CrossMark

The Inclusive Sustainable Transformation Index

Justin Yifu Lin[1] · Célestin Monga[2] · Samuel Standaert[3]

Abstract

In this paper, we put forth an index of Inclusive Sustainable Transformation that captures the extent to which a country has developed a modern industry or services-based economy that at the same time protects the environment and is gender inclusive. This index distinguishes itself from other indicators that track the structural characteristics of the economy by ensuring that the comparisons between countries account for differences in the level of development, in line with New Structural Economics thinking. The index evaluates how well a country scores given its available resources. In addition, by addressing data availability problems using multiple imputation techniques, the index is able to compare performances on a wide range of topics for almost 200 countries over 25 years, including a large group of developing countries that are often left out. In addition to monitoring the progress made towards the establishment of an inclusive and environmentally friendly, modern economy, the index is a useful tool for policy makers and analysts. By decomposing the total score back into its components, it can help identify areas that require additional attention, as well as 'best practices' in countries at similar levels of development.

1 Introduction

With the adoption of the Sustainable Development Goals (SDGs) and the Paris Agreement on climate change, the year 2015 was a major inflection point in the long struggle for a global consensus on these international priorities. While both agreements set out clear policy goals, the question of how to monitor and evaluate their progress largely remained unanswered. In recognition of this, in December of 2014, UN Secretary-General Ban Ki-moon called for a 'comprehensive program of action on data.' However, the Open Working Group who proposed the 17 SDGs recommended tracking as many as 169 individual targets. It is therefore highly improbable that funding is or will be available to carry out

✉ Samuel Standaert
sastand@clemson.edu

1   Peking University, Beijing, China

2   African Development Bank, Abidjan, Côte d'Ivoire

3   Clemson University, Clemson, USA

the necessary data collection in all countries (Jerven 2014).[1] Under such circumstances, it is better to adopt a more focused approach, which is still based on objective data analysis and monitoring but covers only the factors considered to be the most critical indicators of the SDGs. To that end, this paper is centered on three elements that underlie as much as half of the SDGs: (1) environmentally friendly and (2) socially inclusive (3) structural transformation.

Structural change is the foundation of sustainable and inclusive growth and the condition for achieving the SDGs (Monga 2013). Rarely has a country evolved from a low- to a high-income status without continuous structural transformation from an agrarian or resource-based economy towards an industry- or services-based economy. Industrialization is essential for lifting people out of poverty, creating jobs, advancing technology, and generating prosperity around the world. However, industrial development often comes at a high environmental cost. For example, industrial production is currently the largest emitter of greenhouse gasses representing almost 30% of global emissions. Fortunately, it is possible to transform conventional industrial development patterns to prevent dangerous anthropogenic interference with the atmosphere and other environmental destruction. It is equally important to ensure that these economic opportunities are open to all, regardless of gender or other individual characteristics. After all, for economic growth to be truly impactful, it cannot be limited to (less than) half the population.

The global community needs monitoring tools that provide the right incentives to governments, the private sector, and other development stakeholders to actively promote this kind of structural transformation. To that end, this paper outlines the construction of the Inclusive Sustainable Transformation (IST) index, which captures the extent to which a country has developed a modern economy that protects the environment and is gender inclusive.

In addition to its more directed focus, the IST index sets itself apart from other indicators of sustainable development in three ways. First and foremost, it accounts for a country's development status when assigning scores. New Structural Economics tells us that the feasible and desired characteristics of countries change with their level of development (Lin 2012a, b). This idea is explicitly incorporated in the IST index, which assesses a country's progress towards sustainable development relative to that of countries with a similar level of development. Moreover, the conditional nature of this index aligns with the idea that sustainable development is a continuous process of improvement for all countries, rather than a fixed path with clearly defined end goals.

Second, instead of using the latest available data for each indicator like the SDG indexes of Kroll (2015) and Sachs et al. (2016, 2017), the IST index keeps track of the availability of the underlying indicators. To deal with the gaps in their coverage, we introduce a new way of imputing the missing data: Multiple Imputation using State Space models (MISS). The MISS algorithm substantially increases the reliability of the imputed values and ensures that the confidence intervals of the index reflect the extent to which imputed data was used. This allows us to compute the IST index for almost 200 countries from 1990 to 2016.

Finally, the paper uses a (conditional) cumulative density function (CDF) to rescale the indicators and compare the performance of different countries. The CDF tell us

---

[1] A preliminary estimate by Morten Jerven suggests that even minimal data collection for all 169 targets would cost at least $254 billion—that is about twice last year's ODA aid flows—and this does not even include the cost of conducting all household surveys (Jerven 2014).

what the probability is of finding a country (with a similar level of development) that performs worse; the higher this probability, the more progress a country has made relative to its peers. By transforming the indicators in this way, we avoid the discontinuity problems associated with using discrete methods like rankings or thresholds (cf. Kroll's 2015) while retaining a straightforward interpretation. Moreover, it allows us to take the level of development into account without having to impose a fixed grouping of countries based on their level of development. In short, it enables a straightforward comparison of the IST scores over time and between countries, including those with a completely different developmental status.

The remainder of this paper is organized as follows. The next section defines structural transformation, its importance for growth and how industrialization contributes to it. Section 3 surveys the theoretical challenges for building development indexes and presents the IST methodology. The Sects. 4 and 5 discuss which indicators were chosen and give an overview of the resulting IST index and its subcomponents.

2 The Need for Structural Transformation

The importance of structural transformation as a process for generating prosperity and for improving the quality of life around the world cannot be overstated. This process typically involves improving the productivity in the agricultural sector to increase food supply, free up labor and provide savings. These resources can then support the process of industrialization, urbanization and the development of a high-performing service sector that can absorb a growing fraction of the educated labor force. Prolonged growth and economic prosperity require a shift of resources out of traditional agriculture and other low-productivity primary activities into more productive sectors of manufacturing and services in both urban and rural areas. The ensuing expansion and upgrading of 'modern' sectors (including non-traditional agriculture) are at the core of the sustained productivity gains that characterize economic development. Indeed, the consensus among economists is that rising productivity accounts for the bulk of long-term growth (Lin and Monga 2014).

Structural transformation (or structural change) is, therefore, the central focus of economic policy for countries at all levels of development. It has five main features: (*i*) a steadily declining share of agriculture in economic output and employment; (*ii*) a rising share of urban economic activity in industry and modern services; (*iii*) an increasingly sophisticated share of manufactured goods in production and exports; (*iv*) migration of rural workers to urban settings; (*v*) and a demographic transition that typically involves a spurt in population growth before reaching a new equilibrium.

Sustaining high economic performance, improving living standards, and sharing prosperity widely to maintain social cohesiveness and peace require constant movement of resources to new, more productive industries, sectors, and firms, as well as continuous infrastructural and institutional improvement. Throughout this process, the country's structure of factor endowments—the relative composition of natural resources, labor, human capital and physical capital—will be innately different at every level of development. Because of this, the optimal industrial structure and comparative advantage of any given economy will evolve along with its level of development (Lin 2012a, b; Lin and Monga 2013).

2.1 Industrialization as a Source of Growth

The modernization of agriculture and sustainable industrialization are essential features of the structural transformation process. Productivity increases in agriculture provide food, labor, and savings that fuel the process of urbanization and industrialization (Timmer and Akkus 2008). The development of a competitive industrial sector yields an even higher payoff. Economists have established at least since the early 1960s that manufacturing has always played a significant role in the total output of more affluent countries, and that countries with higher incomes are typically those with substantially larger transport and machinery sectors (McMillan and Rodrik 2011). In fact, only in singular circumstances such as an extraordinary abundance of land or resources have countries succeeded in developing without industrializing. Industrialization also promotes inclusive development by expanding the fiscal space for social investments.

Within the industrial sector, manufacturing in particular has transformed the dynamics of the world economy. The globalization of manufacturing is driven by many factors, including profound changes in geopolitical relations among world nations, the widespread growth of digital information, the decline of transportation costs, the development of physical and financial infrastructure, computerized manufacturing technologies, and the proliferation of bilateral and multilateral trade agreements. These developments have permitted the decentralization of supply chains into independent but coherent global networks that allow transnational firms to locate different parts of their businesses around the world. The creative design of products, the sourcing of materials and components, and the manufacturing of products can now be done more cheaply and more efficiently from virtually any region of the planet while final goods and services are customized and packaged to satisfy the needs of customers in faraway markets. The globalization of manufacturing has thus allowed developed economies to benefit from lower wages in developing countries such as China, India, Bangladesh, Costa Rica, Mexico, and Brazil while creating job and learning opportunities in these formerly poor nations.

2.2 Does Manufacturing Still Matter?

In recent decades, innovation, technological developments and new sources of economic growth have led some economists to question whether manufacturing still matters. Manufacturing's share of global value added has steadily declined over the past 30 years as the global value added of services has grown.[2] However, these trends are mainly observed in high-income countries and can be explained by several factors. First, productivity increases and rising standards of living in advanced economies have pushed up wages and forced many industries to delocalize their production to lower-cost nations. Second, increasing levels of efficiency in the world economy have reduced the relative prices of consumer goods while at the same time the demand for services such as healthcare, security, and transportation has increased. Finally, and perhaps even more importantly, manufacturing jobs have a multiplier effect on employment in services, as the development of industries

---

[2] The share of manufacturing in GDP fell from 35% in 1985 to 27%, while the share of services rose from 59 to 70% (UNIDO 2009).

everywhere automatically generates a wide variety of economic activities, from transportation to housing, from hospitality to entertainment.[3]

Concerns about the future of manufacturing as a viable source of economic growth have been investigated empirically by Hausmann et al. (2011). They found over 70% of the income variations among nations can be explained by differences in manufactured product export data alone. The analysis of the composition and scale of a nation's manufacturing sector revealed that sophisticated economies export a large variety of 'exclusive' goods that few other countries can produce. These economies have typically accumulated productive knowledge and developed manufacturing capabilities that others do not have. It therefore appears that national income and economic sophistication (economic complexity) rise in tandem.

Even basic manufacturing expertise can gradually generate new knowledge and lead to new, more advanced products, provided that the right strategic and business decisions are made on industrial and technological upgrading. In the words of Hausmann and Hidalgo (2012, p. 13), economic development is 'a social learning process, but one that is rife with pitfalls and dangers. Countries accumulate productive knowledge by developing the capacity to make a larger variety of products of increasing complexity. This process involves trial and error. It is a risky journey in search of the possible. Entrepreneurs, investors, and policy-makers play a fundamental role in this economic exploration. Manufacturing, however, provides a ladder in which the rungs are more conveniently placed, making progress potentially easier.' In sum, manufacturing still generates economies of scale, sparks industrial and technological upgrading, fosters innovation, and has significant multiplier effects.

3 Composing an Index of Inclusive and Sustainable Transformation

Both the Sustainable Development Goals (SDGs) and the Paris Climate agreement require that all signatories continuously assess and report on the progress made toward their objectives. However, the monitoring of this progress is a major challenge, especially given the differences in the level of development and the production structure of the countries that joined the agreements. As Ahluwalia (2015, p. 5) notes the best one can expect under these circumstances is for economists to 'help to define a set of measurable indicators reflecting various aspects of inclusiveness and sustainability, taking into account availability of data on these indicators, and the scope for improving data availability over time. We could then set targets for each of these indicators and hope that they would be accepted by different stakeholders as representing significant improvement in each dimension.'

In the case of the SDGs, each goal requires a multidimensional policy framework for action. As a result, to keep track of the progress on all 17 goals, 169 target indicators were identified. With this many aspects to consider, assessing overall progress has become even more challenging, as we can expect numerous conflicting narratives on whether progress has been made. In short, there is a need for synthetic indicators that can capture the essence of empirical analyses and convey policy-relevant messages to development stakeholders who would otherwise be overwhelmed trying to make sense of the data generated about each indicator.

---

[3] A study by the U.S. Department of Commerce, Bureau of Economic Analysis, shows that manufacturing has a higher multiplier effect on the American economy than any other sector, with $1.40 in additional value for every $1.00 in manufacturing value added (WEF 2014).

3.1 Measuring with Indexes: Beyond the Utopian Quest for Legitimate Indicators

There is no shortage of composite indexes to track economic development over time and across countries. In fact, there are so many of them that it has become almost impossible for policymakers to make sense of the stories they tell and to identify the specific, actionable policy levers that yield clear economic and social gains. In his critical review of some popular composite indexes of development, Ravallion (2011, p. 2–3) categorized them into two broad types. First, indexes such as the gross domestic product (GDP), for which the choice of the component series and the aggregate function 'are informed and constrained by a body of theory and practice from the literature.' Second, indexes such as the Human Development Index that are based on 'a set of indicators that are assumed to reflect various dimensions of some unobserved (theoretical) concept.' He saw the former as more appropriate indexes, while the latter lacked the necessary analytical legitimacy: 'neither the menu of the primary series nor the aggregation function is pre-determined from theory and practice, but are […] key decision variables that the analyst is free to choose, largely unconstrained by economic or other theories intended to inform measurement practice' (p. 3). To illustrate his point, Ravallion (2011) contrasts an index where the variables and weights are based on a regression model calibrated with survey data with an index where they are set by an analyst who has some concept of economic welfare in mind. Ravallion (2011) refers to the latter as a 'mashup' index.

Such a distinction may seem like an elegant conceptualization of the problem at hand, but it is an artificial one. First and foremost, the contention that GDP should be *the* model index, one legitimized by 'pure' theoretical reasoning and rigorous analytical modeling, is invalidated by the body of academic research that has highlighted its many shortcomings (beyond being a randomly aggregated set of variables that form a series of accounting identities). GDP as an index is not beyond suspicion. Calling it a 'capitalist conspiracy' like Venezuelan President Hugo Chavez did, may have been an extreme form of criticism. Yet, as a model index, GDP carries many shortcomings and paradoxes (see Coyle 2014). While GDP measures product, it ignores central facts such as quality, costs, sustainability, or purpose (Stiglitz et al. 2010).[4]

Moreover, the expectation that economists and other social scientists can construct development indicators that pass the test of 'pure theories' simply because such indicators would be 'based on a regression model calibrated to survey data' is unrealistic. It is well known that regression models are based on a host of assumptions; without them, legitimate inferences cannot be drawn from the model. While there are statistical procedures for testing some of these assumptions, the tests often cannot detect substantial failures. As pointed out by Freedman (2010, p. 14), 'model testing may become circular; breakdowns in assumptions are detected, and the model is redefined to accommodate. In short, hiding the problems can become a major goal of model building.'

---

[4] GDP aims to measure the goods and services produced, but it is often unclear what is behind the definitions of the totality of goods and services produced. 'Hidden below the overall intention, however, lurk choices and decisions invisible in our day-to-day lives. What are goods and services? How are they defined? Whose contributions count? What, in the end, are we growing? Definitions are based on a still-evolving, cumbersome system of criteria. What is counted as investment or income or expenditure, and how to define the difference between 'final' and 'intermediate' consumption follows a logic that often eludes even accountants' (Philipsen 2015, p. 12).

However appealing it may be at face value, the dichotomy between 'credible' indexes based on some theory or calibrated from regression analyses, and 'mashup' ones singled out as 'randomly elaborated,' is problematic because what constitutes an acceptable theoretical basis is always debatable. Such a distinction assumes the existence of a rationally-neutral analyst who can observe and monitor performance with distance, detachment, and balance. Philosophers have long provided good arguments about the impossibility of this type of rational actor. From Darwin to Marx, Nietzsche, Freud, Wittgenstein or Heidegger, there is an accumulated body of evidence that the so-called sovereign rational subject—the detached observer imagined by Kant—actually does not exist.[5] It follows that no development economist is intellectually autonomous, self-transparent and capable of identifying causal relationships and causal mechanisms in a detached manner. It is impossible to deny the role of the pre-conceptual and non-conceptual at the very core of the rational. The notion that subject and object can be separated from one another—the supreme dogma of empiricism—is merely an illusion.

Therefore, any economic theory or model, especially one built on regression analyses, should acknowledge the limits of its generated knowledge. Any index out there reflects an explicit or implicit theoretical analysis of the dynamics of economic development. The real criteria for assessing pertinence and effectiveness should be whether an index provides useful information to strengthen intellectual and policy arguments and whether it helps focus the attention on social and economic goals deemed to be of importance to society.

3.2 Rescaling, Weighing, and Aggregation

The first step in the construction of the IST index is the selection of the indicators of inclusive and sustainable development, which will be outlined in detail in the next section. The remainder of this section discusses the subsequent steps, namely the way in which these indicators are rescaled, weighed and aggregated, as well as how the level of development is taken into account and how missing observations are handled.

Before the indicators can be combined, they first have to be put them on a comparable footing. There are different ways of transforming these variables, many of which are described in the OECD handbook on composite indicators (OECD-JRC 2008). The distribution of values of the indicators in the IST dataset can differ markedly between indicators and different levels of development. This argues against using more common transformations like z-scores or Min–Max as done for example in the Quality of Growth index (Mlachila et al. 2014). Instead, we use the (empirical) cumulative density function (CDF), which expresses the probability of finding a country with a lower score.[6] The main advantage of the CDF is that it provides a clear interpretation: a CDF of zero means that the country performs worse than all other countries with the same level of development, while a CDF of one indicates the opposite. Any score in between can be interpreted as the fraction of countries that scores worse. Moreover, as we will see below, the CDF also enables us to account for the level of development in a straightforward fashion, without discontinuity problems.

[5] See Baynes and Bohman (1987) for an overview and useful discussion of debates about reason and lessons from post-philosophy.

[6] The CDF is a non-parametric transformation, meaning that it does not impose any assumptions on the distribution of the data. At the same time, it can capture whatever pattern is present, whether it is linear, inverted-U, multi-modal or otherwise. Moreover, it can do this regardless of whether the characteristic in question is a binary, discrete or continuous variable (Henderson and Parmeter 2015).

The estimation of the CDF is done year-by-year, using only the information on the distribution of the indicator available at that point in time. This allows the frame of reference to change over time, as the older patterns in the distribution of the indicators left out. As a result, the transformed scores represent a country's position relative to its peers in that year. The CDF will not change if all countries improve (or deteriorate) at the same rate. Similarly, a country's score will decrease if it remains unchanged while its peers improve. In this way, the index underlines the idea that inclusive and sustainable industrial development is a continuous process of improvement for all countries, rather than a fixed path with a clearly defined end goal. An additional benefit of this relative approach is that it undermines what Ravallion (2011) termed rank-seeking behavior: increasing your score until it just exceeds a benchmark. Unless a government is willing to decrease its level of development deliberately, each aspect of the index will need continuous improvement for the country to keep up with its peers.

After transformation, the indicators are combined into the IST index using a simple average. This means that each component receives equal weight in the final index. Moreover, it imposes perfect substitutability between all goals: a high score on manufacturing compensates perfectly for a lower environmental score. However, in the robustness section of this paper (Sect. 5.1), we abandon the assumption of perfect substitutability and instead use a geometric average, which penalizes countries for imbalances in their scores.[7] However, as will be shown in the robustness section, the effect of using a geometric average on the index values is small.

As the weights sum up to one, the IST index has the same range as the transformed indicators: an IST score of one means that the country outperforms all of its peers, the opposite holds true for zero. Anything in between can be interpreted as the average fraction of countries that perform worse.

3.3 Differentiating by Level of Development

A central tenet of New Structural Economics is that the structural characteristics of countries are not a one-size-fits-all (Lin 2012a). Instead, the economic structure that best helps growth will change as the country develops and new characteristics become feasible. Simply put, one cannot expect relatively developing countries such as India, Burundi or Ethiopia to have the same environmental, institutional and economic characteristics as rich countries like Denmark or Japan. Rather than relying on the individual to consider this when using the IST index, we want to embed this thinking directly into the index. It should indicate the performance of a country relative to countries with a similar level of development.

By way of illustration, Panel *a* of Fig. 1 plots the investment in Research and Development in percent of GDP (*R&D*) versus the level of development (*dev*) as measured by the log of the Gross National Income (GNI) per capita in 2012. It shows that the distribution of investment in *R&D* depends strongly on the level of development. The higher *dev*, the more the distribution shifts towards the maximum values of *R&D*.

An intuitive way of taking development into account is to group countries by their development level and only compare their characteristics within these groups. In line with NSE

---

[7] For example, while the couples {0.5, 0.5} and {0.7, 0.3} have the same arithmetic mean, their geometric means are 0.5 and 0.46.
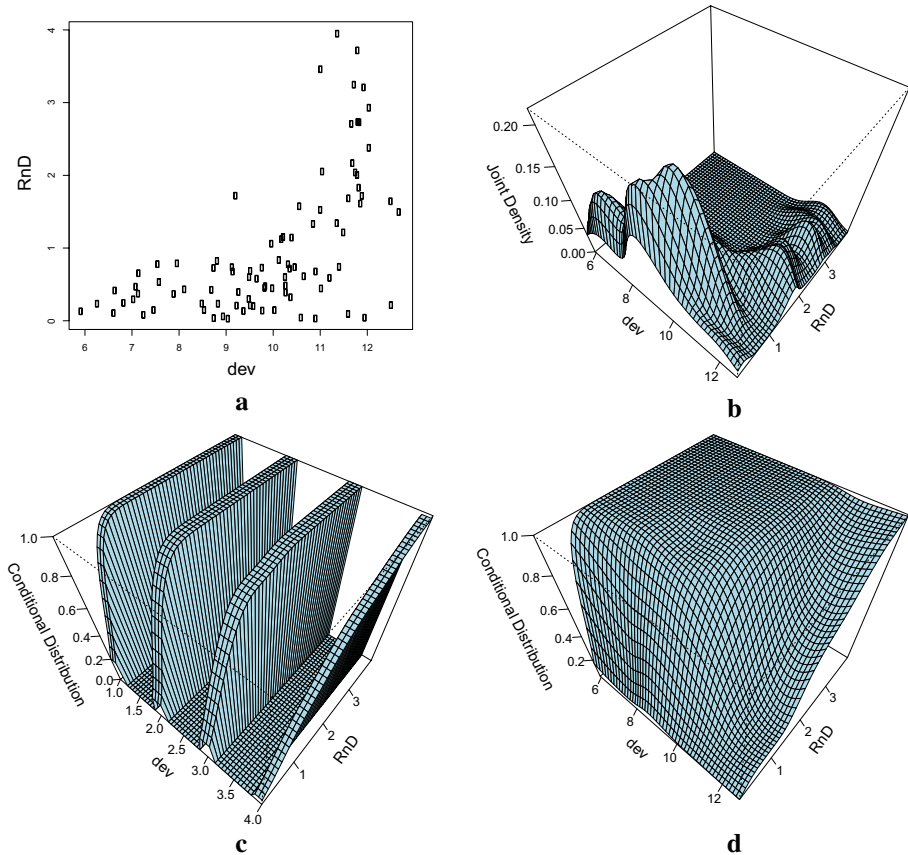
Fig. 1 Conditional kernel density estimator. Panel *a* shows the scatterplot of the investment in R&D as a percentage of GDP *ver*sus the log of GNI per capita (dev) in 2012; Panel *b* the estimated joint probability function of R&D and *dev*; Panel *c* the CDF of R&D computed for each level of development from low income (1) to high-income countries (4); and Panel *d* the conditional cumulative density function of *R&D* given *dev*. **a** Scatterplot of R&D and *dev*, **b** joint density of *R&D* and *dev*, **c** CDF of R&D for each level of development and **d** Conditional CDF of R&D|*dev*

theory, PPP converted per capita income can be used as an indicator of the level of development and the capacity of an economy.[8] For example, the income classification employed by the World Bank identifies four different income levels: low-income, lower-middle and upper-middle income and high-income countries.[9] Panel *b* shows the joint distribution of *y* and *dev* and Panel *c* illustrates how this translates into the conditional CDF for the four income-levels. In general, countries with a high investment in R&D will get a higher score.

---

[8] Alternative measures of development, like the Human Development Index (HDI), are discussed in the robustness section.

[9] As of 1 July 2015, these groups are defined using the Gross National Income per capita in the following way: LIC ≤ 1045; 1045 < LMIC ≤ 4125; 4125 < UMIC ≤ 12,746; and HIC > 14,746. See: http://data.world bank.org/news/2015-country-classifications.

However, for a given percentage of investment in R&D, a low-income country will receive a higher score than a middle or high-income country.

The use of discrete income groups to differentiate levels of development has one important drawback: it creates discontinuities for countries lying on the border. A small change in the level of development can change the group a country belongs to, which can have significant consequences for those variables that strongly depend on it. This can lead to a situation where a small improvement nevertheless leads to a decrease in the index value, only because the country is now compared to an entirely different set of countries. It also biases the comparison between countries that lie on either side of the cut-off point. This problem can be avoided by using a continuous way of controlling for the level of development such as the *conditional* cumulative density function. The conditional CDF gives us the probability of finding a country with a lower score that has the same level of development: $F_{y|dev}(a) = p(y \leq a|dev)$. As illustrated in Panel *d*, the transformed values using the conditional CDF are very similar to those using fixed thresholds, but without the discontinuity problems for countries whose level of development is close to the thresholds.

The conditional cumulative density $F_{y|dev}$ is estimated using a multivariate kernel density estimator. An essential feature of this type of estimator is that it assigns a higher weight to information in the vicinity of the point of interest, both in terms of the variable and the level of development. The bandwidth of the estimator determines what counts as the vicinity: the larger the bandwidth is, the more the performance of dissimilar countries taken into account. By estimating the bandwidth, we can adjust it to match each indicator.[10] If an indicator is strongly dependent on the level of development, only the information of countries with a very similar level of development is taken into account. However, if the level of development does not affect *y*, the increase in the bandwidth ensures that more information is used, allowing us to estimate the CDF with greater certainty.

Finally, we ensured that a higher score is always an improvement in outcomes. If an indicator measures something positive (e.g., the share of renewable energy), the transformed indicator $\hat{y}_i$ indicates the probability of finding countries with a similar level of development (dev_i) that score lower: $\hat{y}_i = F_{y|dev}(y_i|dev_i)$. Conversely, for indicators that measure something negative (e.g., CO_2 emissions) $\hat{y}_i$ shows the probability of finding a country that scores higher: $\hat{y}_i = 1 - F_{y|dev}(y_i|dev_i)$.

3.4 Addressing Missing Values

The final issue that needs to be addressed before the index can be computed is how to deal with missing values. The IST index comprises a large number of indicators, and as can be seen in Table 1, the coverage of those indicators can be markedly different. When we limit the dataset to the period 1990–2016, slightly more than half of all observations are missing. A closer examination of the pattern in the missing data reveals that this cannot be easily solved. Figure 2 illustrates this by mapping the data availability using black (available) and white (missing) rectangles. Each row in this table corresponds to one of the 27 variables. The columns, in turn, show the different combinations in which the indicators are available, with the width of each column signaling how often this combination occurs

---

[10] The size of the bandwidth was estimated using the least squares cross-validation method (using Gaussian kernel). An advantage of this method in the estimation of a conditional probability is that it ignores the influence of extraneous variables. The estimation was done using the non-parametric np package in R (Hall et al. 2004; R Core Team 2015).

Table 1 IST indicators

Name	Description	Countries	Years	Source[b]
O1	*Manufacturing*			
O1.1	Manufacturing value added (% GDP)	207	1990–2015	UNIDO
O1.2	Manufacturing value added per capita (log)	207	1990–2015	UNIDO
O1.3	Share of medium and high tech industry (% value added)	148	1990–2015	UNIDO
O2	*Trade*			
O2.1	Exports of manufactured goods and commercial services per capita (log)	177	1980–2014	WTO
O2.2	Export volume in tons per capita (log)	215	1995–2015	CEPII
O2.3	Participation in global value chains	63	1995–2011	TiVA
O2.4	Share of medium and high tech exports	148	1990–2015	UNIDO
O3	*Innovation*			
O3.1	Research and development expenditure (% GDP)	132	1996–2015	WDI
O3.2	PCT Patents applied by residents per capita (log)	144	1970–2015	OECD
O3.3	Economic Complexity Indicator	121	1995–2016	ECI
O4	*Employment*			
O4.1	Manufacturing employment (% total employment)	156	1980–2016	WDI
O4.2	Labor productivity per worker in 2016 USD (log)	123	1950–2017	CB: TED
O4.3	Human capital	144	1950–2014	PWT8.1
O5	*Gender Inclusiveness*			
O5.1	Gender gap in employment (% male–female)[a]	186	1991–2016	WDI
O5.2	Gender equality rating	85	2005–2016	CPIA
O5.3	Gender Wage Gap (% male median wage)	38	1970–2016	OECD
O6	*Pollution*			
O6.1	CO_2 emissions (kg per GDP, in 2011 PPP)[a]	188	1990–2014	WDI
O6.2	PM2.5 air pollution, mean annual exposure[a]	194	1990–2015	WDI
O6.3	Consumption of ozone-depleting substances per capita (log)[a]	183	1986–2015	UNEP
O6.4	Municipal waste per capita[a]	36	1975–2015	OECD
O6.5	Municipal waste recovery (% of total)	36	1975–2015	OECD

Table 1 (continued)

Name	Description	Countries	Years	Source[b]
O7	*Energy*			
O7.1	Renewable energy share of TFEC (%)	220	1990–2014	SE4A
O7.2	National electrification rate (% population)	215	1990–2014	SE4A
O8	*Resource management*			
O8.1	Access to improved water source (% of population)	203	1990–2015	WDI
O8.2	Ocean Health Index	175	2012–2015	Halpern et al. (2012)
O8.3	Change in forest area (% land area)	210	1991–2015	WDI
O8.4	Terrestrial and marine protected areas (% territorial area)	209	1990–2014	WDI

[a]Indicators for which a higher value signals a worsening of outcomes. They are converted to the conditional probability of finding a country that scores higher, instead of a country that scores lower (cf. Sect. 3.2)

[b]The list of abbreviations can be found in "List of Abbreviations" of Appendix

Fig. 2 Availability of the indicators in the IST index. Note: Available data is represented by a black rectangle and missing values by a white rectangle. The width of the columns indicates the prevalence of the different combinations, which are listed in decreasing order

in the dataset. The columns are sorted in decreasing order of occurrence. The first column tells us that the most common combination is that only the national electrification rate (*O7.1*) is available. The second most prevalent combination (4% of the data) is the total absence of any information. In contrast, there are no observations that are covered by all indicators as there are always at least two indicators missing.

Our first step in tackling the missing data problem is to limit the number of countries for which we compute the index. Specifically, we exclude those countries for which the average data availability over the entire period is less than 25%. This removes mostly small island nations and city-states like Gibraltar, Nauru, or Vatican City, leaving 198 countries in the dataset. However, as Fig. 2 already suggested, limiting the dataset in this way will not solve the missing data problem, as there is no simple pattern in the missing observations. While removing the countries with the worst coverage does reduce the number of missing observations with one quarter, the total fraction of missing values remains high (43%).

In a similar vein, we could leave out the variables that have the lowest availability. For example, by excluding the participation in global value chains (*O2.3*), gender equality (*O5.2*) and the wage gap (*O5.3*), municipal waste (*O.6.4*), waste recovery (*O.65*), the Ocean Health index (*O8.2*) and the protection of terrestrial and marine protected areas (*O8.4*), the number of missing observations can be cut by half. However, changing the variables also affects the meaning of the index, as it amounts to imposing a zero weight on a quarter of the indicators in our dataset. Moreover, all but the most extensive reductions in the scope and span of the index would still leave differences in availability, making it hard to tell whether changes in the index are not due to differences in availability of the underlying indicators.

Rather than drastically limiting the scope of the IST index, we opt instead to solve the missing data problem using multiple imputation, i.e., filling in the gaps with the most likely value of each indicator. Unlike simple imputation (e.g., linear interpolation), multiple imputation draws many different possible values for each missing observation. This gives us a large number of imputed datasets, each of which yields a different index. The final result is computed as the average of these indexes. The variation of the index over different imputations also gives us an indication of the reliability of the results. The more data is missing and the worse the available data is at filling in the gaps, the greater the dispersion in the imputed values and the wider the confidence intervals of the index will be. In summary, multiple imputation allows us to compute the IST index despite the significant

data availability problems without having to reduce the scope of the index or omit variables ex-ante.

To estimate the most likely value of the missing observations, we use a method similar to the Multivariate Imputation by Chained Equations method (MICE) of Buuren and Groothuis-Oudshoor (2011). MICE uses the values of the available indicators to predict what the missing values could have been. For example, it will use the information on the number of patents applied to estimate what the level of investment in research and development could have been. While this works well in a cross-section analysis of countries, MICE ignores time-patterns in the data. As we have a panel dataset where many of the indicators depend strongly on their previous values, we expand the MICE method to a state-space model to also take this time-dependency into account. The so-called Multiple Imputation by State Space models (MISS) improves the quality of the imputations thereby significantly decreasing the size of the confidence intervals. Technical details on the imputation and a comparison of the two techniques can be found in "Multiple Imputation Using State-Space Model (MISS)" of Appendix.

To test the robustness of our imputation algorithm, we also ran the model on a reduced dataset where the data availability of each country was at least 50%. This reduced the total number of countries by half compared to our baseline model (which only imposed 25% availability). However, the imputation of the missing values remained virtually unaffected. For all but one variable, the correlation between the mean imputed values was in excess of 0.99. The change in the sample did affect the size of the confidence intervals, but not in a uniform way. One-third of the sample had confidence intervals that were at least 10% smaller; a third had larger confidence intervals, while the remainder had confidence intervals that were more or less the same.[11]

4 Indicators of Inclusive Sustainable Transformation

Our selection of indicators measuring IST is based to a large extent on the literature dealing with the measurement of progress towards the Sustainable Development Goals. Almost half of the development goals are directly linked to the idea of inclusive and sustainable transformation, including Goal 5 which is 'to achieve gender equality and empower all women and girls'; Goal 8 promoting 'sustained inclusive and sustainable economic growth, full and productive employment and decent work for all'; Goal 9 which is to 'build resilient infrastructure, promote inclusive and sustainable industrialization and foster development'; and Goal 12 which aims to 'ensure sustainable consumption and production patterns.'[12]

For primary sources of indicators, we look at two UN reports on the monitoring of the SGDS: Indicators and a Monitoring Framework for the SDGs (SDSN 2015) and the report of the Inter-Agency and Expert Group on SDG Indicators (ECOSOC 2016). This list of indicators is supplemented with those discussed in the papers by Sachs et al. (2016, 2017) and Kroll (2015) on the readiness of (developed) countries for the SDGs; the WEF and

---

[11] The effect on the final index could not be determined as the index compares all countries to each other and the set of countries in the robustness check is only half that of the baseline model.

[12] Other SDGs that pertain to Inclusive and Sustainable Structural Transformation are goals (6) 'ensure availability and sustainable management of water and sanitation for all'; (7) 'ensure access to affordable, reliable, sustainable and modern energy for all'; and (11) 'make cities and human settlements inclusive, safe, resilient and sustainable'.

IMD's reports on (sustainability adjusted) global competitiveness; and the Human Development Index. From these reports, we retained those indicators that were available for a large group of countries over the past decade. Since the focus is on structural transformation rather than economic growth and development in general, there are differences between the indicators listed in these sources and those selected for the IST index. For example, with respect to inclusiveness, only the indicators related to labor markets are retained, leaving out, e.g., the proportion of seats held by women in parliament. "Comparison of the Indicators Included in the IST Index with Other Indicators of Sustainable Development" of Appendix maps our selection of indicators on the different sources, highlighting variables like participation in global value chains or the economic complexity index that are unique to the IST index. Nevertheless, three-quarters of the indicators are used in one or more of the reports mentioned above, and most are described in great detail in the ECOSOC and SDSN reports.

The IST index can be subdivided into eight categories that each contain between two and five indicators, which are listed in Table 1. In addition to a short description of each indicator, this table lists the availability of each indicator over time, the number of countries it covers and the source of the data.[13]

The first two components of the IST index look at the strength of the manufacturing and export sectors. The former includes the value added by the manufacturing sector per capita, as well as its share in GDP. The latter considers the export of manufactured goods and commercial services and also includes the volume of exports to compensate for sudden shifts in the terms of trade. Because of their increasing importance to global trade, we also include an index that captures the extent to which countries participate in global value chains. Finally, included in both components is the contribution of medium and high tech firms to the value added of the manufacturing and exports, respectively.

The third IST component measures the technological expertise embedded in a country's economy. To that end, we track the overall investment in research and development, the number of patents per capita and the complexity of a country's export basket. The last is measured using the economic complexity index of Hausmann et al. (2011), which combines information on the diversity of goods a country produces with the ubiquity of those goods (i.e., the number of countries that are capable of producing them).

The fourth and fifth components deal with the strength and inclusiveness of the labor market. For the former, we include indicators of the number of people working in manufacturing,[14] their labor productivity and their level of education. Gender equality is measured as the ratio of male and female employment, gender differences in wages as well as the existence and strength of institutional policies promoting equal opportunities for men and women.

The final three components consider the environmental performance. The first one looks at pollution. Air pollution levels are measured by CO_2 emissions, the abundance of fine particle matter in the air and the consumption of ozone-depleting substances. This component

---

[13] There are two differences between the current selection of indicators and the selection of indicators of the working paper published by the African Development Bank, Ghent University and Peking University's Center for New Structural Economics. First of all, the data on patents for indicator *O3.2* now comes from the World Bank's World Development Indicators as opposed to the OECD. Secondly, labor productivity (*O4.2*) is now measured per worker as opposed to per hour. Both changes were made because they significantly increase the number of countries covered.

[14] Services are excluded, as we were not able to distinguish traditional from modern services.

also includes the total municipal waste that is generated but counterbalances it with the percentage of waste that is recycled or composted. The second environmental component looks at the structure of the energy market, in particular, the fraction of the population that has access to modern energy (i.e., electricity) and the share of renewable energy in the total energy consumption. The eighth and final component evaluates the management of environmental resources. This includes the percentage of the population with access to drinkable water, the annual change in forest area (as a percentage of land area), the percentage of terrestrial and marine area that is environmentally protected, and an index tracking the health of the ocean.[15]

Looking at the list of indicators, it is clear that there is often a significant overlap in what is measured. For example, the manufacturing subcomponent (*O1)* contains value added by the manufacturing sector both as a percentage of GDP (*O1.1*) as well as per capita (*O1.2*). It could be argued that it would be better to reduce the number of indicators, given the relatively marginal contribution of this second indicator to measuring the strength of the manufacturing sector. However, there are a number reasons for allowing this apparent surplus of indicators. First, while one indicator can proxy the overall state of affairs, the inclusion of different indicators enables us to build a complete image of the current situation. Second, the overall contribution of some variables might be relatively limited, but their inclusion can be more important for specific groups of countries. One such example is the national electrification rate (*O7.2*), which matters much more for developing than developed countries. Finally, the inclusion of indicators with different availabilities enhances the performance of the multiple imputation algorithm. For instance, while the number of patents per capita (*O3.2*) is available for a longer period, it covers fewer countries than the expenditure on research and development (*O3.1*). In general, even if the selection is disputed, we will show in the robustness section that our results remain virtually unaffected by the exclusion of any one of the indicators.

As is the case for all components of the index, the environmental variables are also scored conditional on the level of development of the country. This might seem incongruous with variables that have global environmental consequences, most noticeably CO_2 emissions. However, the goal of the index is not to measure environmental impact, for which there already exist numerous indicators and indexes of high quality that, for example, also take consumption and offshoring of polluting activities into account (e.g., ecological footprint). Instead, the IST index measures how well the environment is protected given the available means, even if the consequences of pollution are not limited to the country in question. Furthermore, as the index is focused on structural transformation, it does not include a measure of the sustainability of agriculture, although there is some overlap with the indicators that are included.[16]

Finally, while our index only considers the level of development, the capability to score well on a number of indicators depends on more than just the level of development. For example, the location of a country determines access to certain sources of renewable energy, while the lack of access to the sea significantly increases the cost of trade. That being said, many of these problems can be overcome with the right investments, meaning that even in these cases the level of development remains an important factor. Moreover,

---

[15] The IST index for landlocked countries does not include the ocean health variable.

[16] See e.g., http://www.wri.org/publication/indicators-sustainable-agriculture-scoping-analysis.
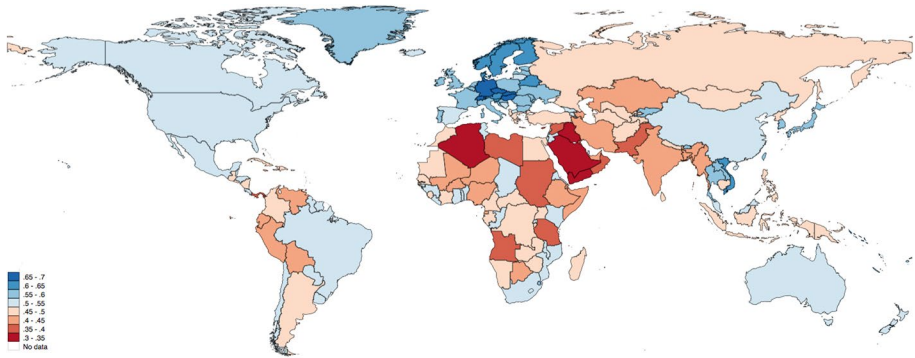
Fig. 3 Map of the IST index in 2015. Note: Blue denotes countries that perform better than average and red those that perform below average. The darker the color, the more above/below average the score. (Color figure online)
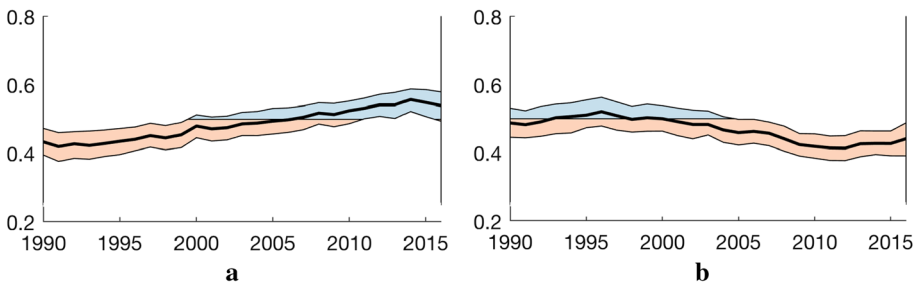

Fig. 4 The IST index over time. **a** China and **b** United Arab Emirates. Notes: IST index of China and the United Arab Emirates from 1990 to 2015. The colored band represents the 95% confidence interval, where below average values are in orange and above average values are in blue. (Color figure online)

technological progress is likely to continue to increase its importance as an impediment towards sustainable, inclusive structural transformation.

5 The Inclusive Sustainable Transformation Index

Using the dataset and methodology described above, we computed the IST index from 1990 to 2016 for 198 countries. As the availability of indicators drops to one in four in 2016, we will focus the discussion of the index on the year before. However, except for an increase in the confidence bands in the final year, the results are very similar in 2016.

Figure 3 shows the worldwide distribution of the IST index in 2015. Countries that score above average are colored blue, and those that score below average are colored red, with the darker colors corresponding to respectively higher or lower values. While in theory, the values of the IST index can lie between zero and one, we find that the actual values of the index lie between 0.3 and 0.7. The values on the individual components lie much closer to the theoretical extremes, indicating that countries that score very high on one component will compensate this with lower scores on other components. Overall, the values of the index tend to be slightly negatively skewed, with below average scores centered

Fig. 5 IST indicators in 2015 for low, middle and high-income countries. **a** Tanzania, **b** Chad, **c** Yemen, **d** ▶
Belarus, **e** Saudi Arabia and **f** Germany. Notes: Radar plot of the transformed IST indicators in 2015. The
90% confidence intervals are indicated by the gray marked area. The indicators are grouped into O1: Manufacturing; O2: Trade; O3: Innovation; O4: Employment; O5: Gender Inclusiveness; O6: Pollution; O7:
Energy; O8: Resource Management. The full description of each indicator can be found in Table 1

on 0.45, and above average scores having a fatter tail. In other words, most countries that
score below average tend to do so only slightly, and there are more countries with a very
high than with a very low score.

By taking the level of development into account when comparing the structural characteristics of countries, the IST index can identify good performance despite a lower level of
development. For example, Vietnam's IST score is higher than all other countries on the
Asian, American, and African continent. At the same time, the overall picture mostly confirms our expectations. Except for Greece, European countries score highly, and while Central and North America also score above average, South America scores below. Countries
in Southeast Asian also tend to score above average and while those in Africa show more
mixed results, many Southeast African countries tend to perform well. The high scores for
some of the high-income European countries like Austria (0.64), Sweden (0.64) and Finland (0.63) are due in part to the fact that they outperform other high-income countries like
Saudi Arabia (0.34), Oman (0.37) and Bermuda (0.38), particularly on the environmental
and equality components. However, some of the best scores on the European continent are
accrued by lower- and upper-middle income countries in Eastern Europe, namely Hungary
(0.66), Slovenia (0.65) and Slovakia (0.65).

When using the index to make comparisons between countries or over time, it is important to keep in mind that they always reflect a country's position relative to those with a
similar level of development in that year. As a result, a decrease in the index does not necessarily mean that a country's absolute achievement deteriorated: it could also mean that
other countries made (more) progress. That being said, the IST index is relatively stable: its
variation between countries is more than twice as large as its variation over time. There are
countries whose score has changed dramatically over the past decade. To illustrate, Fig. 4
show the evolution of the IST score of China (panel a) and the United Arab Emirates (panel
b), together with their 95% confidence intervals. For both countries, the index changes only
gradually and the confidence intervals indicate that the year-to-year changes are all insignificant. Nevertheless, when considering the evolution over a more extended period, there
are 46 countries where changes in the index are big enough that the 95% confidence intervals no longer overlap, including China and the United Arab Emirates.

Before comparing our index with those suggested by Sachs et al. (2017) and Kroll
(2015), it is important to note that these indexes do not consider the time dimension. Both
use only the latest available values for each indicator, which in some cases date back to
2012. Their correlation with the 2015 values of IST is 0.54 and 0.43, respectively. These
rise slightly when we control for GDP per capita: the respective partial correlation coefficients rise to 0.56 and 0.58. As the IST conditions on the level of development, the correlation between IST and per capita GNI is low (0.17), but its correlation with the HDI is
higher (0.34). This is the principal difference with the indexes of Sachs et al. and Kroll,
which are strongly correlated with the level of development (0.60 and 0.67 with GNI/cap;
0.92 and 0.79 with HDI), even though Kroll only compares developed countries.

Having considered the overall IST scores, the next step is to look at the underlying indicators to better understand how certain scores came about. To that end, Fig. 5 shows the
transformed scores of the indicators in the IST index for two low, two middle and two
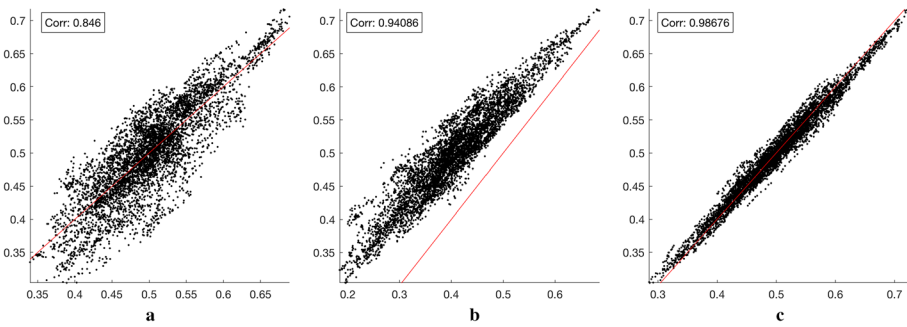
a

b

c

d

e

f
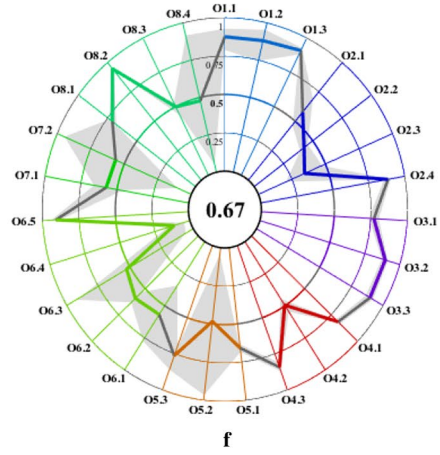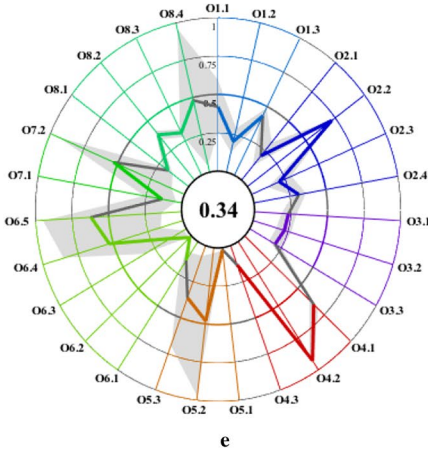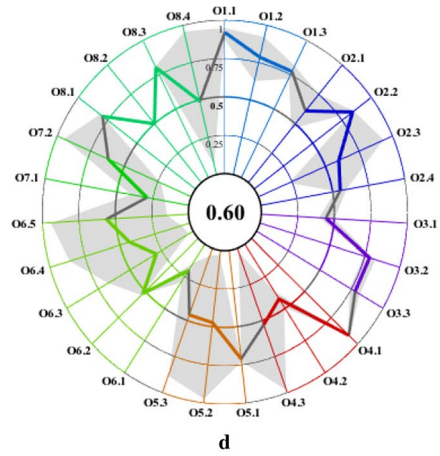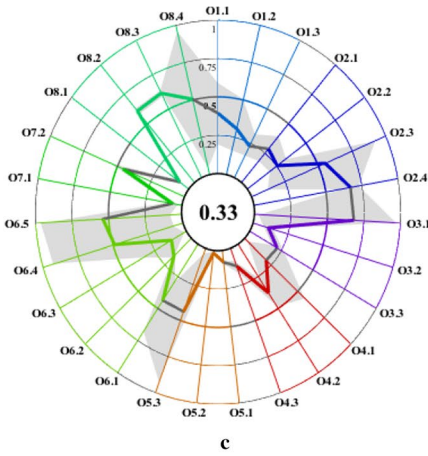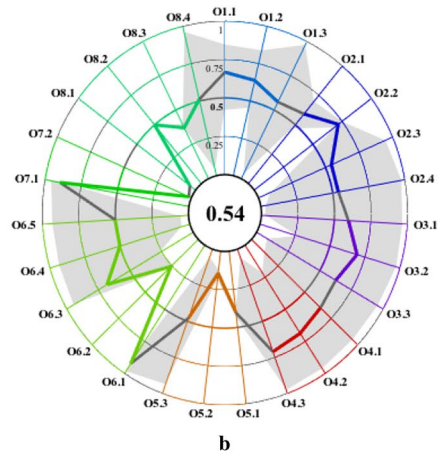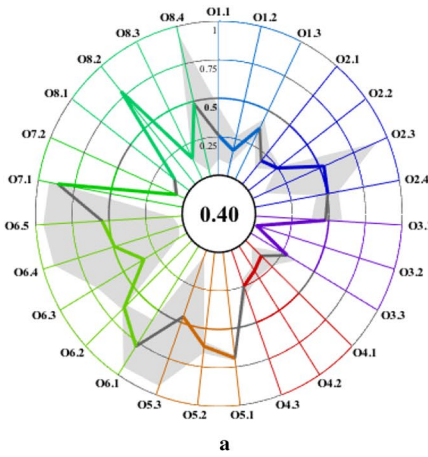
Fig. 6 Scatter plots of the baseline IST index (y-axis) with robustness checks (x-axis). **a** HDI, **b** geometric, **c** grouped. Note: The 45° line is indicated using the red line. (Color figure online)

high-income countries in 2015 using a radar chart. The countries in the left column have one of the lowest scores in their development group, while the countries on the right have of the highest scores. These graphs further illustrate that each component of the index is scored relative to the level of development. Take, for example, the Ocean Health Index (O8.2). Tanzania's transformed score is twice that of Saudi Arabia (0.79 vs. 0.39), even though it has a lower score on the Ocean Health Index (55.6 vs. 66.3). The reason is that its higher level of development means that Saudi Arabia is compared to a group of countries that protect their oceans better, like Germany that has a perfect score on the OHI.

These radar charts can be a useful tool for economic and development policy, as they highlight those policy areas that require more attention, as well as specific problems that need to be addressed. Moreover, they can guide countries towards policies that work for countries with similar levels of development. For example, panel *a* links Tanzania's low score to its poor performance on the manufacturing component. Rather than to try and emulate the economic structure of high-income countries, Tanzania could look at the economic policies of a country like Swaziland that scores highly on Manufacturing value added and Manufacturing value added per capita (over 0.94). In contrast, Germany scores highly on almost all components and as a result, ends up with the highest score in 2015. Nevertheless, Germany's score could even higher if it managed to decrease the total amount of municipal waste that is created. For help in achieving this goal, Germany could take a closer look at Belgium, Korea or Iceland. All three countries have similar levels of development but score exceptionally well on this component: 0.86, 0.91 and 0.94, respectively, versus Germany's 0.11.

5.1 Robustness Checks

In this final section, we determine the IST index' sensitivity to our modeling choices. First, we look at how the results change when an alternative measure of the level of development is used. To that end, we use the UN's Human Development Index (HDI), a composite index that combines GNI per capita with life expectancy and education level. With the HDI as our measure of development, the individual indicators from Table 1 are transformed and the results are combined into a second index: IST^{HDI}. While there are some differences between our baseline index and IST^{HDI}, their overall correlation is high (Fig. 6, panel a). The development category where the most prominent changes take place is in the lower-middle income group. While the holistic nature of the HDI might make it a more appealing

choice as a measure of development, there are two reasons why we use GNI per capita instead. First of all, there is an overlap between HDI and the indicator of human capital included in the index (*O4.3*). Secondly and most importantly, the HDI is only available every five years between 1990 and 2010.[17]

As a second robustness check, we consider the effect of using a geometric average to combine the indicators. Unlike the arithmetic average, which imposes perfect substitutability between the components, the geometric mean penalizes countries with asymmetric component scores. The effect on the ranking by the index is minimal, as the correlation between IST^{GEO} and the baseline IST is 0.94. Nevertheless, as panel b of Fig. 6 shows, the scores are lower using the geometric average. Some countries find their score significantly decreased, like Libya which sees a 44% decrease (from 0.32 to 0.18). While the distribution of IST is negatively skewed, IST^{GEO} has a more symmetric distribution. On the other hand, it also has much wider confidence intervals. Given that the results are so similar, we opt for the arithmetic average as it provides a more straightforward interpretation (i.e., the average fraction of countries that score better).

We also check how the index changes when variables are omitted. In this regard, the index is recomputed 26 times using all but one of the indicators. Regardless of which variable is left out, the index and its standard deviations are almost identical: the correlation of both exceeds 0.98 each time (Fig. 6, panel c). Finally, as some categories contain more variables than others, we also check how the results change when each category, rather than each indicator, receives equal weight in the final index. Similar to the leave-one-out estimations, the results are practically identical.

6 Conclusion

The universal adoption of the Sustainable Development Goals and the successful conclusion of the Paris Climate Summit were seen as turning points in the pursuit of shared global prosperity. However, the monitoring of these goals has become a significant challenge, especially since economies around the world are at different levels of development and have different production structures. This paper proposes the Inclusive Sustainable Transformation (IST) index as a contribution to the monitoring of these global objectives.

The IST index measures the extent to which a country has developed a modern economy that protects the environment and is gender inclusive. In contrast with other development indicators, the level of development is taken into account when the structural characteristics of countries are compared. This is in line with New Structural Economics thinking, which posits that a country's most optimal development strategy depends on its level of development. To make this conditional comparison, we employ a continuous method of transformation (a conditional CDF) that does not bring about structural breaks in the index. Our results show that taking the level of development into account can reveal patterns that are otherwise hidden, revealing a number of countries that performed much better and worse than expected.

Given the ambitious scope of the index, both in terms of countries covered and indicators included, missing data is a big concern. However, we address this problem using

---

[17] The values for the intervening years are typically linearly interpolated, although we used the MISS algorithm.

multiple imputation. This allows us to estimate the relative performance of close to 200 countries and provides us with an estimate of how the reliability of the index is affected by missing data.

Rather than only measuring a country's overall progress, we focus on how the different components of the index contribute to the overall score. To that end, radar graphs accurately show the disaggregated results and allow us to identify those policy areas that are leading or lagging quickly. By breaking the IST scores down to their different components, policymakers and analysts can identify 'best practices' among countries with a similar level of development on a wide range of policies.

Appendix

List of Abbreviations

BACI	Base pour l'Analyse du Commerce International (Gaulier and Zignago 2010)
CB:TED	The Conference Board Total Economy Database
CEPII	Centre d'Études Prospectives et d'Informations Internationales
CIP	Competitive Industrial Performance index
CPIA	Country Policy and Institutional Assessment
ECOSOC	United Nations Economic and Social Council
GFD	Global Financial Development
i-Tip	Integrated Trade Intelligence Portal
IEA	International Energy Agency
IFC	International Finance Corporation
ILO	International Labour Organization
IMD	Institute for Management Development
IMF	International Monetary Fund
MVA	Manufacturing Value Added
OECD	Organization for Economic Co-operation and Development
PWT	Penn World Tables (Feenstra et al. 2015)
SDSN	Sustainable Development Solutions Network
SE4A	Sustainable Energy for All
TiVA	Trade in Value-Added
UNEP	United Nations Environment Programme
UNIDO	United Nations Industrial Development Organization

WB   The World DataBank
WDI   World Development Indicators
WEF   World Economic Forum
WITS   World Integrated Trade Solution
WTO   World Trade Organization

Multiple Imputation Using State-Space Model (MISS)

This section gives a short summary of how gaps in the data are filled in using the MICE and MISS multiple imputation algorithms. Multiple imputation using chained equations (MICE) uses both the available and imputed data of the other variables to estimate possible values for the missing data. However, this creates a self-referential problem. Filling in the gaps in the first variable requires the imputed values of the second variable, but to compute those we need the imputed values of the first variable. As Buuren and Groothuis-Oudshoor (2011) explain, this self-referential problem can be solved by repeatedly running through the estimation algorithm until the values converge.

Formally, we can write the MICE algorithm down as the following loop. Let X_t be a vector containing the original variables at time t, and $\hat{X}_t$ be the vector containing the imputed values (where the bold font is used to denote vectors).[18] Using the superscript, we can select specific variables from these vectors: X_t^i is the value of the ith variable at time t, while X_t^{-i} contains the entire X_t vector except for the variable i. The MICE algorithm works in the following way[19]:

1. Initialize the model by replacing the missing data of each variable with their sample average
2. Estimate a linear model that captures the dependence between the first variable and the rest of the dataset: $X_t^1 = \alpha^1 + \hat{X}_t^{-1} \beta^1 + e_t^1$. This gives us the estimated parameters a^1 and b^1.
3. Fill in the gaps by predicting the value of the first variable using a and b: $\hat{X}_t^1 = a^1 + \hat{X}_t^{-1} b^1$
4. Run step 2 and 3 for all other variables in the dataset.
5. Repeat steps 2 through 4 until the model has converged.

While MICE uses all cross-sectional information, it ignores the information from the past and future values of the indicators. In contrasts, methods like (linear) interpolation use only the past and future values of a variable to fill in possible values for the missing data points. This implies that MICE is preferable over interpolation only when the other variables are better predictors of the missing data than that variable's past and future values.

The Multiple Imputation by State Space models (MISS) combines both techniques by adding this interpolation component to the MICE model. That is, it uses the information contained in the (imputed) data of the other variables, as well as the past and future (imputed) values of the variable itself to determine the most likely value for each missing data point. To combine the information from the other variables with the own past and

---

[18] For convenience, we are writing down the model for a time-series model, but these techniques can be equally applied to a panel dataset.

[19] As all variables are continuous, we can estimate the dependencies between the variables using a linear regression model. However, both MICE and MISS techniques can be adapted to deal with binary or categorical data.

future values, we make use of a state-space model. As is the case with MICE, the self-referential nature of MISS can be solved by iteratively running the algorithm.

As there are entire books devoted to state-space models and how to estimate them (e.g., Kim and Nelson 1999; Durbin and Koopman 2012) we will keep the explanation short and refer the interested reader to these sources. A state-space model is a dynamic model that contains unobserved variables called state variables. It typically consists of two equations. The *measurement equation* describes how the observed variables are related to the unobserved and to-be-estimated state variable. The *state equation* describes the dynamic pattern in the state variables: i.e., how it depends on its previous values. When estimating the state-space model, the most likely values of the unobserved variable are determined as a weighted average of the information in the observed variables and that in the past and future values of the state variable. The weights are determined by how reliable the observed data is versus how strongly the variable depends on its past values.

In this case, the unknown state variable is the to-be-imputed variable $\left(\hat{X}_t^i\right)$. As was the case in the MICE model, we use the (imputed) values of the other variables $\left(\hat{X}_t^{-i}\right)$ as observed variables and assume a linear relation: $\hat{X}_t^j = \alpha^j + \beta^j \, \hat{X}_t^i + \epsilon_t^j$ with $j = \{1, \dots, k\}$ and $i \neq j$. The variance of the error term of ϵ^j captures the extent to which the imputed variable $\hat{X}^j$ is a good predictor. Naturally, if the data is not missing the imputed data has to be identical to the observed data $(\hat{X}_t^i = X_t^i)$. Using matrix notation, these equations can be summarized into the following measurement equation:

$$\begin{bmatrix} \hat{X}_t^{-i} \\ X_t^i \end{bmatrix} = \begin{bmatrix} \alpha^{-i} \\ 0 \end{bmatrix} + \begin{bmatrix} \beta^{-i} \\ 1 \end{bmatrix} \hat{X}_t^i + \begin{bmatrix} \epsilon_t^{-i} \\ 0 \end{bmatrix}$$

As the number of years in the dataset is relatively limited, the autocorrelation of each variable is simply modeled as an autoregressive process with one lag. This gives us the following state equation.

$$\hat{X}_t^i = A^i \, \hat{X}_{t-1}^i + \mu_t^i$$

More details on how to estimate this model can be found in Kim and Nelson (1999).

The MISS algorithm then runs through the following steps:

1. Initialize the model by replacing the missing data with the average value
2. Estimate the k-1 measurement equations describing the relation between the first variable and the other variables in the dataset: $X_t^j = \alpha^j + \beta^j \, \hat{X}_t^i + \epsilon_t$.
3. Estimate the parameters of the state-equation using the actual values of X^1: $X_t^i = A^i X_{t-1}^i + \mu_t^i; \forall i \neq j$.
4. Stack the estimated parameters of the state and measurement equation and use state-space model techniques (i.e., a Kalman filter and simulation smoother) to draw new values for $\hat{X}_t^1$.
5. Run steps 2 and 3 for all other variables in the dataset.
6. Repeat steps 2 through 5 until the imputed values converge.
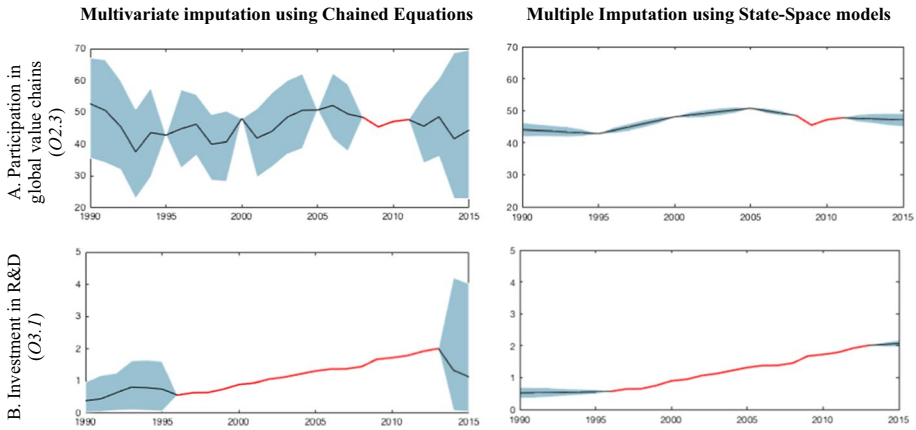
Fig. 7 MISS versus MICE. Note: Plot of the actual (red) and imputed (black) data on global value chains and investment in R&D of China. The values in the graphs on the left were imputed using MICE, while those on the right using MISS. 95% confidence intervals of the imputed values are indicated by the blue shaded area. (Color figure online)

Table 2 Monte Carlo simulation on the reliability of MISS

	10%	20%	30%	40%	50%	60%	70%	80%	90%
Average bias	−0.0062	0.0011	0.0004	−0.0026	−0.0022	−0.0031	−0.0218	0.0266	−0.012
SD bias	0.0367	0.0251	0.0312	0.0366	0.0454	0.0701	0.1327	0.195	0.1865
Average SD	0.3189	0.3222	0.3482	0.3611	0.3749	0.3968	0.4251	0.4796	0.5675

Comparison with MICE

As Fig. 7 illustrates, the effect of MISS on the imputation of missing values can be substantial. Especially for variables that are available every 5 years (panel a) or that depend strongly on their previous values (panel b), the range of imputed values is drastically reduced when using the state-space technique. This decrease in the variance of the imputed values in turn leads to a smaller variance in the transformed indicators and the IST index.

Monte Carlo Simulation

In order to get a better understanding of how the model performs when the number of missing values increases, we ran a Monte Carlo simulation on a generated dataset whose characteristics mimic the dataset of the IST index. Specifically, we first generated nine variables (1000 observations each) that have both an autoregressive part and are moderately correlated to two other variables:

$$X_t = 0.9 * X_{t-1} + \mu_t \quad \text{with } \mu_t \sim N(0, \Sigma)$$

where $\Sigma = (\delta\delta')^{-1}$ and δ is an upper triangular matrix filled with 0.5.

After normalizing the data, we subsequently randomly deleted 10% of the observations of the first variable, 20% of the second, and so on until the last variable only has 10% of his original observations left. The MISS algorithm was then used to try to fill in the gaps in the dataset. The MISS estimator ran for 1100 iterations of which the first 1000 were discarded as burn-in, and the entire Monte Carlo simulation was repeated a hundred times.

The results are shown in Table 2. The first two rows compare the imputed values of the MISS algorithm with the original values. The bias is the difference between the original values of the variable and the average value returned by the MISS algorithm. The first row of Table 2 shows the average bias over the Monte Carlo simulations, while the second row shows the standard deviation of the biases. This reveals that even when the 90% of the data is missing, the MISS algorithm returns the rights coefficients on average. However, the standard deviation of the bias does increase as the number of missing values increases. In line with expectations, row three shows that the confidence with which the MISS algorithm can fill in the missing values gradually decreases as the fraction of missing values increases. Nevertheless, the average standard deviation of the imputed values remains well below 1.4, which is what you would get if these values were filled using random draws from a normal distribution, meaning that they remain informative.

Overall, the Monte Carlo simulations support the earlier finding that when the dataset is reduced to only those countries with more than 50% availability, the results of the MISS algorithm remain the same.

Comparison of the Indicators Included in the IST Index with Other Indicators of Sustainable Development

		Sachs et al. (2017)	Kroll (2015) Sustainable Development Goals: Are the Rich Countries Ready?	Sustainable Development Solutions Network Indicators and a monitoring framework for the sustainable development goals	UN economic and social council Report on inter-agency and expert group on Sustainable Development Goal Indicators	Institute of Management Development World Development yearbook	World Economic Forum Global Competitiveness Report (Sustainability adjusted)	UN Human Development Index
GNI/cap (measures the level of development)	5		Included	Included		Included	Market size index	Included
O1: Manufacturing								
O1.1 Manufacturing value added (% GDP)	2			Included	Included			
O1.2 Manufacturing value added per capita	0							
O1.3 Share of medium and high tech industry	1				Included			
O2: Trade								
O2.1 Exports of goods and services	2					Included	Included	
O2.2 Export volume in tons per capita (log)	0							
O2.3 Participation in global value chains	0							

			Sachs et al. (2017)	Kroll (2015) Sustainable Development Goals: Are the Rich Countries Ready?	Sustainable Development Solutions Network Indicators and a monitoring framework for the sustainable development goals	UN economic and social council Report on inter-agency and expert group on Sustainable Development Goal Indicators	Institute of Management Development World Development yearbook	World Economic Forum Global Competitiveness Report (Sustainability adjusted)	UN Human Development Index
O2.4	Share of medium and high tech exports	1					Included		
O3: Innovation									
O3.1	Research and development expenditure (% GDP)	5	Included	Included	Included	Included	Included		
O3.2	PCT. Patents per capita (log)	4	Included			Green patents		Included	
O3.3	Economic Complexity Indicator	0							
O4: Employment									
O4.1	Manufacturing employment (% total employment)	5	Unemployment	Unemployment (% population)	Included	Included	Included		
O4.2	Labor productivity per hour worked in 2014 USD	0							
O4.3	Human capital	5	Various indicators of schooling			Various indicators on schooling	Secondary school enrollment	Various education levels	Education index

			Sachs et al. (2017)	Kroll (2015) Sustainable Development Goals: Are the Rich Countries Ready?	Sustainable Development Solutions Network Indicators and a monitoring framework for the sustainable development goals	UN economic and social council Report on inter-agency and expert group on Sustainable Development Goal Indicators	Institute of Management Development World Development yearbook	World Economic Forum Global Competitiveness Report (Sustainability adjusted)	UN Human Development Index
O5: Gender Inclusiveness									
O5.1	Gender gap in employment (% Male - Female)	2	Included						
O5.2	Gender equality rating	0					Included		
O5.3	Gender Wage Gap (% male median wage)	3	Included	Included	Included				
O6: Pollution									
O6.1	CO_2 emissions (kg per 2011 PPP of GDP)	5	Included	Production based CO_2 emissions	Included	Included	Included	Included	
O6.2	PM2.5 air pollution, mean annual exposure	5	Included	Included	Included	Included	Included	Included	
O6.3	Consumption of ozone depleting substances	1		Included	Included				
O6.4	Municipal waste per capita	2	Included	Included				Included	

			Sachs et al. (2017)	Kroll (2015) Sustainable Development Goals: Are the Rich Countries Ready?	Sustainable Development Solutions Network Indicators and a monitoring framework for the sustainable development goals	UN economic and social council Report on inter-agency and expert group on Sustainable Development Goal Indicators	Institute of Management Development World Development yearbook	World Economic Forum Global Competitiveness Report (Sustainability adjusted)	UN Human Development Index
O6.5	Municipal waste recovery	3	Included		Included	Included			
O7: Energy									
O7.1	Renewable energy share of TEC	5	Included	Included	Included	Included	Included		
O7.2	National electrification rate	4	Included		Included	Included		Included	
O8: Resource Management									
O8.1	Access to improved water source	4	Included		Included	Included		Included	
O8.2	Ocean Health Index	1	Incl. seperately						
O8.3	Change in forest area	4	Included		Included	Net permanent forest loss	Included		
O8.4	Terrestrial and marine protected areas	4	Incl. seperately	Terrestrial only	Included Incl. seperately	Incl. seperately			

Overall evolution of the IST index from 2000 to 2015

Name	2000	2005	2010	2015	Name	2000	2005	2010	2015
Afghanistan	0.444	0.469	0.463	0.475	Lebanon	0.394	0.457	0.435	0.476
Albania	0.443	0.507	0.499	0.466	Lesotho	0.498	0.560	0.542	0.526
Algeria	0.361	0.343	0.344	0.343	Liberia	0.486	0.492	0.486	0.510
Andorra	0.569	0.551	0.550	0.537	Libya	0.326	0.325	0.324	0.388
Angola	0.377	0.379	0.389	0.386	Lithuania	0.531	0.563	0.555	0.540
Antigua and Barbuda	0.459	0.447	0.486	0.471	Luxembourg	0.537	0.478	0.447	0.468
Argentina	0.470	0.471	0.450	0.454	Macao SAR, China	0.366	0.390	0.396	0.405
Armenia	0.588	0.528	0.492	0.502	Macedonia, FYR	0.535	0.529	0.529	0.531
Aruba	0.497	0.508	0.482	0.485	Madagascar	0.439	0.450	0.458	0.454
Australia	0.511	0.500	0.494	0.504	Malawi	0.536	0.508	0.515	0.515
Austria	0.638	0.641	0.640	0.644	Malaysia	0.577	0.536	0.563	0.548
Azerbaijan	0.478	0.478	0.447	0.420	Maldives	0.438	0.472	0.464	0.490
Bahamas, The	0.454	0.486	0.508	0.522	Mali	0.435	0.455	0.439	0.435
Bahrain	0.353	0.380	0.389	0.385	Malta	0.574	0.550	0.529	0.480
Bangladesh	0.428	0.427	0.424	0.419	Marshall Islands	0.550	0.549	0.535	0.539
Barbados	0.515	0.527	0.549	0.517	Mauritania	0.471	0.451	0.459	0.456
Belarus	0.643	0.639	0.597	0.601	Mauritius	0.517	0.492	0.476	0.473
Belgium	0.605	0.601	0.579	0.599	Mexico	0.498	0.489	0.481	0.515
Belize	0.504	0.498	0.536	0.543	Micronesia, Fed. Sts.	0.580	0.585	0.570	0.554
Benin	0.523	0.516	0.510	0.516	Moldova	0.566	0.578	0.566	0.582
Bermuda	0.401	0.394	0.388	0.376	Mongolia	0.426	0.442	0.463	0.452
Bhutan	0.580	0.593	0.578	0.562	Montenegro			0.517	0.480
Bolivia	0.456	0.423	0.443	0.439	Morocco	0.482	0.488	0.478	0.458
Bosnia and Herze- govina	0.461	0.514	0.528	0.517	Mozambique	0.528	0.544	0.514	0.507
Botswana	0.484	0.454	0.452	0.446	Myanmar	0.355	0.388	0.427	0.413
Brazil	0.491	0.500	0.498	0.505	Namibia	0.455	0.467	0.467	0.485
Brunei Darussalam	0.439	0.407	0.444	0.449	Nepal	0.454	0.448	0.485	0.475
Bulgaria	0.586	0.579	0.561	0.593	Netherlands	0.579	0.575	0.563	0.571
Burkina Faso	0.439	0.442	0.416	0.430	New Caledonia	0.565	0.563	0.559	0.552
Burundi	0.459	0.524	0.533	0.498	New Zealand	0.535	0.535	0.505	0.505
Cambodia	0.426	0.398	0.442	0.468	Nicaragua	0.475	0.463	0.468	0.478
Cameroon	0.500	0.494	0.466	0.472	Niger	0.445	0.437	0.418	0.420
Canada	0.572	0.564	0.539	0.525	Nigeria	0.407	0.377	0.380	0.410
Cape Verde	0.501	0.506	0.492	0.505	Norway	0.586	0.620	0.628	0.625
Central African Republic	0.537	0.514	0.502	0.488	Oman	0.337	0.352	0.407	0.371
Chad	0.503	0.527	0.532	0.537	Pakistan	0.370	0.389	0.402	0.395
Chile	0.507	0.490	0.511	0.505	Palau	0.476	0.482	0.454	0.474
China	0.480	0.494	0.525	0.549	Panama	0.441	0.440	0.433	0.390
Colombia	0.455	0.425	0.432	0.452	Papua New Guinea	0.487	0.490	0.494	0.496
Comoros	0.506	0.517	0.523	0.531	Paraguay	0.476	0.485	0.493	0.520
Congo, Dem. Rep.	0.463	0.467	0.499	0.471	Peru	0.393	0.409	0.413	0.401

Name	2000	2005	2010	2015	Name	2000	2005	2010	2015
Congo, Rep.	0.485	0.526	0.516	0.500	Philippines	0.508	0.490	0.451	0.494
Costa Rica	0.529	0.553	0.533	0.528	Poland	0.567	0.547	0.534	0.555
Cote d'Ivoire	0.519	0.510	0.493	0.487	Portugal	0.472	0.498	0.513	0.555
Croatia	0.592	0.571	0.592	0.578	Puerto Rico	0.442	0.471	0.516	0.497
Cuba	0.498	0.488	0.484	0.485	Qatar	0.429	0.482	0.502	0.520
Cyprus	0.373	0.382	0.381	0.385	Romania	0.508	0.544	0.561	0.577
Czech Republic	0.643	0.635	0.652	0.663	Russian Federation	0.583	0.523	0.482	0.459
Denmark	0.610	0.579	0.602	0.601	Rwanda	0.481	0.482	0.501	0.501
Djibouti	0.393	0.407	0.411	0.406	Samoa	0.593	0.602	0.575	0.588
Dominica	0.552	0.532	0.485	0.503	Sao Tome and Principe	0.521	0.519	0.486	0.515
Dominican Republic	0.517	0.510	0.496	0.497	Saudi Arabia	0.325	0.340	0.356	0.341
Ecuador	0.474	0.426	0.439	0.446	Senegal	0.477	0.477	0.475	0.483
Egypt, Arab Rep.	0.451	0.446	0.468	0.474	Serbia			0.566	0.575
El Salvador	0.465	0.467	0.455	0.459	Seychelles	0.486	0.464	0.444	0.486
Equatorial Guinea	0.478	0.421	0.421	0.416	Sierra Leone	0.447	0.457	0.451	0.494
Eritrea	0.511	0.504	0.511	0.508	Singapore	0.600	0.596	0.608	0.602
Estonia	0.657	0.639	0.650	0.599	Slovakia	0.668	0.666	0.656	0.652
Ethiopia	0.379	0.336	0.395	0.429	Slovenia	0.619	0.618	0.614	0.648
Fiji	0.570	0.584	0.543	0.581	Solomon Islands	0.592	0.595	0.603	0.614
Finland	0.697	0.667	0.664	0.628	Somalia	0.393	0.402	0.407	0.407
France	0.589	0.576	0.566	0.570	South Africa	0.504	0.513	0.501	0.506
French Polynesia	0.533	0.543	0.534	0.515	Spain	0.522	0.510	0.533	0.529
Gabon	0.401	0.419	0.441	0.462	Sri Lanka	0.446	0.457	0.454	0.451
Gambia, The	0.529	0.502	0.488	0.494	St. Kitts and Nevis	0.463	0.464	0.435	0.443
Georgia	0.527	0.540	0.551	0.538	St. Lucia	0.486	0.497	0.462	0.466
Germany	0.668	0.677	0.673	0.667	St. Vincent and the Grenadines	0.522	0.514	0.513	0.505
Ghana	0.509	0.520	0.509	0.519	Sudan	0.437	0.419	0.423	0.399
Greece	0.447	0.446	0.435	0.467	Suriname	0.517	0.575	0.569	0.532
Greenland	0.580	0.573	0.561	0.551	Swaziland	0.542	0.546	0.568	0.563
Grenada	0.524	0.517	0.498	0.488	Sweden	0.715	0.697	0.665	0.643
Guatemala	0.480	0.453	0.476	0.456	Switzerland	0.657	0.684	0.681	0.671
Guinea	0.518	0.533	0.517	0.537	Syrian Arab Republic	0.387	0.402	0.416	0.374
Guinea-Bissau	0.545	0.524	0.509	0.524	Taiwan; province of China	0.561	0.580	0.571	0.573
Guyana	0.516	0.527	0.538	0.535	Tajikistan	0.574	0.565	0.525	0.492
Haiti	0.418	0.430	0.444	0.452	Tanzania	0.461	0.426	0.420	0.397
Honduras	0.506	0.491	0.504	0.518	Thailand	0.555	0.514	0.558	0.577
Hong Kong SAR, China	0.410	0.391	0.401	0.399	Timor-Leste		0.463	0.427	0.440
Hungary	0.653	0.624	0.666	0.662	Togo	0.512	0.527	0.526	0.524
Iceland	0.526	0.528	0.584	0.516	Tonga	0.565	0.545	0.534	0.533
India	0.467	0.473	0.472	0.440	Trinidad and Tobago	0.444	0.432	0.430	0.470
Indonesia	0.458	0.453	0.462	0.466	Tunisia	0.475	0.500	0.516	0.517
Iran, Islamic Rep.	0.426	0.434	0.414	0.429	Turkey	0.376	0.423	0.481	0.465
Iraq	0.351	0.337	0.315	0.347	Turkmenistan	0.458	0.480	0.455	0.476

Name	2000	2005	2010	2015	Name	2000	2005	2010	2015
Ireland	0.585	0.588	0.593	0.571	Tuvalu	0.491	0.496	0.518	0.518
Israel	0.591	0.558	0.540	0.532	Uganda	0.492	0.491	0.481	0.480
Italy	0.512	0.516	0.535	0.551	Ukraine	0.606	0.625	0.598	0.588
Jamaica	0.448	0.455	0.458	0.441	United Arab Emirates	0.500	0.460	0.420	0.428
Japan	0.584	0.609	0.604	0.574	United Kingdom	0.564	0.546	0.554	0.573
Jordan	0.506	0.506	0.530	0.525	United States	0.532	0.512	0.516	0.507
Kazakhstan	0.486	0.474	0.455	0.432	Uruguay	0.526	0.529	0.523	0.514
Kenya	0.484	0.513	0.529	0.532	Uzbekistan	0.525	0.534	0.490	0.462
Kiribati	0.520	0.520	0.514	0.529	Vanuatu	0.561	0.593	0.582	0.565
Korea, Dem. Rep.	0.464	0.470	0.471	0.477	Venezuela, RB	0.445	0.410	0.407	0.403
Korea, Rep.	0.585	0.582	0.587	0.593	Vietnam	0.514	0.534	0.556	0.604
Kuwait	0.367	0.383	0.398	0.381	West Bank and Gaza	0.464	0.464	0.443	0.459
Kyrgyz Republic	0.554	0.562	0.561	0.554	Yemen, Rep.	0.357	0.345	0.348	0.334
Lao PDR	0.474	0.501	0.542	0.567	Zambia	0.451	0.454	0.481	0.499
Latvia	0.633	0.581	0.607	0.575	Zimbabwe	0.510	0.462	0.504	0.500

References

Ahluwalia, M. S. (2015) *Role of economists in policy-making*, WIDER working paper 2015/144, Helsinki, United Nations University World Institute for Development Economics.

Baynes, K., & Bohman, J. (1987). *After philosophy: End or transformation?*. Cambridge, MA: MIT Press.

Buuren, S., & Groothuis-Oudshoor, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67.

Coyle, D. (2014). *GDP: A brief but affectionate history*. Princeton, NJ: Princeton University Press.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford: Oxford University Press.

ECOSOC. (2016). *Report of the inter-agency and expert group on sustainable development goal indicators*. United Nations Economic and Social Council. E/CN.3/2016/1.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn world table. *American Economic Review, 105*(10), 3150–3182.

Freedman, D. (2010). Edited by Collier, D., Sekhon, J. S. and Stark, P. B. Statistical models and causal inference: A dialogue with the social sciences. New York, NY: Cambridge University Press.

Gaulier, G., & Zignago, S. (2010). *BACI: International trade database at the product-level. The 1994–2007 Version*. CEPII working paper 2010–23.

Hall, P., Racine, J., & Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association, 99,* 1015–1026.

Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., et al. (2012). An index to assess the health and benefits of the global ocean. *Nature, 488*(7413), 615–620.

Hausmann, R., & Hidalgo, C. A. (2012). *Economic complexity and the future of manufacturing. The future of manufacturing: Opportunities to drive economic growth*. Cologny/Geneva: World Economic Forum.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2011). *Atlas of economic complexity: Mapping paths to prosperity*. Cambridge, MA: Harvard Center for International development.

Henderson, D. J., & Parmeter, C. F. (2015). *Applied nonparametric econometrics*. New York, NY: Cambridge University Press.

Jerven, M. (2014). Benefits and costs of the data for development targets for the post-2015 development Agenda. *Data for development assessment paper working paper*. Copenhagen: Copenhagen Consensus Center.

Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.

Kroll, C. (2015). *Sustainable development goals: Are the rich countries ready?* Gütersloh: Bertelsmann Foundation.

Lin, J. Y. (2012a). *New structural economics: A framework for rethinking development and policy*. Washington DC: World Bank.

Lin, J. Y. (2012b). *The quest for prosperity: How developing economies can take off*. Princeton, NJ: Princeton University Press.

Lin, J. Y., & Monga, C. (2013). Comparative advantage: The silver bullet of industrial policy. In: J. E. Stiglitz & J. Y. Lin (Eds.), *The industrial policy revolution I: The role of government beyond ideology* (pp. 19–39). New York: Palgrave MacMillan.

Lin, J. Y., & Monga, C. (2014). The evolving paradigms of structural chance. In Bruce Currie-Alder, R. Kanbur, D. M. Malone, & R. Medhora (Eds.), *International development: Ideas, experience, and prospects*. New York: Oxford University Press.

McMillan, M. S., & Rodrik, D. (2011). *Globalization, structural change and productivity growth*. No. w17143. National Bureau of Economic Research.

Mlachila, M., Tapsoba, R., & Tapsoba, S. J. A. (2014). A quality of growth index for developing countries: A proposal. *Social Indicators Research*: 1–36.

Monga, C. (2013). Winning the jackpot: Jobs dividends in a multipolar world. In J. E. Stiglitz, J. Y. Lin, & E. Patel (Eds.), *The industrial policy revolution II—Africa in the 21st century* (pp. 135–171). New York: Palgrave MacMillan.

OECD and Joint Research Centre-European Commission. (2008). *Handbook on constructing composite indicators: Methodology and User guide*. OECD Publishing.

Philipsen, D. (2015). *The little big number: How GDP came to rule the world and what to do about it*. Princeton, NJ: Princeton University Press.

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ravallion, M. (2011). *Mashup indexes of development*, Policy Research Working Paper no. 5532, Washington DC, World Bank.

Sachs, J., Schmidt-Traub, G., Kroll, C., Durand-Delacre, D., & Teksoz, K. (2016). *SDG index and dashboards—Global report*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

Sachs, J., Schmidt-Traub, G., Kroll, C., Durand-Delacre, D., & Teksoz, K. (2017). *SDG index and dashboards—Global report*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

SDSN. (2015). *Indicators and a monitoring framework for sustainable development goals: Launching a data revolution for the SDGs*. New York: Sustainable Development Solutions Network.

Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris.

Timmer, C. P., & Akkus, S. (2008). The structural transformation as a pathway out of poverty: Analytics, empirics and politics. *Center for Global Development Working Paper* 150.

UNIDO. (2009). *Industrial development report 2009—Breaking in and moving up: New industrial challenges for the bottom billion and the middle-income countries*. Vienna.

WEF. (2014). *The global competitiveness report 2014–2015*. Vienna: World Economic Forum.