# Policy Research Working Paper

# Missing SDG Gender Indicators

Kathleen Beegle Umar Serajuddin Brian Stacy Divyanshi Wadhwa



Development Data Group & Development Research Group August 2023



# **Reproducible Research Repository**

A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

# Abstract

The Sustainable Development Goal agenda lays out an ambitious set of 231 indicators to track progress. Countries continue to fall short in terms of reporting on the indicators in general, and this is particularly the case for the subset of 50 gender-related indicators, where countries reported on average on 31 percent of these indicators in at least one year from 2016 to 2020. A closer look at this low coverage reveals four salient fundings. First, this is not just a problem of missing data; lack of reporting on existing data is detected to be a problem. For example, of the 32 gender-related indicators that are sex disaggregated, if countries that had a population estimate also had a sex-disaggregated

estimate (which is almost always feasible), the Sustainable Development Goal gender coverage rate would be 43 percent instead of 31 percent. Second, better statistical systems are a major part of the solution, as statistical system strength is correlated with higher coverage. Third, poorer countries are doing no worse in reporting on gender-related Sustainable Development Goal indicators than high-income countries, despite weaker statistical systems. Lastly, sizable over (and under) performance in reporting, conditional on statistical strength, suggests that country-level advocacy and focus can yield wins in Sustainable Development Goal gender indicator coverage.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

This paper is a product of the Development Data Group and Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at kbeegle@worldbank.org. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

## **Missing SDG Gender Indicators**

Kathleen Beegle, Umar Serajuddin, Brian Stacy, Divyanshi Wadhwa\*

Key words: statistical indicators, gender, national statistical system

**JEL**: C8, J16, I00, O1

All authors are with the World Bank. Corresponding author: Kathleen Beegle kbeegle@worldbank.org. The authors are grateful to comments from Hai-Anh Dang, Anna Fruttero, and Lauren Harrison. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

#### 1 Introduction

The Sustainable Development Goals (SDGs) lay out an ambitious agenda including that of achieving gender equality by 2030. This agenda is paired with a set of goals and targets measured by concrete indicators and is adopted by nearly all countries. SDG 5 focuses on gender equality and sets 9 measurable targets (with 14 indicators) on issues that especially affect women and girls (United Nations, 2022). But gender cuts across a far wider range of the SDGs than just the indicators under Goal 5. For example, SDG 3 on ensuring good health and well-being includes a target on reducing maternal mortality (target 3.1). The SDG agenda also calls for sex disaggregated data across several goals where monitoring of gender disparities is essential for effective policy. For example, SDG 8 on promoting decent work and economic growth sets a target of achieving full employment and equal pay for all women and men (target 8.5). These gender data are promoted as key to understanding if and how patterns of progress differ between women and men or girls and boys (UN Women, 2022). Countries are, however, falling short on reporting on gender-related indicators of the SDGs. This paper analyzes the patterns underlying these data gaps.

The objective of this paper is to look at this globally agreed-upon set of gender data indicators and identify key country patterns related to the existence, or lack of, such data. We focus on the availability of data reporting on the SDGs since they represent an internationally agreed-upon set of goals to meet and for countries to report on (UNSD, 2022).

Missing gender data is not a new concern. There are different approaches to diagnosing the causes of the lack of gender data. One approach put forth by Bonfert et al (2022) emphasizes four obstacles to more gender data (Figure 1): (i) lack of data sources such as core and/or specialized surveys, censuses or relevant administrative data (that is, the data simply are not collected); (ii) methodological flaws in data collection (e.g., collecting land holdings of households but not

identifying which household member has the rights/ownership to this land); (iii) insufficient processing of existing data; and (iv) lack of dissemination even when data are available and processed.





Related, but not identical, Buvinic Furst-Nichols, and Koolwal (2014) discuss gender data gaps as driven by four gaps: (i) lack of regular production at the country level; (ii) lack of international standards; (iii) lack of information across domains; (iv) lack of granularity, i.e., lack of large, detailed datasets making possible disaggregation.

In this paper we look at the production and reporting of SDG indicators on gender, clearly laying out the availability (alternatively the lack of) data along indicator types – uniquely gender focused versus cross-cutting. We then focus on the challenges posed by insufficient processing of available data or the lack of dissemination even when processed data and constructed indicators are available. While focusing on improving the statistical systems is an important part of the agenda to fulfill the goal of reporting on gender-related SDGs, some rapid improvements can be made from existing data.

Source: Bonfert et al. 2022

#### 2 Gender indicators for the SDGs

Although nearly all countries have agreed to report on the SDG indicators, major gaps exist in indicator availability since the SDG agenda's inception in 2015 (Dang and Serajuddin 2020). Gender-related SDG indicators are no exception. There are 231 unique SDG indicators. Many of the indicators, even if not obviously related to gender, nonetheless have sub-indicators, such as, indicators by sex, age, or disability status. The UN global SDG indicators database provides access to the data compiled for tracking progress toward fulfilling the SDGs. We use this data source, rather than individual NSO websites, because data submitted to the UN Global SDG monitoring database goes through a standardized process including a certain level of quality control and documentation review.

We explore the coverage of the 50 gender-related SDG indicators out of the 231 unique indicators.<sup>1</sup> As noted earlier, gender-related SDG indicators are not limited to SDG 5 on gender equality, but rather span indicators across 10 out of 17 of the SDG goals. All 50 SDG-gender indicators are Tier 1 or 2 SDG indicators.<sup>2</sup> While most are related to sex disaggregation of data

<sup>&</sup>lt;sup>1</sup> These 50 indicators closely match the UN Women minimum set of 52 quantitative gender indicators from the SDGs (United Nations Economic and Social Council 2012), subsequently revised to be 51 quantitative indicators, with a few exceptions. These exceptions are: (i) indicators 4.7.1, 4.a.1, and 13.3.1 are in the UN Women minimum set but not here as in our view they are not gender-related or sex-disaggregated measures. (ii) indicator 1.1.1 includes the "working poor" (employed population below international poverty line) by sex component which is in Table 1 but not in the UN Women list. Relatedly, Open Data Watch (2019) refers to 32 SDG gender indicators and another 36 "additional" SDG gender indicators. The difference between their 68 and our 50 SDG gender indicators is that some of theirs are, in our assessment, gender neutral in terms of the present drafting of the indicator (such as 1.5.1 Number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population).

 $<sup>^2</sup>$  Tier 1 indicators, according to the UN, are indicators that are conceptually clear, have an internationally established methodology and standards available, and data are regularly produced by countries for at least 50% of countries and of the population in every region where relevant. Tier 2 indicators are conceptually clear and have an internationally established methodology and set of standards, but are not regularly produced by countries. There are no gender equality indicators in the third category, Tier 3, which is defined as an indicator with no internationally established methodology or standards established, and, thus, these indicators are likely to have the lowest rate of coverage. (UNSD, 2022).

(32 of the 50), the 18 others are related to goals specific to females– highlighting that genderrelated SDG indicators are not only about sex disaggregation.

Table 1 shows the share of countries for which there is at least one annual data point in the five-year period from 2016 to 2020 for each of the 50 gender-related indicators.<sup>3,4</sup> The average coverage rate of indicators is around 34% for 181 countries; that is, an average country will have data reported in the SDG website for about 17 out of 50 indicators. Over 90% of the world's population lives in a country where less than half of the 50 SDG gender indicators are available for any year in this 5-year period. Indeed, the gender-related SDG indicators are more likely to be unreported than other indicators. For the overall 181 SDG indicators, the average reporting rate is 65% for this same period.<sup>5</sup>

Next, we unpack several notable aspects of the availability (or lack of) SDG gender indicators. For Tier 1 SDG indicators (18 out of 50), arguably those that will or should have the greatest availability, availability is much higher; countries have a recent value for only about half (51%) of the indicators. For Tier 2 indicators (32 out of 50), the average country has a recent value

<sup>&</sup>lt;sup>3</sup> Countries with populations of less than 200,000 (34 out of 215 countries in the UN SDG Database) were excluded from this analysis. These are: American Samoa, Andorra, Antigua and Barbuda, Aruba, Bermuda, British Virgin Islands, Cayman Islands, Channel Islands, Curacao, Dominica, Faroe Islands, Gibraltar, Greenland, Grenada, Guam, Isle of Man, Kiribati, Liechtenstein, Marshall Islands, Micronesia, Fed. Sts., Monaco, Nauru, Northern Mariana Islands, Palau, San Marino, Seychelles, Sint Maarten (Dutch part), St. Kitts and Nevis, St. Lucia, St. Martin (French part), St. Vincent and the Grenadines, Tonga, Turks and Caicos Islands, Tuvalu, Virgin Islands (U.S.). Encarnacion et al (2022) note that the poorest performers in terms of lowest SDG-gender indicators are small islands and nations. These countries have, on average, very low reporting rates for SDGs, including, but not only, those related to gender. In general, small islands and nations are under-performers in terms of statistical performance conditional on their income and human captial index level (Dang et al 2021).

<sup>&</sup>lt;sup>4</sup> The SDG database includes actual (survey/census or other primary data source estimates) as well as additional modeled estimates for indicators when primary sources are not available for the country. We do not use modeled estimates.

 $<sup>^{5}</sup>$  This is our own calculation. Dang and Serajuddin (2020) report lower rates of SDG indicator reporting in part because they focus on an earlier period (2012-2016) and because they include small islands and nations.

for only a quarter (24%) of the indicators. Annex 1 presents the availability of Tier 1 and Tier 2 SDG gender-related indicators by region and by country income grouping.

For the 14 indicators under Goal 5, the country average availability is 37%, only marginally higher availability compared with the average availability for all gender-related indicators. Figure 2 shows this distribution. No country has more than 10 of these 14 indicators in the 5-year period. Forty-one countries report three or fewer indicators. Annex 1 presents the availability of SDG5 gender-related indicators by region and country income grouping.



Figure 2. Availability of SDG5 gender-related indicators (N=181 countries)

Note: The figure shows the coverage of the 14 SDG 5 indicators where coverage is defined as having at least one annual data point in the five-year period from 2016 to 2020 for an indicator as compiled by the UN.

Among indicators that require sex disaggregation (32 out of the 50), both the population data and the sex-disaggregated data are not reported for any country for five indicators (such as for indicator 10.2.1). For four of these 32 indicators, the sex-disaggregated and population coverage rates match, as we would expect if the underlying data identified individual sex, was

collected for both males and females, and was processed accordingly. We would not expect the sex-disaggregated coverage rate to exceed the population coverage, and it never does. Moreover, if the country has a sex-disaggregated data point they also have a population estimate. But the reverse is not the case. For six of these 32 indicators (19%), while there is some reported data for the population indicator, there is no sex-disaggregated data reported. As an example, 56% of countries report a population rate for SDG indicator 10.2.1 (the proportion of people living below 50 percent of median income), yet no country reports this statistic by sex. These missing data are not the result of missing sex in underlying data source (in this case, household surveys). The measure itself (living below an income threshold) is defined at the household level and so one can produce a sex-disaggregated estimate based on the households in which individuals reside. For example, Munoz Boudet et al. (2021) report poverty rates by sex.<sup>6</sup>

In the remaining 17 cases (out of 32), where there is some reported data for both the population and by sex-disaggregation, in a handful of cases there are large gaps between the percentage of countries with a recent value by sex and those reporting a population estimate (i.e. comparing the last two columns in Table 1 when both columns are non-zero). For SDG indicator 1.3.1, on the proportion of population covered by social protection floors/systems, 79% of countries have a recent value for the population, but only 8% of countries have a sex disaggregated data point. A less drastic example is SDG indicator 4.1.1, related to early childhood education: 64 percent of countries have a population estimate for this indicator but only 53 percent have an estimate by sex.

<sup>&</sup>lt;sup>6</sup> This is not related to the issue of measuring income or poverty at the individual level versus a household measure. It is simply the point that if a population estimate is produced based on a household-level measure, then there is no methodological argument against producing additional estimates for population sub-categories (be it urban/rural, male/female, for children, etc).

These results show that the problem of missing SDG gender measures is, in part, a problem with processing existing data rather than the lack of the primary data collection, since the underlying sources (typically household surveys) almost always (if not always) collect the sex of household members. If countries with a population estimate also reported data by sex, the SDG gender coverage rate would rise from 31% to 43%.

Evidence from other sources underscore the problem of available data not being reported. In a review of national statistics for 12 countries related to sex-disaggregated data on asset ownership, employment, and entrepreneurship, Bonfert et al (2023) find that such data exist but are not made available on average for 9 out of 24 indicators. Examining gender statistics from the UK, when comparing the Office of National Statistics (ONS) website for gender data to the reporting in the UN SDG site, we find that about 31% of gender-related SDG indicators (16 out of 50) are on the ONS website but not in the UN SDG site. On the other hand, only 4 of the 50 are in the UN SDG site but not on the ONS website. The remaining 30 SDG gender indicators are in both (21) or neither (9).

Turning to regional patterns in coverage, South Asia has the highest coverage rate of SDG gender indicators at 36% availability, about 5% higher than the global average (31%). There are no clear patterns in terms of which groups or types of indicators have higher or lower indicator coverage by region.

Notably, the high-income countries do not have higher coverage of gender-related SDG indicators (Figure 4). GDP per capita is not associated with better coverage of gender statistics in the UN SDG database (Figure 5). This is also the case for SDG indicators overall: high income countries do not have higher rates of reporting of all 181 SDG indicators (rates are high income 64%, upper middle income 72%, lower middle income 70%, and low income 65%). Yet, high

income countries perform notably better on the Statistical Performance Indicators and Index (SPI) – the World Bank's new official tool to measure country statistical capacity and is being added to the SDG indicators under SDG17 (Dang et al. 2023).

One explanation for this paradox on reporting SDGs and statistical strength overall is that richer countries may have been slow to report SDGs compared with low- and middle-income countries that have experience in engaging with the Millennium Development Goals (MDGs) (MacFeely 2018). A second explanation lies in the specific focus of some SDGs. This difference in the overall performance of national statistical systems and the reporting on gender SDGs might be explained by presence in the latter of indicators which relate to phenomena that are arguably infrequent or rare for high income countries (or perceived as such). For example, data on child marriage and on female genital mutilation (covered in SDG target 5.3) are rarely collected in OECD countries. OECD (2022) describes the extra lengths needed to get such data from alternate sources in order to be able to report on this SDG.

A third explanation, related to the second one above, is the presence of systematic and large-scale data collection under the Demographic and Health Survey (DHS) and the Multiple Indicator Cluster Survey (MICS) programs, which are focused on low-income countries (and often financed with non-national resources). These surveys are often the source of gender-related data, especially in the domains of female health and empowerment. To assess this, we examine the main data sources for the 50 SDG gender indicators. The DHS or MICs is the source for at least one country data point for 13 out of 50 indicators but only extensively (well over half of the data points) for 5 indicators.<sup>7</sup> We find very slight evidence that the DHS/MICs data source explains lower

<sup>&</sup>lt;sup>7</sup> For the 13 indicators where the DHS/MICs is ever the source, the share of data points which are from DHS/MICs are: 2% (indicator 4.1.1), 4% (indicator 5.b.1), 14% (indicator 5.a.1), 16% (indicator 3.7.2), 24% (indicator3.1.2),

coverage of gender-related SDGs in high-income countries relative to lower income countries. When excluding these 5 DHS/MICs-dominant indicators, high-income countries have basically the same coverage (31%) as low (22%) and lower middle (31%) countries. And they continue to lag behind upper middle income countries (35.4%).



Figure 3. Availability of SDG gender indicators by region (N=181 countries)

The 25th percentile of the SDG gender availability is 26%. The median is 32% and the 75th percentile is 38%. Means are unweighted. Circles are sized according to population

<sup>36% (</sup>indicator 1.4.2), 36% (indicator 4.5.1), 41% (indicator 3.7.1), 86% (indicator 16.2.3), 93% (indicator 5.3.1), 97% (indicator 5.6.1), 97% (indicator 4.2.1), and 98% (indicator 5.3.2).



Figure 4. Availability of SDG gender indicators by income group (N=181 countries)

The 25th percentile of the SDG gender availability is 26%. The median is 32% and the 75th percentile is 38%. Means are unweighted. Circles are sized according to population



Figure 5. GDP per capita and SDG gender indicators (n=181 economies)

Source: GDP per capita from the World Development Indicatrs (NY.GDP.PCAP.PP.CD). The availability of Gender SDG Indicators based on authors calculations based on the UN Global SDG Database.

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
1	1.1.1	1	Proportion of the population living below the international poverty line by sex, age, employment status and geographic location (urban/rural)		0.0%	61.3%
	1.1.1	1	Proportion of the population living below the international poverty line by sex, age, employment status and geographic location (urban/rural)	Employed population below international poverty line, by sex and age (%)	0.0%	14.9%
	1.2.1	1	Proportion of population living below the national poverty line, by sex and age		0.0%	53.6%
	1.2.2	2	Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions (multidimensional poverty)		21.5%	28.2%
	1.3.1	1	Proportion of population covered by social protection floors/systems, by sex, distinguishing children, unemployed persons, older persons, persons with disabilities, pregnant women, newborns, work-injury victims and the poor and the vulnerable	Proportion of population covered by at least one social protection benefit, by sex (%)	9.4%	86.7%

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	1.4.2	2	Proportion of total adult population with secure tenure rights to land, (a) with legally recognized documentation, and (b) who perceive their rights to land as secure, by sex and type of tenure	Any data: Proportion of people with legally recognized documentation of their rights to land out of total adult population, by sex (%); AND/OR Proportion of people who perceive their rights to land as secure out of total adult population, by sex (%); AND/OR Proportion of people with secure tenure rights to land out of total adult population, by sex (%)	12.7%	13.3%
2	2.2.3	1	Prevalence of anaemia in women aged 15 to 49 years, by pregnancy status		NA	65.2%
	2.3.2	2	Average income of small-scale food producers, by sex and indigenous status		5.0%	5.0%
	3.1.1	1	Maternal mortality		NA	97.8%
	3.1.2	1	Proportion of births attended by skilled health personnel		NA	81.2%
3	3.3.1	1	New HIV infections in uninfected population, by sex, age and key populations		63.5%	78.5%
	3.7.1	1	Proportion of women of reproductive age (aged 15-49 years) who have their need for family planning satisfied with modern methods		NA	46.4%
	3.7.2	1	Adolescent birth rate (aged 10-14 years; aged 15-19 years)		NA	82.9%

Table 1. SDG indicators related to	o gender (N=181 countries)
------------------------------------	----------------------------

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
4	4.1.1	1	Proportion of children and young people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex	Proportion of children and young people achieving a minimum proficiency level in reading and mathematics (%)	62.4%	75.1%
	4.2.1	2	Proportion of children aged 24-59 months who are developmentally on track in health, learning and psychosocial well-being, by sex	Proportion of children aged 36–59 months who are developmentally on track in at least three of the following domains: literacy-numeracy, physical development, social-emotional development, and learning	28.7%	29.8%
	4.2.2	1	Participation rate in organized learning (one year before the official primary entry age), by sex		77.9%	79.0%
	4.3.1	2	Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex		38.1%	39.8%

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	4.5.1	2	Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated	Any data: Adjusted gender parity index for participation rate in organized learning (one year before the official primary entry age), (ratio); AND/OR Adjusted gender parity index for the proportion of teachers with the minimum required qualifications, by education level (ratio); AND/OR Adjusted gender parity index for participation rate in formal and non-formal education and training (ratio); AND/OR Gender parity index for youth/adults with information and communications technology (ICT) skills, by type of skill (ratio); AND/OR Adjusted gender parity index for achieving a minimum proficiency level in reading and mathematics (ratio); AND/OR Adjusted gender parity index for completion rate, by location, wealth quintile and education level; AND/OR Adjusted gender parity index for achieving at least a fixed level of proficiency in functional skills, by numeracy/literacy skills (ratio)	NA	95.0%
	4.6.1	2	Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex		7.7%	8.8%

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
5	5.1.1	2	Whether or not legal frameworks are in place to promote, enforce and monitor equality and non-discrimination on the basis of sex	Any data: Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement, $0 - 100$ ) – Area 1: overarching legal frameworks and public life; AND/OR Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement, $0 - 100$ ) – Area 2: violence against women; AND/OR Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement, $0 - 100$ ) – Area 3: AND/OR mployment and economic benefits; AND/OR Legal frameworks that promote, enforce and monitor gender equality (percentage of achievement, $0 - 100$ ) – Area 4: marriage and family	NA	52.5%
	5.2.1	1	Proportion of ever-partnered women and girls aged 15 years and older subjected to physical, sexual or psychological violence by a current or former intimate partner in the previous 12 months, by form of violence and by age		NA	81.8%
	5.2.2	2	5.2.2 Proportion of women and girls aged 15 years and older subjected to sexual violence by persons other than an intimate partner in the previous 12 months, by age and place of occurrence		NA	0.0%
	5.3.1	1	Proportion of women aged 20-24 years who were married or in a union before age 15 and before age 18	Any data: Proportion of women aged 20-24 years who were married or in a union before age 18 (%); AND/OR Proportion of women aged 20-24 years who were married or in a union before age 15 (%)	NA	47.0%

Goal	Indicator	Tier Description		Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	5.3.2	1	Proportion of girls and women aged 15-49 years who have undergone female genital mutilation/cutting, by age		NA	11.0%
	5.4.1	2	Proportion of time spent on unpaid domestic and care work, by sex, age and location		17.1%	17.7%
	5.5.1	1	Proportion of seats held by women in (a) national parliaments and (b) local governments	Any data: Proportion of seats held by women in national parliaments (% of total number of seats); AND/OR Proportion of elected seats held by women in deliberative bodies of local government (%)	NA	98.3%
	5.5.2	1 Proportion of women in managerial positions		Proportion of women in senior and middle management positions (%)	NA	50.3%
	5.6.1	2	Proportion of women aged 15-49 years who make their own informed decisions regarding sexual relations, contraceptive use and reproductive health care		NA	21.0%
	5.6.2	2	Number of countries with laws and regulations that guarantee full and equal access to women and men aged 15 years and older to sexual and reproductive health care, information and education		NA	40.9%
	5.a.1	2	(a) Proportion of total agricultural population with ownership or secure rights over agricultural land, by sex; and (b) share of women among owners or rights-bearers of agricultural land, by type of tenure	Any data: Proportion of people with ownership or secure rights over agricultural land (out of total agricultural population), by sex (%); AND/OR Share of women among owners or rights-bearers of agricultural land, by type of tenure (%)	14.9%	15.5%

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	5.a.2	2	Proportion of countries where the legal framework (including customary law) guarantees women's equal rights to land ownership and/or control		NA	19.9%
	5.b.1 2 Proportion of individuals who own a mobile telephone, by sex		Proportion of individuals who own a mobile telephone, by sex		48.1%	61.3%
	5.c.1	2	Proportion of countries with systems to track and make public allocations for gender equality and women's empowerment		NA	32.6%
	8.3.1	2	Proportion of informal employment in total employment, by sector and sex		37.6%	37.6%
	8.5.1	2	Average hourly earnings of employees, by sex, age, occupation and persons with disabilities		44.8%	45.3%
0	8.5.2	1	Unemployment rate, by sex, age and persons with disabilities		74.0%	75.1%
8	8.7.1	2	Proportion and number of children aged 5- 17 years engaged in child labour, by sex and age	Proportion of children engaged in economic activity, by sex and age (%)	30.4%	30.4%
	8.8.1	2	Fatal and non-fatal occupational injuries per 100,000 workers, by sex and migrant status	Any data: Fatal occupational injuries among employees, by sex and migrant status (per 100,000 employees); AND/OR Non-fatal occupational injuries among employees, by sex and migrant status (per 100,000 employees)	30.9%	38.7%

Goal	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	8.8.2	2	Level of national compliance with labour rights (freedom of association and collective bargaining) based on International Labour Organization (ILO) textual sources and national legislation, by sex and migrant status		0.0%	70.2%
10	10.2.1	2	Proportion of people living below 50 percent of median income, by sex, age and persons with disabilities		0.0%	56.4%
	11.2.1	2	11.2.1 Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities		0.0%	0.0%
11	11.7.1	2	11.7.1 Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities		0.0%	0.0%
	11.7.2	2	11.7.2 Proportion of persons victim of physical or sexual harassment, by sex, age, disability status and place of occurrence, in the previous 12 months		0.0%	0.0%
	16.1.1	2	Number of victims of intentional homicide per 100,000 population, by sex and age		55.8%	57.5%
16	16.1.2	2	Conflict-related deaths per 100,000 population, by sex, age and cause	· · · · · · · · · · · · · · · · · · ·	0.0%	0.0%

al	Indicator	Tier	Description	Sub-Indicator (if any)	Sex Disaggregation Available	Any Data Available
	16.1.3	2	Proportion of population subjected to (a) physical violence, (b) psychological violence and (c) sexual violence in the previous 12 months	Any data: Proportion of population subjected to physical violence in the previous 12 months, by sex (%); AND/OR Proportion of population subjected to robbery in the previous 12 months, by sex (%); AND/OR Proportion of population subjected to sexual violence in the previous 12 months, by sex (%)	9.4%	19.9%
	16.2.2	2	Number of victims of human trafficking per 100,000 population, by sex, age and form of exploitation		0.0%	59.1%
	16.2.3	2	Proportion of young women and men aged 18- 29 years who experienced sexual violence by age 18		19.9%	19.9%
	16.7.1	2	Proportions of positions in national and local institutions, including (a) the legislatures; (b) the public service; and (c) the judiciary, compared to national distributions, by sex, age, persons with disabilities and population groups	Any data: Ratio for female members of parliamen (Ratio of the proportion of women in parliament i proportion of women in the national population w the age of eligibility as a lower bound boundary), Lower Chamber or Unicameral; AND/OR Ratio f female members of parliaments (Ratio of the proportion of women in parliament in the proporti women in the national population with the age of eligibility as a lower bound boundary), Upper Chamber	NA	0.0%
	16.7.2	2	16.7.2 Proportion of population who believe decision-making is inclusive and responsive, by sex, age, disability and population group		0.0%	0.0%

Note: Percent of countries with any reporting on the indicator in the five years (2016-2020). NA indicates that the indicator is not relevant in regards to sex disaggregation. Source: UN SDG Global Database. https://unstats.un.org/sdgs/dataportal

#### 3. SDG gender indicator availability and country statistical performance

Next we assess how a country's overall statistical performance relates to the availability of gender data, and identify countries that may have strong systems overall but are underperforming on gender statistics. To do so, we compare the availability of gender statistics to scores on the World Bank's Statistical Performance Indicators (SPI) (Dang et al 2023).

The World Bank's Statistical Performance Indicators (SPI) measure statistical performance for 174 countries covering over 99% of the world population. The indicators are grouped into five pillars: (1) data use, which captures the demand side of the statistical system; (2) data services, which looks at the interaction between data supply and demand such as the openness of data and quality of data releases; (3) data products, which reviews whether countries report on global indicators;<sup>8</sup> (4) data sources, which assesses whether censuses, surveys, and other data sources are created; and (5) data infrastructure, which captures whether foundations such as financing, skills, and governance needed for a strong statistical system are in place. Within each pillar is a set of dimensions, and under each dimension is a set of indicators to measure performance. The indicators going back to 2004.<sup>9</sup> The indicators are summarized as an index, termed the SPI overall score, with scores ranging from a low of 0 to a high of 100. We use the SPI data for 2019.

<sup>&</sup>lt;sup>8</sup> The data products pillar measures whether countries have recent SDG indicators across the 17 goals available in the UN Global SDG Indicators database.

<sup>&</sup>lt;sup>9</sup> The data for the indicators are from a variety of sources, including databases produced by the World Bank, International Monetary Fund (IMF), United Nations (UN), Partnership in Statistics for Development in the 21st Century (PARIS21), and Open Data Watch—and in some cases, directly from national statistical office websites.

There is a positive relationship between countries' SPI overall scores and the availability of SDG gender indicators (Figure 6). For pillar 3, which is overall SDG coverage, likewise there is a positive relationship with SDG gender indicators (Figure 7). This is not surprising since pillar 3 encompasses the SDG gender indicators itself. Interestingly, although SPI and the SDG gender indicator coverage are positively correlated and SPI is positively correlated with country income level (not shown), as noted earlier, SDG gender indicator coverage is slightly negatively correlated with country income. A breakdown of correlations is reported in Annex 2.

It is among the countries in the poorest quintile of statistical system scoring where the gap in gender data availability is largest. Among the countries within 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, or top quintile of the SPI score, the mean gender-related SDG indicators availability is between 17 and 20. However, for the countries in the bottom quintile of the SPI score, only 12 gender indicators out of 50 are available on average.



Figure 6. SDG gender indicators and SPI score (n=161 economies)

Source: SPI Overall Score from the World Development Indicators (IQ.SPI.OVRL). The availability of Gender SDG Indicators based on authors calculations based on the UN Global SDG Database.



Figure 7. SDG gender indicators and all SDG indicators (n=161 economies)

Next, we explore patterns of over- and under-performing countries in gender data availability as compared to what is expected given a country's SPI score. Specifically, we regress SDG gender indicator availability on SPI score. We take the residual from predicted values as a measure of over and under performance, conditional on SPI. We convert the residual by multiplying it by 50 (the number of SDG gender indicators) to get an estimate of the additional (or fewer) SDG gender indicators a country produces as compared to the predicted number. We also report the difference from the predicted value as share of the predicted value (which can be negative or positive). Figure 8 shows the count and Figure 9 shows the percentage change for top and bottom 15 countries.

For example, Serbia reported on 58% of the SDG gender indicators. Based on the SPI overall score of the country, it was expected to produce only 40% of the SDG gender indicators.

When we multiple this 18 percentage point difference by the total number of SDG gender indicators (50 indicators), we get the additional SDG indicators that were reported from what would be predicted (around 9 indicators).





Over and under performers calculated for Availability of Gender Indicators by using OLS regression of Availability on SPI overall score in 2019 and calculting residuals from this regression.





As discussed above, a simple OLS regression of the availability of gender SDG indicators on the SPI overall score indicates that around 40 percent of the variation in the availability of gender SDG indicators is explained by the statistical performance of a country, leaving around 60 percent unexplained. Many factors beyond SPI might be correlated with the availability of gender SDG indicators, including a country's region, its income level, population size, or the level of female empowerment. Table 2 reports on regressions controlling for some additional country-level traits. These include three measures of gender equality: World Bank Women, Business and the Law index (World Bank 2019), OECD's Social Institutions and Gender Index (SIGI) (OECD 2019), and the Gender Inequality Index (GII) (UN 2020). Controlling for region, the correlations between SDG gender indicators and both SPI and GDP per capita still hold. There is some indication that countries with worse gender inequality measures fare better in terms of SDG gender indicator reporting.

rubie in Regrebbion e	i u unuom	<i>ij</i> or <i>b</i> <b>2 c</b> <i>g</i>	,	one countr	y traits	
	(1)	(2)	(3)	(5)	(6)	(7)
SPI Overall Score	0.20***	-	0.43***	0.50***	0.57***	0.60***
	(0.04)		(0.06)	(0.06)	(0.06)	(0.08)
Log GDP per capita			-5.40***	-5.05***	-4.35***	-4.19**
			(0.64)	(0.75)	(1.08)	(1.26)
Pop under 1.5 million (0/1)			-2.95	-3.22	-0.60	-1.23
			(2.10)	(2.08)	(2.06)	(2.92)
WBL			0.04	0.02	-0.02	-0.17+
			(0.04)	(0.05)	(0.06)	(0.10)
UNDP GII					15.54*	
					(7.33)	
OECD SIGI					. ,	0.10
						(0.09)
<i>Region (0/1):</i>						
Middle East/North Africa		-5.59*		0.95	1.52	-0.63
		(2.76)		(2.28)	(2.26)	(3.34)
North America		-0.86		-4.96	-2.82	-4.69
		(4.80)		(5.33)	(4.65)	(4.99)
South Asia		7.14*		8.26*	6.18+	3.04
		(3.33)		(3.25)	(3.38)	(3.88)
Sub-Saharan Africa		3.52		3.33	1.58	4.08
		(2.56)		(2.40)	(2.47)	(3.31)
Ν	163	163	163	163	140	101
R Sq.	0.13	0.15	0.42	0.51	0.51	0.41

Table 2. Regression of availability of SDG gender indicators on country traits

Notes: \*\*\* indicates statistical significance at 1%, \* at 5% and + at 10%. Constant term included.

The WBL has a range of 1-100, with a high score indicating more gender equal laws and regulations. The OECD SIGI SIGI measures the extent of gender discriminatory legislation and restrictive social norms and practices, where a high score indicates greater gender inequality. Likewise, the GII measures poor outcomes for women in regards to reproductive health, empowerment, and the labor market, and a high score indicates greater gender inequality.

#### 4. Discussion

The world's Sustainable Development Goal (SDG) agenda lays out an ambitious set of indicators to track progress. While the overall coverage of the 231 indicators certainly needs

improvement, the coverage for the 50 gender-related indicators is especially low. On average, countries have 30% of these indicators for at least one year between 2016 and 2020, compared to a rate of 65% for all 181 SDG indicators. Moreover, the subset of 14 indicators under the specific Goal 5 on gender equality fare only slightly better. This low coverage is not a problem of ill-defined indicators. These 50 indicators are classified as either Tier 1 (18 indicators) or Tier 2 (32 indicators) in terms of statistical complexity; so the methodology to collect such data is established.

Clearly, the world needs more reporting on gender-relevant indicators, but how much of this problem is one of lack of data (i.e. no survey exists) versus a failure in the reporting process? For 23 indicators, a lack of data reporting seems to be a cause of missing SDG gender indicators. In these instances, population estimates are being reported, but the sex-disaggregated counterpart is not reported to the same degree though this disaggregation, if not always, is nearly always feasible. This gap in SDG gender indicator reporting seems to be low-hanging fruit, addressed by ensuring that sex-disaggregated information is processed and reported.

For the other indicators, we cannot as easily disentangle if data reporting is the source of the problem or, rather, the lack of relevant surveys/administrative data. For seven indicators, there are no countries with any data point in the five-year period. Statistical system strength as measured by the SPI score is positively correlated with SDG gender data availability; better statistical systems are an important part of the solution. Still, when assessing the performance of gender data availability by a country income level, poor countries are not doing worse despite often weaker statistical systems. The SDG agenda was set in a way that all countries, irrespective of their development status or income level, were to report their progress on all targets, which is in contrast with the MDGs era, which was largely focused on low- and middle-income countries. High-income countries may have been slow to adjust to this shift in the agenda and have, therefore,

underreported statistics (MacFeely, 2018). Certainly there may also be cases where they do not collect certain statistics because of the lack of relevance to their country contexts, as noted in the case of child marriage and female genital mutilation, but we do not find evidence to support this as a driving factor of the result that poorer countries do as well as high-income countries in reporting SDG gender-related indicators.

Beyond these factors, we are left with unexplained variation in gender data availability across countries. This is partly captured in the notable over (and under) performance in reporting gender-related SDGs relative to statistical system strength. One can take a somewhat optimistic perspective in combining this with the two previous findings – that some portion of under-reporting is not driven by lack of data but by under-reporting, and that country income is not driving higher rates of reporting. Even without major investments in statistical systems or the years it may take for such investments to yield results, with some concerted effort, it is possible to achieve big wins in SDG gender indicator coverage.

Meanwhile, it is important to note that while the SDG framework offers the world a consensus set of indicators selected as part of a global consultative process, there are other important country-level gender-related indicators available outside the SDG system. Sources such as the UN Women Data Hub and the World Bank's Gender Data Portal offer compilations of national statistics produced by countries and curated by international agencies.

And lastly, this paper might lay some foundation to develop a systematic and comprehensive approach to tracking country statistical performance regarding gender data. The recently developed Statistical Performance Indicators (SPI) provides a conceptual framework to track the progress of country data and statistical systems, with 5 pillars on data use, data services, data products, data sources, and data infrastructure (Dang et al. 2023). A "Gender SPI" could build from the SPI to focus on specific areas related to gender data production and dissemination, and serve as a means of identifying progress or stalls in national statistical systems with regard to gender data. Such an index would aid in tracking and prioritization of investments by countries and development partners to close gender data gaps.

# References

Bonfert, Anna Tabitha, Sarah Bunker, Kiran Corea, Heather Moylan, and Kolobadia Nayihouba. 2023. "How to assess gender data gaps in the economic domain." World Bank Data Blog (March 1, 2023)

Bonfert, Anna Tabitha, Talip Kilic, Heather Moylan, and Miriam Muller. 2022. "Three ways to tackle gender data gaps – and 12 countries embracing the challenge." World Bank Data Blog, February 7, 2022.

Buvinic, Mayra, Rebecca Furst-Nichols, and Gayatri Koolwal. 2014. "Mapping Gender Data Gaps."

Dang, Hai-Anh H., John Pullinger, Umar Serajuddin, and Brian Stacy. 2023. "Statistical Performance Indicators and Index: A New Tool to Measure Country Statistical Capacity." *Scientific Data* 10:146.

Dang, Hai-Anh H, and Umar Serajuddin. 2020. "Tracking the Sustainable Development Goals: Emerging Measurement Challenges and Further Reflections." *World Development* 127: 104570.

Devarajan, S. 2013. Africa's Statistical Tragedy. *Review of Income and Wealth, Series 59, Special Issue*.

Encarnacion, Jessamyn, Ramya Emandi, and Papa Seck. 2022. "It will take 22 years to close SDG gender data gaps." UN Women Research Highlight (September 2022).

Inter-agency and Expert Group on Gender Statistics (IAEG-GS). 2017. The United Nations Minimum Set of Gender Indicators. United Nations Statistics Division, 6 June 2017. https://genderstats.un.org/files/Minimum%20Set%20indicators%20web.pdf.

MacFeely, S. 2018. The 2030 Agenda: An Unprecedented Statistical Challenge. *Friedrich Ebert Stiftung International Policy Analysis*.

Munoz Boudet, Ana Maria, Antra Bhatt, Ginette Azcona, Jayne Yoo, and Kathleen Beegle, Kathleen. 2021. "A Global View of Poverty, Gender, and Household Composition." World Bank Policy Research Working Paper 9553. World Bank, Washington, DC.

OECD 2019. Social Institutions and Gender Index, SIGI 2019 Global Report: Transforming Challenges into Opportunities.

OECD. 2022. The Short and Winding Road to 2030: Measuring Distance to the SDG Targets. OECD Publishing, Paris.

Open Data Watch. 2019. Bridging the Gap: Mapping Gender Data Availability in Africa.

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N., & Dabalen, A. 2015. "Data Deprivation: Another Deprivation to End." World Bank Policy Research Working Paper No 7252.

United Nations. 2020. Human Development Report 2020.

United Nations. 2021. The Sustainable Development Goals Report 2020.

United Nations. 2022. Achieve gender equality and empower all women and girls.

United Nations Economic and Social Council. 2012. "Gender statistics: Report of the Secretary-General."

United Nations Statistics Division (UNSD), Sustainable Development Goals. https://unstats.un.org/sdgs/dataportal

UNSD. 2018. Gender-relevant SDG indicators (80 indicators).

UNSD. 2022. Inter-agency and Expert Group on SDG Indicators.

UNSD. 2022. Tier Classification of Global SDG Indicators.

UN Women. 2022. SDG Monitoring.

World Bank. 2022. Women, Business, and the Law.





Figure A1.1. Availability of Tier 1 SDG gender indicators by region (N=181 countries)

The 25th percentile of the SDG Tier 1 gender availability is 44%. The median is 50% and the 75th percentile is 61%.

Figure A1.2. Availability of Tier 1 SDG gender indicators by income group (N=181 countries)





Figure A1.3. Availability of Tier 2 SDG gender indicators by region (N=181 countries)







Figure A1.5. Availability of SDG5 gender indicators by region (N=181 countries)

Figure A1.6. Availability of SDG5 gender indicators by income (N=181 countries)



The 25th percentile of the SDG gender availability is 29%. The median is 36% and the 75th percentile is 50%.

# Annex 2

## GDP per capita correlation with:

- Gender SDG Availability: -0.17\*
- SPI Overall Score: 0.67\*\*\*
- SPI Pillar 1 (Data use) Score: 0.42\*\*\*
- SPI Pillar 2 (Data services) Score: 0.53\*\*\*
- SPI Pillar 3 (Data products [Overall SDG Indicator availability]) Score: 0.05
- SPI Pillar 4 (Data sources) Score: 0.74\*\*\*
- SPI Pillar 5 (Data infrastructure) Score: 0.7\*\*\*

# Gender SDG Availability correlation with:

- SPI Overall Score: 0.48\*\*\*
- SPI Pillar 1 (Data use) Score: 0.53\*\*\*
- SPI Pillar 2 (Data services) Score: 0.44\*\*\*
- SPI Pillar 3 (Data products [Overall SDG Indicator availability]) Score: 0.69\*\*\*
- SPI Pillar 4 (Data sources) Score: 0.34\*\*\*
- SPI Pillar 5 (Data infrastructure) Score: 0.25\*\*\*

Note: + indicates statistical significance at 10% level and \*\*\* at the .1% level.