Since the SPI is currently available only for 2016, another promising direction is to collect time series data for the new index, both going forward and for several years past, and expand the current framework to allow for dynamic changes over time.

References

Alkire, Sabina and James Foster. (2011). "Counting and multidimensional poverty measurement." *Journal of Public Economics*, 95(7): 476-487.

Alkire, Sabina, José Manuel Roche, Paola Ballon, James Foster, Maria Emma Santos, and Suman Seth. (2015). *Multidimensional Poverty Measurement and Analysis*. United Kingdom: Oxford University Press.

Atkinson, Tony, Bea Cantillon, Eric Marlier, and Brian Nolan. (2002). *Social Indicators: The EU and Social Inclusion*. New York: Oxford University Press.

Atkinson, Anthony B. (2003). "Multidimensional deprivation: contrasting social welfare and counting approaches." *Journal of Economic Inequality*, 1(1): 51-65.

Beegle, K., Christiaensen, L., Dabalen, A., & Gaddis, I. (2016). *Poverty in a Rising Africa*. Washington DC: The World Bank.

Casella, George and Roger L. Berger. (2002). *Statistical Inference*, 2nd Edition. California: Duxbury Press.

Chakravarty, Satya R. and Conchita D'Ambrosio. (2006). "The Measurement of Social Exclusion". *Review of Income and wealth*, *52*(3): 377-398.

Dang, Hai-Anh and Andrew L. Dabalen. (in press). "Is Poverty in Africa Mostly Chronic or Transient? Evidence from Synthetic Panel Data". *Journal of Development Studies*.

Dang, Hai-Anh and Peter Lanjouw. (2017). "Welfare Dynamics Measurement: Two Definitions of a Vulnerability Line and Their Empirical Application". *Review of Income and Wealth*, 63(4): 633-660.

Decancq, Koen and María Ana Lugo. (2013). "Weights in multidimensional indices of wellbeing: An overview." *Econometric Reviews*, 32(1): 7-34.

Fantom, N. and N. Watanabe (2008). "Improving the World Bank's Database of Statistical Capacity". *African Statistical Newsletter*, 2(2): 21-22.

Food and Agriculture Organization of the United Nations, Asia and Pacific Commission on Agricultural Statistics (2012). *Report on Initial Country Assessments*. http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Workshops_Events/APCAS_24/Paper_after/APCAS-12-7-__Report_Country_Assessment.pdf.

Food and Agriculture Organization of the United Nations (2014). *Guidelines for Assessing Country Capacity to Produce Agricultural and Rural Statistics*. June. http://gsars.org/wp-content/uploads/2014/09/Guidelines_Country-Assessment_FINAL.pdf .

Foster, James, Joel Greer, and Erik Thorbecke. (1984). "A Class of Decomposable Poverty Measures." *Econometrica*, 52(3): 761-766.

Foster, James, Suman Seth, Michael Lokshin, and Zurab Sajaia. (2013). *A Unified Approach to Measuring Poverty and Inequality: Theory and Practice*. Washington D.C.: World Bank.

Hidalgo, César A. and Ricardo Hausmann. (2009). "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences*, 106(26): 10570-10575.

Kreps, David M. (1979). "A Representation Theorem for 'Preference for Flexibility'." *Econometrica*, 47(3): 565-577.

Ngaruko, Floribett. (2008). "The World Bank's Framework for Statistical Capacity Measurement: Strengths, Weaknesses, and Options for Improvement". *The African Statistical Journal*, Volume 7. November.

Organisation of Islamic Cooperation (2012). Current State of Statistical Capacity in the OIC Member Countries. February.

Partnership in Statistics for Development in the 21st Century (PARIS21) (2002). *Statistical Capacity Building Indicators Final Report*. https://www.paris21.org/sites/default/files/scbi-final-en.pdf.

Sanga, D., Dosso, B., and Gui-Diby, S. (2011). "Tracking Progress towards Statistical Capacity Building Efforts: The African Statistical Development Index". *International Statistical Review*, 79(3): 303-329.

Stamp, Josiah. (1929). *Some Economic Factors in Modern Life*. London: P.S. King and Sons.

Tapsoba, Sampawende J-A., Codjo Neree Noumon, and Robert C. York. (2017). "Can Statistical Capacity Building Help Reduce Procyclical Fiscal Policy?" *Journal of International Development*, 29(4): 407-430.

UNICEF. (2018). *Progress for Every Child in the SDG Era*. New York: UNICEF.

United Nations (2016). *Evaluation of the Contribution of the United Nations Development System to Strengthening National Capacities for Statistical Analysis and Data Collection to Support the Achievement of the Millennium Development Goals (MDGs) and other Internationally-agreed Development Goals.* Technical report.

United Nations Economic Commission for Europe. (UNECE). (2014). Global Assessment Report: National Statistical System of Mongolia. https://www.unece.org/fileadmin/DAM/stats/documents/technical_coop/GA_Mongolia_EN.pdf.

United States Census Bureau. Tool for Assessing Statistical Capacity (TASC)—Version 2.0. https://www2.census.gov/software/tasc/tasc-booklet.pdf

World Bank. (2018a). *Country Income Classification*. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

---. (2018b). *World Development Indicators Online*.

Table 1: Comparing SCIs with other statistical capacity measurement tools

Tools[28] / Criteria	PARIS21	UNECE	FAO	UNECA	US Census Bureau	WB
Country-level data	No	Yes	No[29]	Yes	No	Yes
Countries covered	N.A.	10 (ECA countries)	N.A.	43	N.A.	146 (old); 218+ (new)
Time covered	N.A.	One year	N.A.	N.A.	N.A.	2004-2016 (old SCI); 2016- (new SPI)
Designed data collection method	Primary data collection	Primary data collection	Primary data collection	Primary data collection	Primary data collection	Secondary data collection
Selected indicators						
Number of indicators	16 (quantitative); 18 (qualitative)	5 broad categories, 40-50 specific aspects[30]	24 sub-categories	42 variables	110	25 (old); 42 (new)
Dimensions covered	System-wide indicators; agency-related indicators; data-related indicators.	Legal Basis; National Statistical System; Statistics Authority; Data Sources and Processing Statistical Domains.	Institutional infrastructure; Resources; Statistical methods and practices; Availability of statistical information; Critical constraints	Functioning of the National Statistical Systems; Statistical infrastructure; Data dissemination; Human capital development; Funding	Institutional capacity; Planning and management; Mapping; Sampling; Questionnaire content and testing	Revised SCIs: Methodology, standards and classification; Censuses and Surveys; Dissemination Practices and Openness; Availability of Key Indicators
Operational relevance	• Useful both for national management and international comparison; • More useful to countries that are "statistically challenged". • Facilitate coordination among the organizations involved in technical assistance.	• Recommendations from the reports could be incorporated into statistical capacity-building programs and strategies; • These strategies are then made operational through annual statistical programs of work and implemented by the beneficiary countries.	• CAQ can be used to create country profiles and group countries into quartiles or to identify priority interventions that could be implemented at the regional level. • The in-depth assessment will further provide insights on agricultural and rural statistics.	• Support the monitoring and evaluation of RRSF identify for each African country weaknesses and strengths in order to support interventions; • Provide a general idea of the performance of African countries' statistical systems.	• Aid NSOs in identifying areas of improvement; • Assist NSOs and donors to justify the need for funding for training; • Provide a measure of the impact of capacity building activities by being administered at two points in time, before and after.	• Provide an assessment of country statistical capacity over time in a cost-effective, sustainable way; • Provide guidance to the WB teams in assessing the progress and sustainability of the Bank supported projects; • Inform Systematic Country Diagnostics; • Provide a monitoring tool for countries' SDGs data production capacity.

[28] Partnership in Statistics for Development in the 21st Century (PARIS21): Statistical Capacity Building Indicators; United Nations Economic Commission for Europe: Global Assessment; Food and Agriculture Organization of the United Nations: Country Capacity Indicators; United Nations Economic Commission for Africa: African Statistical Development Index; US Census Bureau: Tool for Assessing Statistical Capacity, World Bank: Statistical Capacity Indicator.
[29] The pilot study was conducted in Asia Pacific countries in 2012. In-depth country assessment will be carried out, following the established guidelines.
[30] Referred to the most recent GA report on National Statistical System of Mongolia. No specific indicators were referred to, but the report focused on assessing different specific areas of the Mongolian statistical system.

Table 2: Number of Indicators by Data Sources

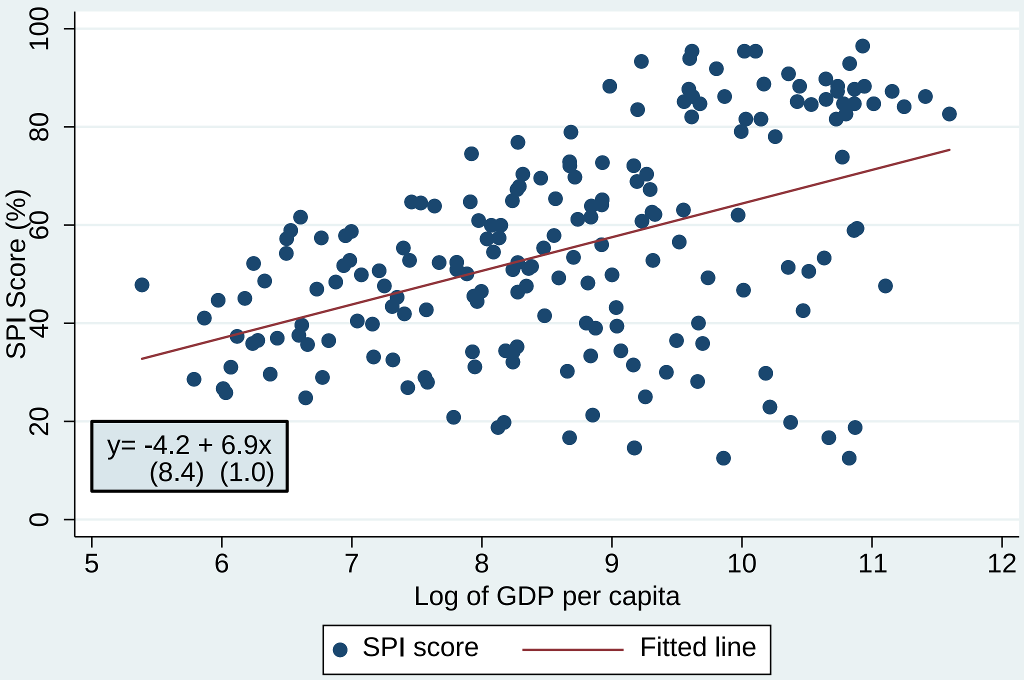
No	Data Sources	Number of Indicators Total	Number of Indicators Percent.
Category A: International Organization			
1	*IMF*		
	Dissemination Standards Bulletin Board (DSBB)	3	6
	Government Finance Statistics (GFS)	1	2
	International Financial Statistics (IFS)	2	4
	World Economic Outlook (WEO)	3	6
	Subtotal	9	19
2	*Other United Nations Agency*		
	UN Economic Commission for Europe (UNECE)	1	2
	UN Industrial Development Organization (UNIDO)	1	2
	UN Statistical Division (UNSD)	2	4
	Food and Agriculture Organization of the United Nations	1	2
	International Labour Organization (ILO)	2	4
	Subtotal	6	13
3	*World Bank*		
	Interational Household Survey Network Catalogu (IHSN)	9	19
	PovCalNet	4	9
	Microdata Library	4	9
	World Development Indicators (WDI)	15	32
	Demographic and Health Surveys (DHS)	1	2
	Subtotal	24	51
Category B: National Statistical Agency			
	National Statistics Office (NSO)	8	17
Total		47	100

Note: One indicator can be constructed based on data from more than one source.

Table 3: Comparing the SPI and the SCI

	SCIs	SPI
Coverage		
Country coverage	146	218+
Time covered	2004-2016	2016 onwards
Selected indicators		
Number of indicators	25	42
Dimensions covered	Methodology; Source Data; Periodicity and Timeliness	Methodology, Standards and Classifications; Censuses and Surveys; Dissemination Practices and Openness; Availability of Key Indicators
Aggregation method	Simple arithmetic average	Revised weighted average
Operational relevance	1) Track the strengths and weaknesses of country statistical capacity overtime in a cost-effective manner; 2) Track the progress and sustainability of Bank-financed projects in statistical capacity building; 3) With a focus on MDGs in Periodicity section, the old SCIs provided a monitoring tool for country MDGs data production capacity.	1) Provide an objective, justifiable assessment of country statistical capacity over time with comprehensive, up-to-date information; 2) Provide guidance to the WB teams in assessing the progress and sustainability of the Bank supported projects; 3) Inform Systematic Country Diagnostics; 4) Provide a monitoring tool for countries' SDGs data production capacity.
Limitations/Identified weaknesses	1) Output focused, failing to address the infrastructure/resource part of the NSOs; 2) Relative narrow scope of dimensions and indicators; 3) Unable to reflect the change of data landscape and new data requirements brought up by the SDGs.	1) Selection of indicators, especially under the Availability of Key Indicators section, is constrained by the development stage of SDGs indicators. This leaves room for future development of the SCIs structure.

Figure 1: SPI Scores versus Country Income, 2016



Note: Estimated coefficients are shown from an OLS regression of SPI score on log of GDP per capit
standard errors are in parentheses.

Figure 2: SPI Scores versus Country Economic Complexity, 2016

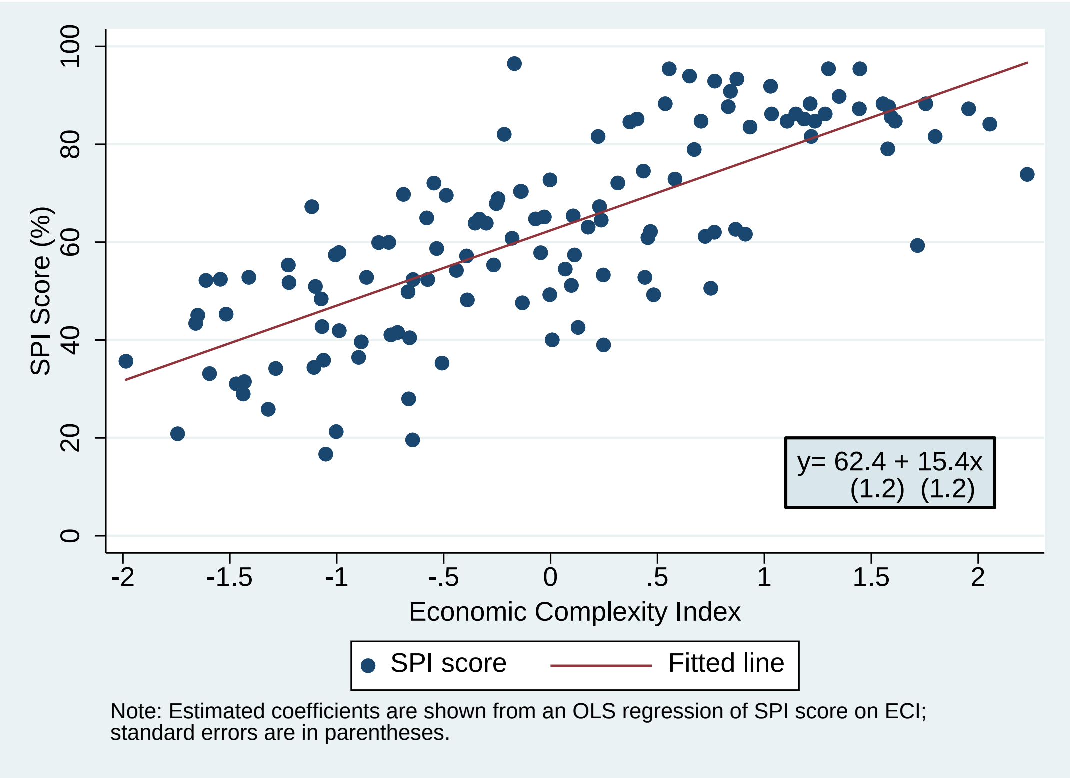


Figure 3: SPI Scores by Region, 2016

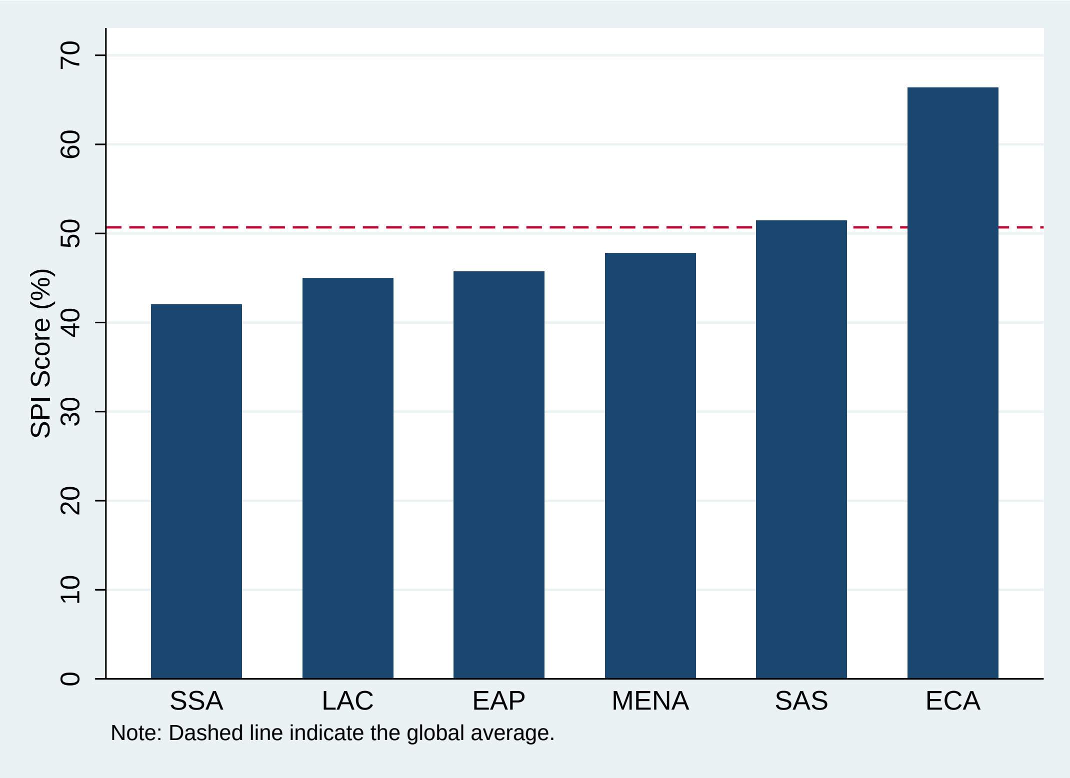


Figure 4: Dimension SPI Scores by Region, 2016

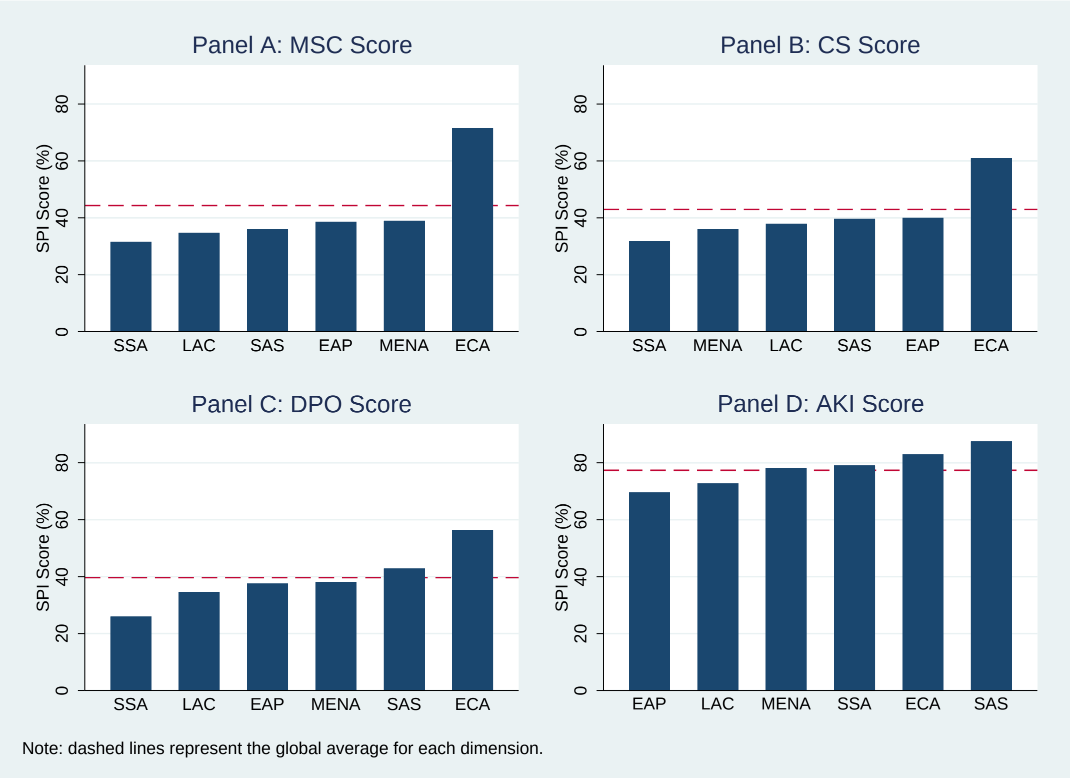
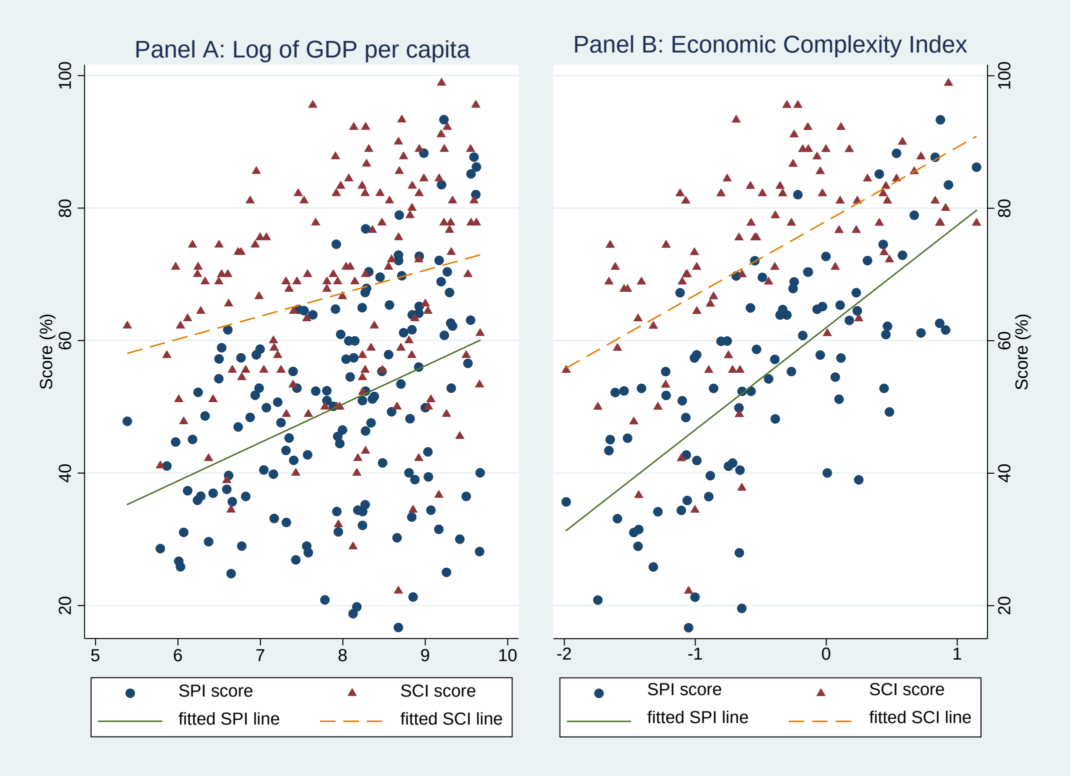


Figure 5: SPI and SCI Scores versus Country Income and Country Economic Complexity, 2016



Appendix 1: Proofs, Axiomatics, and a Numerical Example

Part A: Proofs

Proposition 1. Decomposability into Subsets of Variables

Since S and S' are two nonempty and exclusive sets of variables whose union is $\{a_1, \ldots, a_V\}$, we have

$$w \cdot a = w^S \cdot a^S + w^{S\prime} \cdot a^{S\prime} \tag{1.1}$$

Divide both sides of Equation (1.1) by $w \cdot u$, we have

$$\frac{w \cdot a}{w \cdot u} = \frac{w^S \cdot a^S}{w \cdot u} + \frac{w^{S\prime} \cdot a^{S\prime}}{w \cdot u} \tag{1.2}$$

or equivalently

$$C = \frac{w^S \cdot u^S}{w \cdot u} C^S + \frac{w^{S\prime} \cdot u^{S\prime}}{w \cdot u} C^{S\prime} \tag{1.3}$$

Indeed, for Equation (1.3), the term on the left hand side follows from Equation (1.2) using the definition of C. The two terms on the right hand side for the same equation are obtained by multiplying the term on the right hand side of Equation (1.2) respectively with $\frac{w^S \cdot u^S}{w^S \cdot u^S}$ and $\frac{w^{S\prime} \cdot u^{S\prime}}{w^{S\prime} \cdot u^{S\prime}}$ and then rewriting using the definitions of C^S as $\frac{w^S \cdot a^S}{w^S \cdot u^S}$, and $C^{S\prime}$ as $\frac{w^{S\prime} \cdot a^{S\prime}}{w^{S\prime} \cdot u^{S\prime}}$.

A similar proof applies where we partition the set $\{a_1, \ldots, a_V\}$ into D subsets, such that C can be decomposed into D associated dimensional indices $C^{(d)}$.

Proposition 2. Decomposability into Subsets of Countries

Let N be the number of countries considered and let A denote the $N \times V$ achievement matrix whose n^{th} row a^n is the achievement vector for country $n = 1, \ldots, N$. Then the index applied to the collection $\{1, \ldots, N\}$ of countries is defined by $C = C(A; w) = \mu(A; w)$. Since the weighted mean $\mu(A; w)$ of matrix A is given by an arithmetic mean of $N \times V$ elements

$$\mu(A; w) = \sum_{n=1}^{N} \sum_{v=1}^{V} \left(\frac{1}{N} \frac{w_v}{w_1 + \cdots + w_V}\right) a_v = \sum_{n=1}^{N} \frac{1}{N} \mu(a_v; w) \tag{1.4}$$

there are two ways to obtain C. The first way is to average across the dimension as with Proposition 1. The second way is to calculate the index for each country n first as $C(a^n; w)$, and then obtain the average for the whole world (or region). This can be formally stated as follows

$$C(A; w) = \frac{1}{N} \sum_{n=1}^{N} C(a^n; w) \tag{1.5}$$

Proposition 3. Equivalence between Dichotomous and Multi-Valued Indicators

Let $S = \{a_1, \ldots, a_v\}$ and $S' = \{a_{v+1}, a_{v+2}, \ldots, a_T\}$. The T dichotomous variables in S' can be replaced with one variable x having the values 0, 1/T, …, (T-1)/T, and 1, where x inherits a total weight of w_T from the original T variables. That is,

$$x = C^{S\prime} \tag{1.6}$$

and

$$w_x = \sum_{k=v+1}^{T} w_k = \frac{w^{S\prime} \cdot u^{S\prime}}{w \cdot u} \tag{1.7}$$

Let $C(a^{S'}; w^{S'})$ be the counting index for the subset of variable S'. We then have

$$C(a; w) = \frac{w^S \cdot u^S}{w \cdot u} C^S + \frac{w^{S'} \cdot u^{S'}}{w \cdot u} C^{S'} \quad (1.8)$$

or

$$C(a; w) = \frac{w^S \cdot u^S}{w \cdot u} C^S + w_x x \quad (1.9)$$

where Equation (1.8) follows from the decomposition of the two indices by sets of variables (from Proposition (1)), and Equation (1.9) follows from replacing the second term on the right-hand side of Equation (1.8) with the corresponding terms in Equations (1.6) and (1.7).

Part B: Numerical Example

Figure 1.1 graphically illustrates the nested counting method for an example having nine variables and three dimensions. The first two dimensions have two variables apiece; the third dimension has two subdimensions. The first of these subdimensions has two variables while the second subdimension has three. The example has been constructed according to the nested approach, with the width of each dimension, subdimension or variable indicating its relative importance. The implied weight for each variable – which is the width for the variable divided by the overall width – is listed for each at the base of the figure. Now suppose that the nine achievements from right to left take on the values a = (1,0,1,1,1,0,1,0,0), where each "1" variable is shaded in the figure, and each "0" is indicated by a lack of shading. The counting approach essentially adds up the associated weights, or equivalently the shaded widths divided by the total width (as with Equation (1)), resulting in

$$C = 1/6+1/6+1/6+1/12 +1/18 = 23/36 \quad (1.10)$$

or about 0.639.

Equivalently, the aggregate value for the counting measure is obtained by taking the inner product of the weighting vector $w = (w_1, \dots, w_9) = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{12}, \frac{1}{12}, \frac{1}{18}, \frac{1}{18}, \frac{1}{18})$ and the profile $a =$ (1,0,1,1,1,0,1,0,0) (as with Equation (2)), yielding

$$C = (w_1 + w_3 + w_4 + w_5 + w_7)/1 = 23/36 \quad (1.11)$$

as before. Notice that $M = 1 - C = 13/36$ is a complementary measure of the extent to which targets have yet to be met, as indicated by the share of the figure that is left unshaded.[31]

Returning to the example with profile $a =$ (1,0,1,1,1,0,1,0,0), weighting vector $w = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{12}, \frac{1}{12}, \frac{1}{18}, \frac{1}{18}, \frac{1}{18})$ and three dimensions, it follows that

$$C = \frac{1}{3} C^{(1)} + \frac{1}{3} C^{(2)} + \frac{1}{3} C^{(3)} \quad (1.12)$$

where $C^{(1)} = 1/2$, $C^{(2)} = 1$, and $C^{(3)} = 5/12$. Summing the right side of Equation (1.12) yields 23/36 as before, which illustrates Equation (5). Dividing each term in this equation by the value of C, the relative contribution of each dimension is approximately 26%, 52%, and 22%, respectively.

[31] M is analogous to the individual poverty measure underlying the adjusted headcount ratio M_0 of Alkire and Foster (2011) with a "union" identification step.

It is likewise possible to define indices for subdimensions. In the example there are two subdimensions for dimension 3, the first of which includes variables 5 and 6 and the second of which includes variables 7, 8 and 9. Where $C^{(31)}$ and $C^{(32)}$ denote the associated indices, it follows that $C^{(31)}$ = 1/2 and $C^{(32)}$ = 1/3, and that

$$C^{(3)} = \frac{1}{2}C^{(31)} + \frac{1}{2}C^{(32)} = 5/12 \qquad (1.13)$$

where the contribution of the first subdimension to the dimensional index is 60% and the contribution of the second subdimension is 40%. Overall, the decomposition formula over these basic groups of variables – namely the dimensional and sub-dimensional variables – is given by

$$C = \frac{1}{3}C^{(1)} + \frac{1}{3}C^{(2)} + \frac{1}{6}C^{(31)} + \frac{1}{6}C^{(32)} \qquad (1.14)$$

Notice that in the process of defining weights for the index, a tree structure T of dimensions and subdimensions has been spelled out for the variables. For instance, the above example has the structure $T = (a^{(1)}, a^{(2)}, (a^{(31)}, a^{(32)}))$, where $a^{(1)} = (a_1, a_2)$, $a^{(2)} = (a_3, a_4)$, $a^{(31)} = (a_5, a_6, a_7)$, and $a^{(32)} = (a_8, a_9)$ are the basic groups at the end of the tree structure. The extra parentheses around $a^{(31)}$ and $a^{(32)}$ indicate that they are both part of the same dimension. The basic groups, in turn, together form a partition P of variables into, say, K many basic groups of variables p_k for $k = 1, \dots, K$ having the property that if one variable in p_k is in a given dimension, then all variables in the p_k are in that dimension; and if one variable in p_k is in a given subdimension, all variables in p_k are in that subdimension. For instance, in the above example there are $K = 4$ basic groups and the partition is given by $P = (p_1, p_2, p_3, p_4) = (a^{(1)}, a^{(2)}, a^{(31)}, a^{(32)})$.[32]

Proposition 3 can be illustrated with the help of the example in Figure 1.1. Let $S = \{a_1, \dots, a_6\}$ and $S' = \{a_7, a_8, a_9\}$, where the latter three variables comprise subdimension 2 of dimension 3. Suppose that these three variables are related as described above, and thus yield the four options (0,0,0), (1,0,0), (1,1,0), and (1,1,1). The three dichotomous variables in S' can be replaced with one variable x having the values 0, 1/3, 2/3, and 1, where x inherits a total weight of 1/6 from the subdimension. The resulting seven variables and weights are denoted by $a' = (a_1, \dots, a_6, x)$ and $w' = \left(w_1, \dots, w_6, \frac{1}{6}\right)$. For example, for $a = (1,0,1,1,1,0,1,0,0)$, the original counting index is

$$C(a; w) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{12} + \frac{1}{18} = 23/36 \qquad (1.15)$$

while for the analogous $a' = (1,0,1,1,1,0,\frac{1}{3})$, the alternative index is

$$C(a'; w') = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{12} + \frac{1}{6} \cdot \frac{1}{3} = 23/36 \qquad (1.16)$$

Consequently, the user can choose either approach without altering the results.

Finally, Equation (9) can be illustrated with the help of the example in Figure 1.1 where the initial six variables in a' have values (1,0,1,1,1,0). How does $C_\alpha(a'; w')$ change as the final variable rises from 0 to 1/3 to 2/3 to 1? Note that the index takes on the value 0.58 when the final variable is 0 and the value 0.75 when the final variable is 1, regardless of α. What is different for different α is how the two intermediate levels line up. For the original index with α = 1, each increment in the variable leads to the same increase in the index value along the way from 0.58 to

[32] The main decomposition axiom is defined with respect to the basic groups, analogous to equation (5).

0.75. For α = 1/2 the index values are 0.58, 0.68, 0.72, and 0.75 as the final variable increases. Countries that move from 0 to the first rung of the variable generate the largest increase in $C_{1/2}$. Likewise, for α = 2 the respective index values are 0.58, 0.60, 0.66, and 0.75. It is now the movement from the penultimate to the highest rung that generates the greatest improvement in index C_2. This suggests that the class C_α of indices might be helpful as a diagnostic tool in support of the main index C.

Figure 1.1: Illustration for the Dimensions and Weights

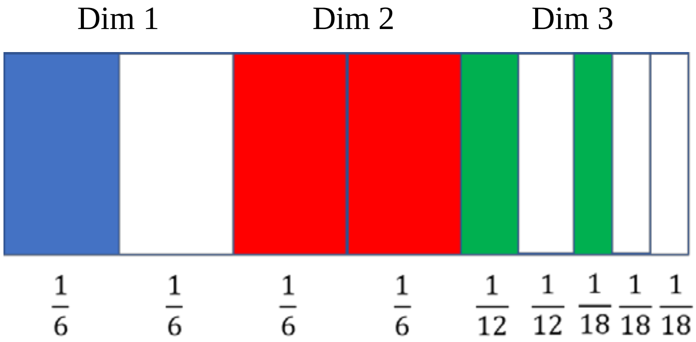


Figure 1.2: SPI Scores versus Country Income Level, 2016

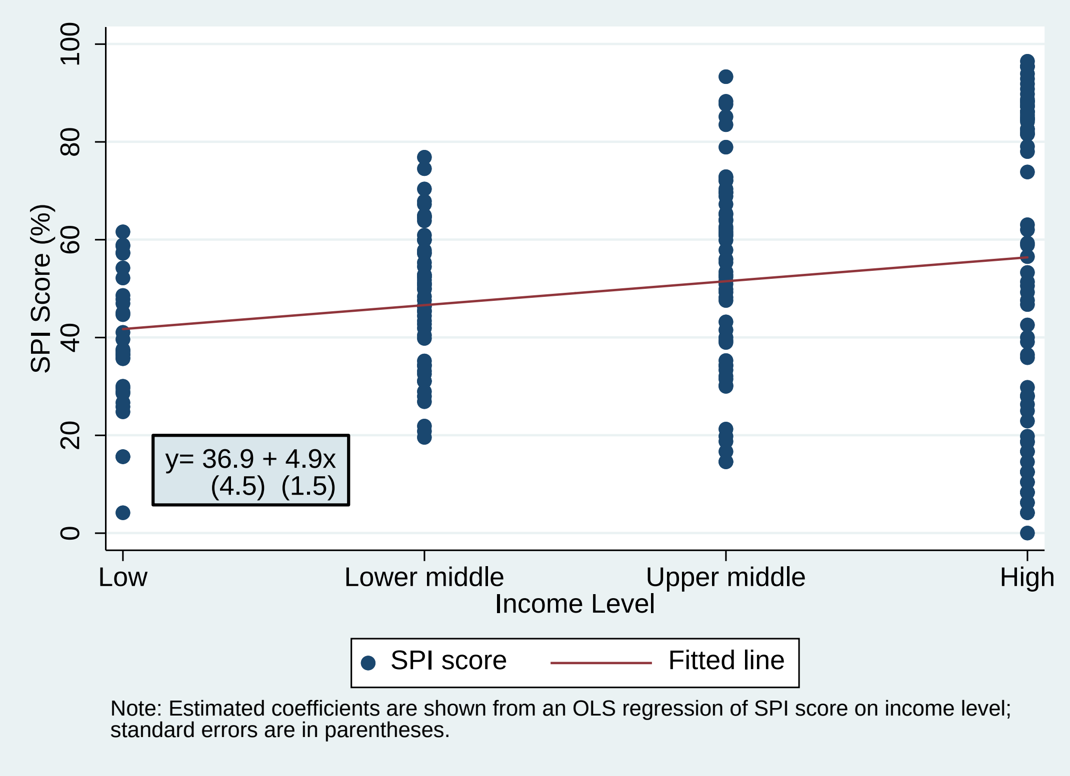
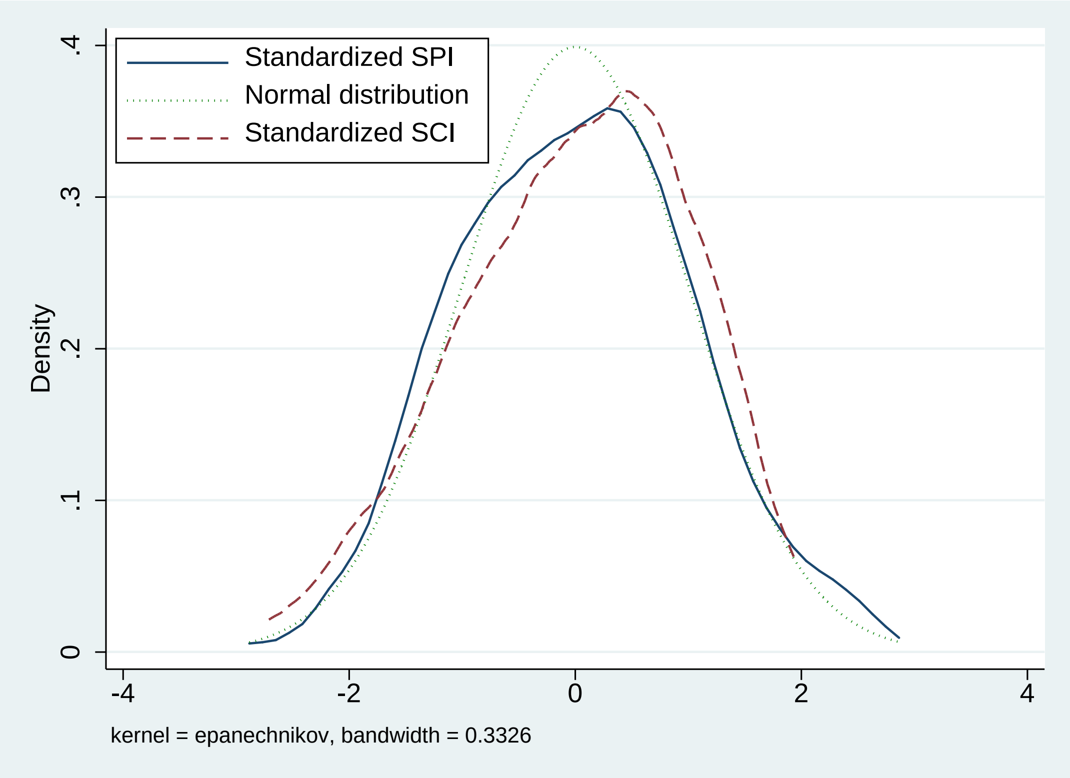


Figure 1.3: The Standardized Distribution of the SPI vs. the Standardized Distribution of the SCIs for the Same Countries, 2016



Appendix 2: SPI Scoring Methodology Matrix

Methodology, Standards & Classifications							
#	**Indicator**	**Score 1**	**Score 0.5**	**Score 0**	**Weight**	**Sources**	**Link**
1	System of National Accounts in use	SNA2008/ESA 2010	SNA1993/ ESA 1995	Otherwise	1	World Development Indicators Primary Data Documentation (WDI/PDD); International Monetary Fund's (IMF) World Economic Outlook (WEO), Table G of Statistical Appendix	http://data.worldbank.org/products/wdi http://www.imf.org/en/publications/weo
2	National Accounts base year	Annual chain linking	Within past 10 years	Otherwise	1	World Development Indicators Primary Data Documentation (WDI/PDD); International Monetary Fund's (IMF) World Economic Outlook (WEO), Table G of Statistical Appendix	http://data.worldbank.org/products/wdi http://www.imf.org/en/publications/weo
3	Classification of national industry	Latest version is adopted (ISIC Rev 4, NACE Rev 2 or a compatible classification)	Previous version is used (ISIC Rev 3, NACE Rev 1 or a compatible classification)	Otherwise	1	United Nations Monthly Bulletin of Statistics Database, Index of Industrial Production	https://unstats.un.org/UNSD/mbs/app/DataSearchTable.aspx
4	CPI base year	Annual chain linking	Within past 10 years	Otherwise	1	IMF: International Financial Statistics (IFS) / Country notes	http://www.elibrary.imf.org/browse?freeFilter=false&pageSize=10&sort=datedescending&t_7=urn%3ASeries%2F041
5	Classification of household consumption	Follow Classification of Individual Consumption by Purpose (COICOP)	N.A.	Otherwise	1	IMF: Dissemination Standards Bulletin Board (DSBB)	http://dsbb.imf.org/Default.aspx

6	Classification of status of employment	Follow International Labour Organization, International Classification of Status in Employment (ICSE-93)	N.A.	Otherwise	1	IMF: Dissemination Standards Bulletin Board (DSBB)	http://dsbb.imf.org/Default.aspx
7	Central government accounting status	Consolidated central government accounting follows noncash recording basis	Consolidated central government accounting follows cash recording basis	Otherwise	1	IMF: Government Finance Statistics (GFS) Yearbook/Guide to country tables	http://www.elibrary.imf.org/browse?freeFilter=false&pageSize=10&sort=datedescending&t_7=urn%3ASeries%2F043
8	Compilation of government finance statistics	Follow the latest Government Finance Statistical Manual (2014)/ ESA2010 or GFSM 2001	N.A.	Otherwise	1	IMF: WEO, Table G of Statistical Appendix	http://www.imf.org/en/publications/weo
9	Compilation of monetary and financial statistics	Follow the latest Monetary and Finance Statistics Manual (2000) or Monetary and Finance Statistics: Compilation Guide (2008)	N.A.	Otherwise	1	IMF: IFS / Country notes	http://www.elibrary.imf.org/browse?freeFilter=false&pageSize=10&sort=datedescending&t_7=urn%3ASeries%2F041
10	SDDS/e-GDDS subscription	Subscribing to IMF SDDS standards	Subscribing to IMF e-GDDS standards	Otherwise	1	IMF: SDDS/e-GDDS website	http://dsbb.imf.org/Pages/GDDS/Home.aspx

11	CRVS	Vital registration complete	N.A	Otherwise	1	WDI book: Primary Data Documentation	http://data.worldbank.org/products/wdi
12	Business process	GSBPM is in use	N.A	Otherwise	1	United Nations Industrial Development Organization (UNIDO) United Nations Economic Commission for Europe (UNECE)	https://statswiki.unece.org/display/GSBPM/United+Nations+Industrial+Development+Organization+(UNIDO):+use+of+GSBPM https://statswiki.unece.org/display/GSBPM/Case+Studies+of+Metadata+use+with+GSBPM+and+GSIM

Censuses and Surveys							
Censuses							
#	**Indicator**	**Score 1**	**Score 0.5**	**Score 0**	**Weight**	**Sources**	**Link**
1	Population & Housing census	Population census done within last 10 years	Population census done within last 20 years	Otherwise	1	United Nations Statistical Division (UNSD): 2020 World Population and Housing Census Programme	https://unstats.un.org/unsd/demographic/sources/census/censusdates.htm#ASIA
2	Agriculture census	Agriculture census done within last 10 years; OECD and/or EU member	Agriculture census done within last 20 years	Otherwise	1	Food and Agriculture Organization (FAO): World Programme for the Census of Agriculture	http://www.fao.org/world-census-agriculture/countries/en/
3	Business/establishment census	Business/establishment census done within last 10 years; OECD and/or EU member	Business/establishment census done within last 20 years	Otherwise	1	NSO websites	https://unstats.un.org/home/nso_sites/
Surveys							

#	**Indicator**	**Score 1**	**Score 0.6**	**Score 0.3**	**Score 0**	**Weight**	**Sources**	**Link**

4	Household Survey on income/consumption /expenditure/budget/ Integrated Survey	3 or more household surveys done within past 10 years; OECD and/or EU member	2 household surveys done within past 10 years;	1 household survey done within past 10 years;	None within past 10 years	1	Microdata library / World Bank, Development Research Group PovcalNet International Household Survey Network Catalog (IHSN) IHSN Gender Data Navigator	http://microdata.worldbank.org/index.php/home http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx http://catalog.ihsn.org/index.php/catalog http://datanavigator.ihsn.org/
5	Agriculture survey	3 or more agriculture surveys done within past 10 years; OECD and/or EU member	2 agriculture surveys done within past 10 years;	1 agriculture survey done within past 10 years;	None within past 10 years	1	Microdata library / World Bank, Development Research Group PovcalNet International Household Survey Network Catalog (IHSN)	http://microdata.worldbank.org/index.php/home http://catalog.ihsn.org/index.php/catalog http://datanavigator.ihsn.org/ http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx

							IHSN Gender Data Navigator	
6	Labor Force Survey	3 or more labor force surveys done within past 10 years; OECD and/or EU member	2 labor force surveys done within past 10 years;	1 labor force survey done within past 10 years;	None within past 10 years	1	International Labour Organisation (ILO) Microdata Repository	http://www.ilo.org/surveydata/index.php/catalog/central#_r=&collection=&country=94,140&dtype=&from=2006&page=1&ps=&sid=&sk=&sort_by=nation&sort_order=&to=2016&topic=&view=s&vk= http://www.ilo.org/dyn/lfsurvey/lfsurvey.list?p_lang=en http://microdata.worldbank.org/index.php/home http://catalog.ihsn.org/index.php/catalog http://datanavigator.ihsn.org/ http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx
7	Health/Demographic survey	3 or more health surveys done within past 10 years; OECD and/or EU member	2 health surveys done within past 10 years;	1 health survey done within past 10 years;	None within past 10 years	1	Demographic and Health Surveys	https://dhsprogram.com/What-We-Do/survey-search.cfm?pgType=main&SrvyTp=type
8	Business/establishment survey	3 or more business/establishment surveys done within	2 business/establishment surveys	1 business/establishme	None within	1	Microdata library / World Bank, Development	http://microdata.worldbank.org/index.php/home

		past 10 years; OECD and/or EU member	done within past 10 years;	nt survey done within past 10 years;	past 10 years		Research Group PovcalNet International Household Survey Network Catalog (IHSN) IHSN Gender Data Navigator	http://catalog.ihsn.org/index.php/catalog http://datanavigator.ihsn.org/ http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx

Availability of Key Indicators						
#	**Indicator**	**Score 1 - Data is available once in three years, prior to reference year**	**Score 0**	**Weight**	**Source**	**Link**
1	Poverty headcount ratio at national poverty lines/ availability of similar key indicators	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
2	Mortality rate, under-5 (per 1,000 live births)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators

3	Child immunization (proportion of one-year-old children immunized against measles)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
4	Primary completion rate, both sexes (%)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
5	Adult literacy rate, population 15+ years, both sexes (%)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
6	Proportion of population using safely managed drinking water services	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
7	Unemployment, total (% of total labor force)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
8	Manufacturing value added as a proportion of GDP	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
9	Gross capital formation (% of GDP)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
10	GDP implicit price deflator (annual % growth)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators

11	Net trade in goods and services (BoP, current US$)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators
12	Prevalence of undernourishment (% of population)	Yes	No	1	WDI database	http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators

Dissemination Practices & Openness						
1) Dissemination capacity of NSO						
#	**Indicator**	**Score 1**	**Score 0**	**Weight**	**Source**	**Link**
1	NSO has an Advance Release Calendar and it is published	Yes	No	1	UNSD Partners National Statistics Office (NSO) websites	https://unstats.un.org/home/nso_sites/
2	NSO has a listing of surveys and microdata sets (or NADA)	Yes	No	1	NSO websites / Accelerated Data Program of IHSN	https://unstats.un.org/home/nso_sites/ http://www.adp.ihsn.org/country_activities
3	NSO has a data portal	Yes	No	1	NSO websites	https://unstats.un.org/home/nso_sites/
4	Timeseries indicators are available for download in reusable format for free	Yes	No	1	NSO websites	https://unstats.un.org/home/nso_sites/

5	Metadata is available providing definition, methodology, standards or classifications for existing data series	Yes	No	1	NSO websites	https://unstats.un.org/home/nso_sites/
6	NSO has conducted a user satisfaction survey	Yes	No	1	NSO websites	https://unstats.un.org/home/nso_sites/
7	Geospatial data available on relevant agency website	Yes	No	1	NSO / Relevant agency websites	https://unstats.un.org/home/nso_sites/
2) Openness of data						
#	**Indicator**	**Score**		**Weight**	**Source**	**Link**
8	Open Data Inventory	ODIN score		0	Open Data Watch	http://odin.opendatawatch.com/
9	Open Data Barometer	ODB score		0	The Open Data Barometer	http://opendatabarometer.org/
10	Open Data Index	ODI score		0	Global Open Data Index	https://index.okfn.org/
Maximum score for sub-category "Openness of data": 0						