**Social Sustainability Global Database 2.0: A leap forward in data disaggregation**

**Technical Note**

Paola Ballon and Omar Alburqueque

*Social Sustainability and Inclusion Global Practice*

June 2024

Contents

[**1.** **Background** 1](#_Toc163821652)

[**1.1.** **The Social Sustainability Framework** 1](#_Toc163821653)

[**1.2.** **The Importance of Social Sustainability** 2](#_Toc163821654)

[**1.3.** **The Data Gap in Social Sustainability** 2](#_Toc163821655)

[**2.** **The Social Sustainability Global Database (SSGD) v2.0** 3](#_Toc163821656)

[**2.1.** **What is the SSGD v2.0?** 3](#_Toc163821657)

[**2.2.** **SSGD v2.0 features** 6](#_Toc163821658)

[**2.3.** **A descriptive view of the SSGD v2.0** 7](#_Toc163821659)

[**3.** **SSGD v2.0 data visualization tools** 16](#_Toc163821660)

[**4.** **An application exercise using the SSGD v2.0** 19](#_Toc163821661)

[**4.1.** **A country profile for Peru** 19](#_Toc163821662)

[**4.2.** **Comparative assessment: Peru and other Andean countries** 22](#_Toc163821663)

[**4.3.** **Regional analysis: Associations with trust in people in LAC** 23](#_Toc163821664)

[**4.4.** **Subnational analysis: Trust in people in Andean countries** 24](#_Toc163821665)

[**5.** **Use and replicability of the SSGD v2.0** 26](#_Toc163821666)

[**5.1.** **Standardization processes and general guidelines in the use of the SSGD v2.0** 26](#_Toc163821667)

[**5.1.1.** **Wide data format** 26](#_Toc163821668)

[**5.1.2.** **Long data format** 27](#_Toc163821669)

[**5.2.** **Automation processes and replicability of the SSGD v2.0** 28](#_Toc163821670)

[**6.** **Further improvements** 33](#_Toc163821671)

[**7.** **Appendix** 34](#_Toc163821672)

1. **Background**
   1. **The Social Sustainability Framework**

Social Sustainability refers to the principle that every individual is included in the development process, with the expectation that both they and future generations will see benefits from such development. According to Barron et al. (2023), for a community or society to be considered socially sustainable, it must demonstrate the capacity to collaboratively address challenges, effectively deliver public goods, and distribute resources in a manner deemed fair and legitimate by all parties involved, thereby ensuring the well-being of all its members over time.

This concept is built on four components: social cohesion, inclusion, resilience, and process legitimacy. Social cohesion refers to the sense of shared purpose and trust within a society, facilitating cooperation towards common goals and the formulation of effective, sustainable solutions. Inclusion ensures that every person has access to markets, services, and various social, cultural, and political spaces, enabling all society members to prosper. Resilience describes a society's ability to prevent conflicts, including interpersonal violence, and to recover from or adapt to external shocks. Lastly, process legitimacy concerns how policies and programs are designed and executed, ensuring that these actions are perceived as fair, credible, and acceptable by all community members.

The practical application of these components involves complex, nonlinear interactions that are highly dependent on the specific context, underscoring the intricate dynamics present within any community or society. This framework operates within a policy arena, a space where public resources are allocated, and decisions are made through discussions, negotiations, and compromises among individuals, governments, and various stakeholder groups. Ensuring broad access to this arena, particularly for marginalized and vulnerable groups, and incorporating mechanisms such as information sharing, feedback loops, and other forms of social accountability, is crucial for addressing tensions. Figure 1 shows a diagram that summarizes the conceptual framework of Social Sustainability.

**Figure 1**: Conceptual Framework of Social Sustainability

A diagram of a process

Description automatically generated

Source: Barron et al. (2023)

* 1. **Why does Social Sustainability matter?**

**Social sustainability is instrumental to achieving and maintaining growth, poverty reduction, and combatting climate change.** Global challenges such as climate change and conflict pose unique threats to development. Ending extreme poverty and promoting shared prosperity on a livable planet requires more than just traditional growth and redistribution; social sustainability is essential for building inclusive and resilient societies. In such societies, people can cooperate, respond effectively to crises, access essential services and markets for a decent and prosperous life, and participate in decisions impacting them and future generations. Societies with poor social sustainability face vulnerabilities that undermine their social contract, exhibit low trust in people and institutions, lack resilience and opportunities, and suffer from inadequate social and governmental accountability. These weaknesses lead to compounded exclusion, particularly affecting vulnerable groups. This perspective is central to the WBG's renewed mission to end poverty and enhance shared prosperity on a livable planet. Recognizing the importance of economic growth, resilience, cohesion, and inclusion as key to ongoing poverty alleviation, this approach is captured in the [New WBG Scorecard FY24-FY30](https://documents1.worldbank.org/curated/en/099121223173511026/pdf/BOSIB1ab32eaff0051a2191da7db5542842.pdf).

**Social Sustainability provides a comprehensive lens to better understand how to bring about broad-based inclusive growth.** Inclusive growth is not just about reducing monetary poverty or income inequality; it must concentrate on enhancing access for those traditionally marginalized from economic and social opportunities, encompassing more than just their income, consumption, and wealth. Exclusion can be based on gender, disability, age, sexual orientation, ethnicity, or race. Social sustainability underscores the ability, agency, and dignity of people to access markets, services, and spaces. It also fosters a sense of shared purpose and trust, enabling them to handle shocks over time, within a society that recognizes the legitimacy of policies and programs enacted by those in authority.

* 1. **Gaps of Disaggregated Data in Social Sustainability**

**While the attention to social sustainability and its connections with growth, poverty reduction, and climate change is increasing, its evidence base needs to be further improved**. The World Bank has over 80 data platforms, 31 of which report social data, totaling around 1,500 social indicators. These include both Bank-collected and external data. While the number of platforms is notable, they are not well integrated or harmonized, either within the Bank or externally. This lack of integration is relevant as poverty and weak social sustainability is closely related but not systematically linked in the Bank's data. For instance, global poverty estimates suggest about 685 million people were living in monetary poverty in 2022, according to WBG estimates. Additionally, a third of the world's population is at risk of being excluded from markets, services, and participation in cultural and political spaces (Cuesta et al., 2022). However, we are not sure how these issues overlap. Thus, understanding and addressing the links between weak social sustainability and various forms of poverty is crucial for fostering inclusive growth and designing effective interventions.

**Increasing the existing availability of disaggregated data of vulnerable populations is key for better understanding broad-based inclusive growth**. There is a pressing need to enhance the quality, scope, and accessibility of data segmented by gender, ethnicity, sexual orientation, disability, and geolocation. While over 80 data platforms exist within the WBG, only a handful, such as the Poverty and Inequality Platform and the Gender Data Portal, attempt to tackle the complexities of group intersectionality. Focusing on disaggregation by critical socio-demographic factors can unveil the realities of those at the margins, thereby paving the way for more effective, inclusive policies and interventions. In this spirit, the Social Sustainability Global Database (SSGD) v2.0 aims to help bridge the data gap. As will be shown in the next section, the SSGD v2.0 intends to encompass the various dimensions of Social Sustainability, providing targeted information for population groups facing vulnerabilities and for geographical areas.

1. **The Social Sustainability Global Database (SSGD) v2.0**

The Social Sustainability Global Database (SSGD) v2.0 represents a leap forward in data disaggregation and spatial analysis that addresses the multiple dimensions of social sustainability. This new version goes beyond national aggregated data and population group breakdowns to include time analysis, subnational analysis, and geospatial visualization that overlays with climate data among other sources. SSGD v2.0 includes 85 leading indicators of the four pillars of social sustainability: inclusion, resilience, social cohesion, and process legitimacy, plus 43 development indicators of associated development outcomes such as economic growth, inequality, and poverty (See Appendix A for a codebook with a complete list of indicators). The SSGD v2.0 comprises 222 countries distributed among the seven WBG regions. Figure 2 shows the structure of the SSGD v2.0.

**Figure 2**: SSGD v2.0 structureA diagram of a diagram of a social policy

Description automatically generated with medium confidence

Source: Own elaboration.

A review of publicly accessible, internationally reliable, and widely used data sources identified 18 data sources of social sustainability-related indicators. These sources include public opinion surveys, institutional databases, and repositories, research and analysis projects, law and governance initiatives, and studies on the rights of sexual and gender minorities, among others. As the construction of disaggregated metrics on social sustainability requires using microdata, primarily surveys, it was crucial to consider the sampling design and sampling weights to ensure the estimates are representative of each disaggregation level addressed by the SSGD v2.0. Table 1 presents a complete list of these data sources along with source links, and below we provide a brief description of them.

**Table 1**: Data sources used to construct the SSGD v2.0

|  |  |  |
| --- | --- | --- |
| **Source** | | **Link to data portal** |
| 1 | Global Monitoring Database (GMD) | Restricted Access. Only available for WBG staff at datalibweb. |
| 2 | Afrobarometer (AF) | <https://www.afrobarometer.org/> |
| 3 | Arab Barometer (AB) | <https://www.arabbarometer.org/> |
| 4 | Asian Barometer (ASB) | <https://www.asianbarometer.org/> |
| 5 | Latinobarometro (LB) | <https://www.latinobarometro.org/lat.jsp> |
| 6 | World Values Survey (WVS) | <https://www.worldvaluessurvey.org/wvs.jsp> |
| 7 | European Values Study (EVS) | <https://europeanvaluesstudy.eu/> |
| 8 | World Justice Project (WJP) | <https://worldjusticeproject.org/> |
| 9 | Worldwide Governance Indicators (WGI) | <https://www.worldbank.org/en/publication/worldwide-governance-indicators> |
| 10 | World Development Indicators (WDI) | <https://databank.worldbank.org/source/world-development-indicators> |
| 11 | Armed Conflict Location & Event Data Project (ACLED) | <https://acleddata.com/> |
| 12 | The Global FINDEX Database (FINDEX) | <https://www.worldbank.org/en/publication/globalfindex> |
| 13 | CIVICUS | <https://monitor.civicus.org/> |
| 14 | United Nations Development Programme (UNDP) | <https://hdr.undp.org/data-center/human-development-index#/indicies/HDI> |
| 15 | Women, Business, and the Law (WBL) | <https://wbl.worldbank.org/en/wbl> |
| 16 | Economist Intelligence Unit (EIU) | <https://www.eiu.com/n/campaigns/democracy-index-2022/> |
| 17 | Sovereign Environmental, Social, and Governance indicators (ESG) | <https://esgdata.worldbank.org/?lang=en> |
| 18 | Equality of Opportunity for Sexual and Gender Minorities (EQOSOGI) | <https://www.worldbank.org/en/publication/equality-of-opportunity-for-sexual-and-gender-minorities> |

Source: Own elaboration.

Among the eighteen data sources, the **Global Monitoring Database (GMD)**, produced by the World Bank, is one of the key ones for constructing indicators of social sustainability that are comparable across countries. The GMD harmonizes multi-topic household surveys, such as national household budget surveys and Living Standards Measurement Surveys, to construct a globally comparable database of harmonized rates for tracking and monitoring international poverty and shared prosperity indicators, disaggregated by gender, age, employment, and location among others.

A second key source for constructing metrics of social sustainability is social barometers that collect information about the attitudes and values of people by region. Among these, we have **Afrobarometer (AF), Arab Barometer (AB), Asian Barometer (ASB), and Latinobarometro (LB). These barometers collect the experiences and public opinions of adult citizens[[1]](#footnote-1) on** democracy, governance and politics, economic issues, cultural and religious beliefs, attitudes towards immigration and ethnic diversity, gender equality, trust, social issues such as crime and security, and international relations and globalization, aiming to give citizens a voice in policymaking monitoring, analyzing social trends and promoting the global dialogue on democratic governance and societal change. Barometers include a sample size ranging from 1000 to nearly 2400 interviewees for most countries[[2]](#footnote-2) and are representative at the national and sub-national levels.

Like barometers, the **World Values Survey (WVS)** and the **European Values Study (EVS)** provide a rich source of data for constructing social sustainability indicators. The World Values Survey is a global initiative that measures, monitors, and analyzes aspects such as support for democracy, political participation, trust, economic issues, tolerance of foreigners and ethnic minorities, support for gender equality, the role of religion, and changing levels of religiosity, the impact of globalization, attitudes toward the environment, work, family, politics, national identity, culture, diversity, insecurity, and subjective well-being. On the other hand, the European Value Study collects information about politics and governance, economic issues, religion and morality, attitudes towards refugees, immigrants, and racial/ethnic minorities, views on gender roles, perceptions of social justice, globalization, and the national identity of European citizens. Both surveys are oriented toward adult citizens aged 18+, provide national and subnational representation, and include around 1200 people surveyed for most countries[[3]](#footnote-3).

A fourth key data source is nationally available data on governance, rule of law, conflict, economic development, financial inclusion, civic space, human development, women, business, and the law, democracy, environmental and social governance, and equality of opportunity for sexual and gender minorities. The **World Justice Project (WJP)** provides information on the Rule of Law worldwide, emphasizing its role in ensuring justice, economic opportunity, and respect for fundamental rights. The **Worldwide Governance Indicators (WGI)** compiles data from a variety of sources to provide comprehensive assessments of governance quality globally. The **World Development Indicators (WDI)** is a comprehensive database maintained by the WBG that provides a detailed and wide-ranging set of data on various aspects of global development. The **Armed Conflict Location & Event Data Project (ACLED)** provides detailed, real-time data on political violence and protests worldwide. The **Global Findex Database (FINDEX)**, developed by the WBG, provides data on global access to financial services, digital payments, and the behaviors fostering financial resilience. The **CIVICUS Monitor (CIVICUS)** offers near-real-time data on the health of civil society and civic freedoms in 196 countries, it provides updated assessments of civic space, informing advocacy and policy efforts aimed at protecting civic freedoms.

Other sources include the **United Nations Development Programme (UNDP)** which provides the widely disseminated Human Development Index, a statistical tool used to measure a country's overall achievement in its social and economic dimensions. The **Women, Business and the Law (WBL)** project, undertaken by the World Bank Group, systematically gathers data to measure the legal and regulatory barriers affecting women's economic participation across 190 economies. Similar to UNDP, the **Economist Intelligence Unit (EIU)**, a division of The Economist Group, provides the Democracy Index, an annual index that ranks countries based on their electoral process and pluralism, functioning of government, political participation, political culture, and civil liberties to assess the state of democracy worldwide. The **Sovereign ESG Data Portal (ESG)**, developed by the WBG, is designed to provide environmental, social, and governance data related to sovereign debt. Finally, the **Equality of Opportunity for Sexual and Gender Minorities (EQOSOGI)**, a relatively recent project carried out by the WBG, provides data on the degree of legal protection against sexual and gender minorities in countries.

* 1. **New features in the SSGD v2.0**

This new version goes beyond national aggregated data and population group breakdowns to include time analysis, subnational analysis, and geospatial visualization that overlays with climate data among other sources. The SSGD v2.0 not only offers disaggregated estimates for indicators related to social sustainability but also features a user-friendly structure. Following the WBG policy of open data and promotion of [reproducibility](https://reproducibility.worldbank.org/index.php/submit-package) of generated data, the SSGD v2.0 enables users to replicate the database from the raw data to the construction of rates in an automatic manner. To do so, the codes are written so that the variables that are constructed using standardized labeling (prefix and suffix) that standardizes is automatic and standardized comprehensive automation and standardization processes succinctly described in a publicly accessible repository (for more details, refer to section 5).

Given the diverse sources contributing to the SSGD v2.0, the database employs harmonization standards for all its indicators, ensuring comparability across various datasets. It also preserves the original metadata for each indicator, facilitating user access to the sources and providing insight into additional characteristics of the indicators. Regarding population groups, the SSGD v2.0 facilitates the breakdown of national figures by gender (men/women), age groups (classification A: 15-24 years, 25+ years; classification B: 15-29 years, 30-59 years, 60+ years), location (urban/rural), and ethnicity and religion (majority/minority groups in both categories). On a geographical level, it allows for the disaggregation of indicators to the first administrative level (ADM1), aligning with the World Bank's approved administrative boundaries ADM0 and ADM1 ([source link](https://datacatalog.worldbank.org/int/search/dataset/0038272/world-bank-official-boundaries)).

Moreover, due to the varying timelines of different sources, establishing a temporal framework was essential to organize the data and facilitate a time analysis of the indicators. The SSGD v2.0 identifies two time periods, termed waves, where wave 1 spans the years 2015 to 2018, and wave 2 covers 2019 to 2022. Figure 3 summarizes the SSGD v2.0 features in a diagram.

**Figure 3**: SSGD v2.0 features

A diagram of a software development process

Description automatically generated with medium confidence

Source: Own elaboration.

Many features of the SSGD v2.0 introduce significant novelties when compared to its predecessor, version 1.0. Initially, the SSGD had only 71 main indicators and 21 development indicators. While it maintained a harmonized framework for indicator definitions across different data sources, it lacked a temporal foundation for time analysis. In contrast, SSGD v2.0 provides more transparent access to data and the processes involved in its compilation (through public dissemination of codes used for data processing). It incorporates a wider array of data sources, updates to the latest versions of these sources, enables disaggregation by subnational territorial units, and enriches the metadata associated with its indicators.

* 1. **A descriptive view of the SSGD v2.0**

This section explores the SSGD v2.0’s data coverage and source utilization, emphasizing disaggregation and time across leading indicators. Table 2 displays the availability of data for each of the 85 primary indicators, analyzing the extent of coverage in terms of the number of countries for which data estimates are available. This information is presented for each survey wave. Moreover, the table identifies all indicators that allow for decomposition by population group (age, gender, location, ethnicity, religion) or offers subnational breakdowns at the administrative level 1 (ADM1). Note that some indicators do not allow any decomposition, such as the Rule of Law Index, developed and estimated by the World Justice Project, which considers the current legislature in the national territory. Other examples are found in those indicators estimated by the WBG through its Worldwide Governance Indicators portal.

**Table 2: Leading indicators’ distribution by wave and group decomposition**



**Table 2: Leading indicators’ distribution by wave and group decomposition (continued)**



**Table 2: Leading indicators’ distribution by wave and group decomposition (continued)**



**Table 2: Leading indicators’ distribution by wave and group decomposition (continued)**



Appendix A contains supplementary material to this technical report that offers an extended version of Table 2. There the user will find detailed information on the data availability both by wave and population group information. As part of the supplementary materials, a codebook for the SSGD v2.0 is provided.

Figure 4 reveals that a significant portion of the indicators originates from the GMD, various barometers (Afro, Arab, Asian, and Latino), and the World Values Survey. These sources provide information for at least 18 indicators each. Conversely, sources such as CIVICUS or FINDEX contribute to a single indicator each (civic space score for CIVICUS, and the percentage of individuals with a bank account for FINDEX, respectively). It is important to highlight that multiple sources can supply data for a single indicator (refer to the "Source" column in Table 2). This is facilitated by the indicator harmonization framework employed in the construction of the SSGD v2.0.

**Figure 4: Number of leading indicators by source**

A graph of numbers and a number of objects

Description automatically generated with medium confidence

Source: Own elaboration.

Due to the varying timelines across the data sources utilized in constructing the SSGD v2.0, the decision was made to categorize the information into two distinct periods: 2015-2018 (Wave 1) and 2019-2022 (Wave 2). This approach primarily addresses the significant diversity in the data collection frequencies among the different sources. For certain surveys, categorizing data into these periods is straightforward, as their collection cycles align neatly with one of the two specified intervals. However, in other instances, a single survey round may encompass countries surveyed during both Wave 1 and Wave 2 (e.g. Asian Barometer). Table 3 presents the outcomes of this temporal harmonization for each source, detailing the number of countries covered per source and/or survey cycle. Similarly, Table 4 shows the information contained in Table 3 but breaks down the values according to population groups (by gender, age, location, ethnicity, and religion) and subnational administrative unit.

**Table 3: Number of countries available by wave and source**



Source: Own elaboration

**Table 4: Number of countries available by group and source**



Source: Own elaboration.

Figure 5 illustrates the data availability by showing the number of countries for which there are estimates available in every leading indicator in the SSGD v2.0. It displays a panel for each dimension of SSI, along with a temporal comparison showing the data availability across each wave and their joint availability. Generally, aggregated databases like the WGI provide extensive coverage across a broad number of countries and over numerous years or periods. This extensive coverage means indicators such as Voice and Accountability and Control of Corruption offer data for both waves in at least 200 countries. Conversely, indicators derived from barometers, notably those under the Social Cohesion dimension, are often restricted to a specific subset of countries within the regions where these instruments are deployed, resulting in more limited data availability.

**Figure 5: Number of countries by leading indicators and wave**

|  |  |
| --- | --- |
| **Social Inclusion indicators** | **Resilience indicators** |
| **A graph of numbers and a number of countries/regions  Description automatically generated** | **A diagram of a number of countries/regions  Description automatically generated** |
| **Social Cohesion indicators** | **Process Legitimacy indicators** |
| **A diagram of a number of people  Description automatically generated with medium confidence** | **A close-up of a graph  Description automatically generated** |
| Source: Own elaboration. |  |

1. **Data visualization**

The SSGD v2.0 introduces a data visualization tool, accessible to the public through a Tableau-developed dashboard. This tool exhibits three main features: a country profile panel for detailed national insights, benchmarking panels for cross-comparison among countries, regions, and globally, and a panel for identifying associations between indicators. It employs bar charts, box plots, maps, and scatter plots.

The country profile and benchmarking functionalities include filters for region, income group (as per World Bank Group classifications), and classification within fragile and conflict-affected states (FCS). Likewise, each country can be chosen directly by its name. Except for the association panel, all panels offer the ability to select specific waves, categories, or analysis groups (e.g., gender, age groups), indicator dimensions (e.g., social inclusion, development indicators), and the indicators themselves. On the other hand, the association's panel requires selecting a pair of variables for scatter plot analysis, with options to filter by region, wave, and category.

**Figure 6**: SSGD v2.0 tools and their possible uses

A diagram of a company profile

Description automatically generated

Source: Own elaboration

This tool's outputs can be used for creating products across various WBG units, for instance, they could help in the development of Country Climate and Development Reports (CCDR) and other WBG reports by providing essential data insights. Also, they could serve as a foundational tool for crafting knowledge narratives, corporate data reporting, supporting qualitative studies, establishing socioeconomic typologies, and facilitating comprehensive analyses across time, space, and other databases like the World Development Indicators (WDI), Global Monitoring Database (GMD), and Green, Resilient, and Inclusive Development (GRID). Figure 6 summarizes the SSGD v2.0 tools and their possible uses.

On the other hand, Figure 7 highlights key features of the SSGD v2.0 Tableau dashboard. Box 7A presents the dashboard's main panel, which summarizes the content of the dashboard and enables access to all other panels. Box 7B displays the country profile panel, where tabulations for all development indicators of a chosen country are shown by default. Box 7C presents a comparative analysis of the same indicator between two pairs of countries or periods. This benchmarking allows us to compare the profiles of two countries during the same period, as well as the profile of a single country across different periods. Box 7D features a regional benchmarking panel, displaying a boxplot by region (as classified by the World Bank) for a specific indicator and period. Box 7E presents a global benchmarking panel presenting a heat map comparison based on a selected indicator. Lastly, Box 7F introduces the association panel, offering a scatter plot that visualizes the relationship between two indicators selected by the user.

**Figure 7: SSGD v2.0 Tableau dashboard features**

|  |  |  |
| --- | --- | --- |
| 1. **Main interface** | 1. **Country Profile** | 1. **Country Benchmarking** |
|  |  |  |
| 1. **Regional Benchmarking** | 1. **Global Benchmarking** | 1. **Associations** |
|  |  |  |

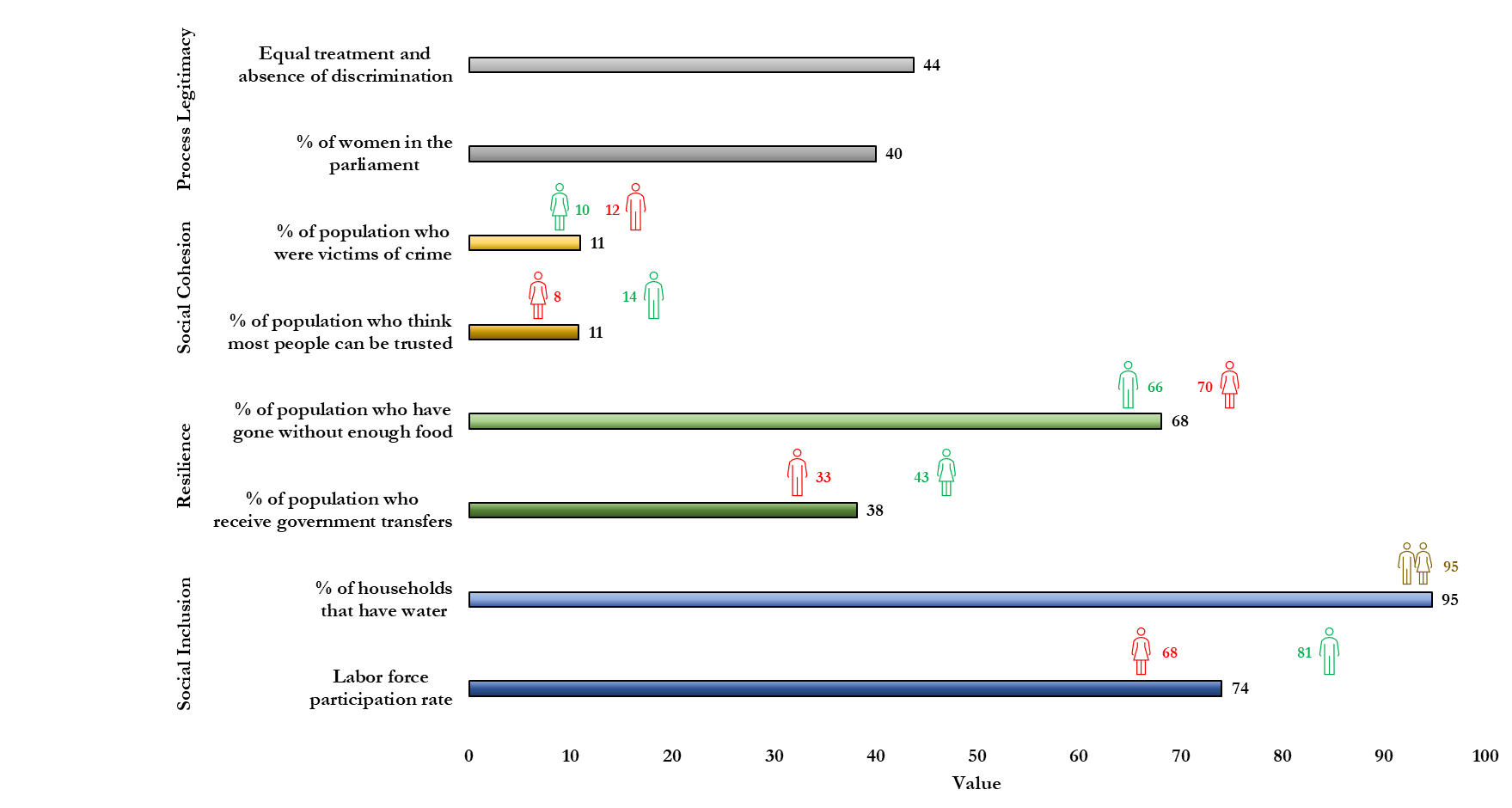
1. **An application exercise using the SSGD v2.0**

This section demonstrates a practical application of the SSGD v2.0 data, using Peru—an upper-middle-income country in the Latin American and Caribbean (LAC) region—as a case study. Initially, a comprehensive country profile for Peru is crafted, selecting eight key indicators, two representing each dimension of Social Sustainability. Our analysis will highlight differences in these indicators across various population groups, and geographical regions, and their changes over time. Subsequently, Peru is compared with neighboring Andean nations, Ecuador and Bolivia, using four indicators—one from each social sustainability dimension—to observe their progression nationally over time. Lastly, a regional analysis is conducted within the LAC region, centering on trust within communities as a crucial variable, to explore associations and draw insights.

* 1. **A country profile for Peru**

As highlighted in the previous paragraph, the results for Peru showcase data on eight social sustainability indicators from wave 2 (2019-2022). These include two indicators related to Social Inclusion (Labor force participation rate and Share of households that have access to water), two associated with Resilience (Share of the population that has gone without enough food to eat and Share of the population ), two concerning Social Cohesion (Share of the population that was victim of a crime and Share of the population that says most people can be trusted), and two tied to Process Legitimacy (Equal treatment and absence of discrimination score and Share of women in the parliament).

**Figure 8: Country Profile for Peru in Wave 2 (2019-2022). National estimates and breakdown by gender**



Source: Own elaboration.

Figure 8 highlights significant disparities in Peru. On the one side, nearly the entire population has access to water, and about three-quarters of the working-age population is part of the labor force (both employed and unemployed). Additionally, nearly half of the members in the Peruvian Congress are women, indicating progress towards achieving gender parity. However, challenges persist: almost 70% of Peruvians report experiencing food scarcity in the past year, only 10% express trust in others, and the country scores 44 out of 100 on equal treatment and absence of discrimination. Gender disparities are evident as well, with only 8% of women reporting trust in others compared to 14% of men. Women are also marginally more affected by food insecurity (70% versus 66% for men), and there is a significant gender gap in labor force participation, disadvantaging women by approximately 13 percentage points.

**Table 5: Country Profile for Peru in Wave 2 (2019-2022). National estimates and breakdown by population groups**



Source: Own elaboration.

Table 5 further dissects Peru's national data across different demographic groups such as age, location, ethnicity, and religion, unveiling pronounced disparities. At the age group level, labor force participation significantly favors adults (30 to 59 years) with an 85% participation rate, compared to 68% among youth (15 to 29 years) and 56% among seniors (60 years and older). Food insecurity is more acute among adults, with at least 70% affected, compared to 64% in younger populations. Interestingly, young people report a higher incidence of crime victimization (15%) than adults (up to 9%). The discrepancies between urban and rural residents are significantly larger in all indicators. Rural populations have higher labor force engagement but less access to water, receive more government transfers, face greater food insecurity, have a higher trust in others, and are less likely to be crime victims compared to their urban counterparts. Differences also emerge across ethnic and religious groups, especially in terms of government transfers, food insecurity, and crime victimization.

**Figure 9: Scorecard for the evolution of selected indicators. Peru, national estimates**A diagram of a social inclusion

Description automatically generated with medium confidence

Source: Own elaboration.

The SSGD v2.0 dataset could also be used to create scorecards for evaluating progress across various indicators. Figure 9 features a scorecard comparing the performance of eight key indicators between Wave 1 (2015-2018) and Wave 2 (2019-2022) for Peru. The findings illustrate mixed outcomes: on one hand, Peru has achieved notable improvements, such as enhanced access to drinking water, increased government transfers to citizens, and a higher representation of women in parliament. On the other hand, there have been declines in labor force participation, trust in other people, and scores for equal treatment and non-discrimination, alongside rises in food insecurity and crime rates. These shifts may be closely associated with the stringent quarantine measures implemented during the COVID-19 pandemic in this country.

* 1. **Comparative assessment: Peru and other Andean countries**

The comparative study between Peru and its Andean neighbors, Ecuador and Bolivia, focuses on the evolution of four key social sustainability indicators across waves: labor force participation rate, the prevalence of food insecurity, levels of trust in other people among the population, and the proportion of women in parliament.

Figure 10 displays the national data for these indicators across the three countries by wave. Panels 10A and 10B show the results for waves 1 and 2 respectively. The symbols next to the estimated values in panel 10B not only indicate the direction of change over time but also highlight whether these changes represent advancements (green) or regressions (red) for each country. Unfortunately, all three countries have experienced declines in the levels of trust in people and increases in food insecurity. Nonetheless, Peru and Ecuador have both witnessed growth in women's representation in parliament, with Ecuador also reporting a rise in labor force engagement. Bolivia, however, has experienced a sharp decline in the percentage of women in the parliament, while the labor force participation has stayed the same.

**Figure 10: Time analysis panel for selected indicators between Peru and other Andean countries. National estimates.**

|  |
| --- |
| 1. **Wave 1 (2015-2018) results** |
|  |
| 1. **Wave 2 (2019-2022) results** |
|  |

Source: Own elaboration.

* 1. **Regional analysis: Associations with trust in people in LAC**

Another application for the SSGD v2.0 consists of providing visualization tools that highlight the relationship between the population's trust in others and two key indicators: the Strength of Legal Rights Index (SLRI) and the Regulatory Quality Estimate (RQE) among countries in the LAC region. The SLRI evaluates how well collateral and bankruptcy laws protect the rights of borrowers and lenders, thereby facilitating lending, scored from 0 to 12. Higher values signal laws more effectively designed to broaden access to credit. Conversely, the RQE assesses the government's capacity to develop and enforce policies and regulations that support private sector growth, with its values ranging from -2.5 to 2.5. Higher scores on RQE suggest a more conducive regulatory environment for the private sector, while lower scores indicate potential obstacles in policy development and execution that could impede economic progress.

Panel 11A in Figure 11 presents a boxplot illustrating the proportion of the population reporting trust in others across different SLRI levels, categorized as high (8 to 12 points), medium (2 to 7 points), and low (0 to 1 point). The data reveal a clear trend: trust in others increases with higher SLRI levels. This pattern underscores the impact of effective legal frameworks in cultivating a trusting environment, where robust legal protections for transactions and contracts not only enhance security but also mitigate exploitation risks, reflecting the presence of well-functioning institutions that collectively foster greater social trust and cooperation.

On the other hand, panel 11B displays a scatter plot correlating the proportion of the population that trusts others with the RQE value. Notably, Uruguay and Chile exhibit the highest marks in both indicators, in contrast to Venezuela, which ranks lowest. This scatter plot exhibits a positive correlation between trust in others and regulatory quality, indicating that well-crafted and executed government policies can generate a climate of trust and cooperation. High regulatory quality, characterized by transparency, fair interactions, and effective protections against misuse, promotes interpersonal trust by ensuring a secure and equitable environment for economic and social endeavors.

**Figure 11: Associations between the share of the population who thinks people can be trusted and selected indicators for the LAC region.**

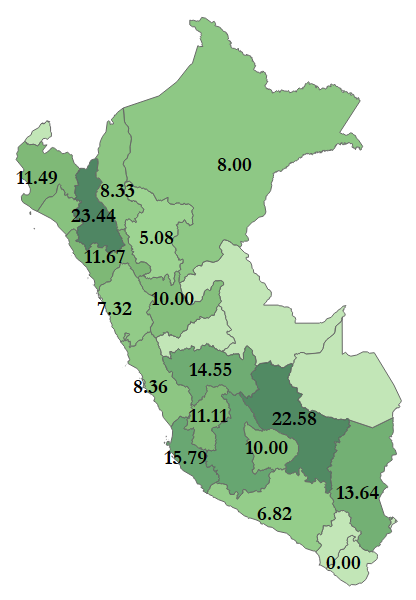
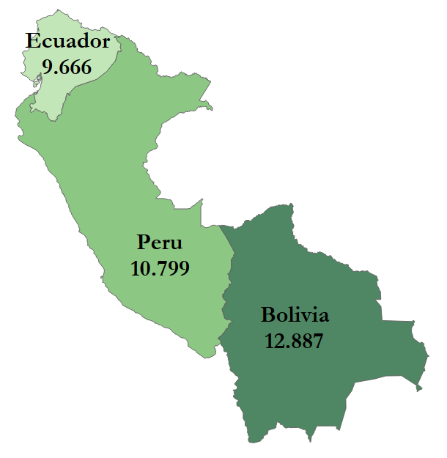
|  |  |
| --- | --- |
| 1. **Strength of Legal Rights Index** | 1. **Regulatory Quality Estimate** |
|  |  |
|  |  |

Source: Own elaboration.

* 1. **Subnational analysis: Trust in people in Andean countries**

To conclude our presentation of application examples using SSGD v2.0, we offer maps illustrating the diversity in the proportion of people expressing trust in others, focusing on Peru and its neighboring countries, Ecuador and Bolivia. These visualizations cover the period from 2019 to 2022, as indicated in Wave 2. Figure 12 (left) employs a green gradient to depict trust levels, with darker shades indicating higher trust. Bolivia exhibits the highest level of trust, with 12.9% of its population reporting trust in others, followed by Peru with 10.8%, and Ecuador with 10%. Additionally, Figure 12 (right) details the breakdown of trust levels within Peru's departments, and its primary administrative divisions, revealing significant variability. Note that some data points are omitted due to the lack of information from the source for certain regions.

**Figure 12**: Share of the population who trust in people. Peru and other Andean countries. Wave 2 (2019-2022)



Source: Own elaboration using Tableau.

1. **Use and replicability of the SSGD v2.0**

The use and replicability of the SSGD v2.0 will be discussed around two processes: standardization and automation. On the one hand, a standardization process is performed regarding the indicators' construction, labeling, and metadata. On the other hand, the automation process refers to the replicability of the SSGD v2.0. These processes guarantee the ease of use of the SSGD v2.0 and the transparency of all computations involved.

* 1. **Standardization processes and general guidelines in the use of the SSGD v2.0**

The standardization procedures can differ based on the specific format of the SSGD v2.0 data. The SSGD v2.0 is available in two formats: wide (ssgd\_wide.dta) and long (ssgd\_long.dta). The following sections detail the standardization processes and general guidelines for the use of each data format.

* + 1. **Wide data format**

**Construction of the indicators**

Many of the leading indicators of SSGD v2.0 come from surveys whose content may change across the years; moreover, a single indicator could come from more than one source (e.g., barometers, WVS). For these reasons, a standardization process was opted for, covering all the surveys used, as well as the periods in which they are available. As mentioned before, Appendix A provides a codebook that shows the definitions and questions used to construct all leading indicators. When an indicator comes from two or more sources, specific concepts are adopted depending on the source. The standardization process does not apply to development indicators since they come from a single source. This also happens to some leading indicators derived from open data portals rather than surveys (e.g., WDI, WGI, and WJP).

**Labeling**

Concerning labeling, all variable names[[4]](#footnote-4) share the same structure or syntax. All variable names start with a term that indicates the dimension or the condition of being a development indicator. There are five terms used: i) si (stands for Social Inclusion), ii) re (Resilience), iii) sc (Social Cohesion), iv) pl (Process Legitimacy), and v) di (Development indicator). Then, a short but unique name for the variable is included. Finally, a final term that refers to the wave or period is added, which can be either w1 or w2. To separate each component in the variable name (i.e., dimension, short name, and wave), the character “\_” is used as a separator symbol. In summary, all the variable names included in the SSGD v2.0 share the following structure:

varname: [dim]\_[short name]\_[wave]

For instance, the name "re\_govtra\_w1" informs that this is a leading indicator classified in the Resilience dimension, includes information for wave one, and has the short name "govtra", which refers to the share of the population who received transfers from the government.

**Metadata**

The SSGD v2.0 includes metadata for all computations involved. In particular, it shows the calculation's source and year for each variable. Prefixes are used to distinguish the metadata from the indicators themselves. The term “sou” indicates the source, while “per” refers to the year or period corresponding to the variable. By working with the previous example, the variables "sou\_re\_govtra\_w1" and "per\_regovtra\_w1" indicate the source and year of origin for the calculation of the indicator "re\_govtra\_w1" in wave one, respectively. For instance, in Peru, the value calculated for the indicator "re\_govtra\_w1" was obtained from the LB, and the survey year corresponds to 2018. Note that the separator "\_" is also applied in using prefixes in the metadata.

* + 1. **Long data format**

**Construction of the indicators**

Same as in wide format.

**Labeling and metadata**

The long data format offers much more information than the wide format as it summarizes all calculations or estimates in a single column called value. The rest of the columns included refer to either one of the following features:

**Country and subnational unit identifiers**: countryname (country name), countrycode (ISO-3 level code for country), region (according to the WBG classification), incomegroup (income group according to the WBG classification), fragile (Fragile and conflict-affected states, FCS), adm0\_code (similar to countrycode but using WBG country codes), adm1\_code (WBG code for subnational units corresponding to the ADM1 level).

**Indicator identifiers**: dimension (four pillars of Social Sustainability and development indicators), area (includes gender, age groups, ethnicity, subnational, etc.), category (includes concrete population groups like male, female, 15-24 years, etc. and subnational territories specific to each country).

**Information about the indicator**: short (which includes syntax similar to variable labeling in the wide format except that it ignores wave information, i.e. varname: [dim]\_[short\_name]), variable (which is defined as the union of both short and category columns[[5]](#footnote-5), i.e. variable: [short]\_[category]), indicator (displays a short definition for the indicator), and definition (displays a long definition for the indicator).

**Metadata related to indicators**: indicator\_type (indicator type, e.g. percentage, magnitude, index, etc.), wave (which can be wave 1: 2015-2018 or wave 2: 2019-2022), period (the exact year for the indicator was calculated), source (source of origin, e.g. GMD, WVS, etc.), scale (scale of the indicator, e.g. percentage, constant 2017 USD, etc.), and range (minimum and maximum values taken by the indicator).

**Relevance to specific Social Sustainability and Inclusion Global Practice (SSIGP) areas**: rel\_cesa is a dummy variable that highlights those indicators that are relevant for Citizen Engagement and Social Accountability; rel\_scorecld does the same for indicators linked to Social Cohesion and Resilience (SCORE), and Community and Local Development (CLD); finally, rel\_inclusion does the same for indicators linked to Inclusion. Keep in mind that while dimension offers a general and corporative classification of all the indicators, rel\_cesa, rel\_scorecld, and rel\_inclusion provides alternative classifications based on SSIGP areas.

As an application, let’s assume the user writes the following code in Stata:

br countryname value period if countrycode=="PER" & wave==1 & short=="re\_govtra" & area=="Subnational"

Then, the user will be able to see in the value column the share of the population who received transfers from the government for all the available subnational units of Peru in wave one (2015-2018). Of course, the column period will give us the exact year corresponding to this data.

For specific applications, the wide data format is optimal, such as when users aim to explore associations between leading indicators through scatterplots. Conversely, the long data format is better suited for other uses, for example, the dashboard linked to the SSGD v2.0 utilizes the long data format.

* 1. **Automation processes and replicability of the SSGD v2.0**

All script files used to develop the SSGD v2.0 are publicly available in the SSGD v2.0 [GitHub repository](https://github.com/Paola-Ballon/Social-Sustainability-Global-Database-v2.0/tree/main/documentation). Such files were designed and organized to ensure the replicability of this database by anyone who has access to the World Bank Group’s intranet. This section briefly discusses the files used to develop the SSGD v2.0 and presents guidelines for their proper use.

Seven master files have been coded in the R and Stata languages to ensure ease of replicability of the SSGD v2.0. The user only needs to work with these master files, classified into the following five blocks: 1) User profile, 2) Data download, 3) Generation of indicators, 4) Data processing, and 5) Data merge. Figure 13 shows a sequential diagram that summarizes all the blocks and lists the master files included in each block.

**Figure 13: Replicability of the SSGD v2.0**

**Software requirements**

Before discussing every block of the replication process in detail, a list of requirements is provided that every user must fulfill to replicate the SSGD v2.0 without problems. First, the user must have both Stata 16 MP and the latest version of R installed on their device. These are minimum requirements; for instance, a user could use Stata 17/18, but the MP version is strongly recommended. Also, the user can choose any development environment software to access the R scripts; RStudio is a popular choice among R users. Second, the user must install the datalibweb and wbopendata packages in Stata before running any do file in the SSGD v2.0 repository. The datalibweb package enables the user to access harmonized household surveys from the World Bank’s microdata library, such as GMD ([source link](http://spqsapps.worldbank.org/qs/ECA/_layouts/15/WopiFrame.aspx?sourcedoc=/qs/ECA/SiteAssets/datalibweb_Guidelines_1page.pdf&action=default), restricted access to WBG staff). On the other hand, the wbopendata package enables the user to access over 3000 indicators from the World Bank databases ([source link](https://datahelpdesk.worldbank.org/knowledgebase/articles/889464-wbopendata-stata-module-to-access-world-bank-data)).

**User profile**

In the user profile block, the user must declare, for one time, the path or working directory where she decided to copy the "SSGD v2.0" folder. This way, the rest of the files will adapt to any user-defined folder. The master files present in this block are "ssgd\_v2\_user\_profile.R" and "ssgd\_v2\_user\_profile.do". It is essential to declare the same path or working directory in both master files before executing them. By default, the working directory for both files is "C:/Users/PC/Desktop/" but the user may change it.

**Data download**

The data download block consists of a semiautomatic process of obtaining raw data to construct the indicators included in the SSGD v2.0. In other words, it entails downloading microdata and other data files. The master files included in this block are "ssgd\_v2\_data\_download\_p1.R" and "ssgd\_v2\_data\_download\_p2.do", the user must execute them in the order described right after running the master files of the previous block.

The data download block is described as semiautomatic because part of it is executed automatically using the master files. In contrast, the rest must be done manually by the user. Except for GMD, all data sources offer data for free. However, in some cases, access to data requires registration or filling out a data request form, so it is not always possible to obtain the information through web scrapping or related computational procedures[[6]](#footnote-6).

Table 6 lists the databases used in the SSGD v2.0 and identifies those requiring manual data download. For such cases, the procedure involves the following steps: 1) access a particular URL (depending on the source), 2) register or fill out a data request form, 3) download the data, and 4) assign it to a folder contained in "SSGD v2.0" main folder. The user is not required to create or modify any folder within the "SSGD v2.0" main folder since all folders are created automatically along with the execution of the master file "ssgd\_v2\_data\_download\_p1.R". All downloaded files are stored in the “raw\_data” folder (within the “SSGD v2.0” folder). Appendix B provides guidelines for each database where the user must download data manually.

**Table 6: Databases used: Access and download process**

|  |  |  |  |
| --- | --- | --- | --- |
| **Source** | **Free Access** | **Registration required** | **Download Type** |
| GMD | No | No | Automatic |
| AF | Yes | No | Automatic |
| AB | Yes | No | Automatic |
| ASB | Yes | Yes | Manual |
| LB | Yes | No | Manual |
| WVS | Yes | Yes | Manual |
| EVS | Yes | Yes | Manual |
| ACLED | Yes | Yes | Manual |
| CIVICUS | Yes | No | Included |
| FINDEX | Yes | No | Automatic |
| WDI | Yes | No | Automatic |
| WGI | Yes | No | Automatic |
| WJP | Yes | No | Automatic |
| EIU | Yes | No | Included |
| ESG | Yes | No | Automatic |
| WBL | Yes | No | Automatic |
| EQOSOGI | Yes | No | Included |
| UNDP | Yes | No | Automatic |

**Generation of indicators**

Once the user completes the semiautomatic data download process, she must execute the master file "ssgd\_v2\_gen\_indicators.do" file, which generates all the indicators in the SSGD v2.0. It is necessary to run all the previous steps correctly so that the master file can create the indicators properly and without execution errors. The user will notice that the “proc\_data” folder (within the “SSGD v2.0” folder) serves to store every file generated by this master file. The “proc\_data” name indicates that it is processed data (unlike the "raw\_data" folder).

**Data processing**

After creating all indicators, the information is classified by wave. The master file "ssgd\_v2\_data\_proc.do" will perform this task. Likewise, this master file aims to establish "dominance relationships" between the different databases used since there are cases where the same indicator is present in more than one data source.

There are three dominance relationships established in SSGD v2.0. In the first place, the FINDEX database dominates other databases for the “ownban” indicator (% of people with a bank account) since the FINDEX database is global (includes almost all countries) and presents the information with fixed periodicity. Secondly, the AF database dominates the AB one only for the indicators they have in common and for the following countries: Morocco, Sudan, and Tunisia. The reason is AF provides more information in general, and the years included are more recent. Finally, barometers (AF, AB, ASB, and LB) dominate the WVS database in their common indicators and countries since the countries in the WVS database appear once every several years.

**Data Merge**

For the final block in the replication process of the SSGD v2.0, the user must run the “ssgd\_v2\_merge.do” to merge all the indicators processed so far. Unlike the processed data in the previous block, the "final\_data" folder will contain the final product of this master file, i.e., the data files "ssgd\_v2\_wide.dta" and "ssgd\_v2\_long.dta". Figure 14 presents a chronological diagram summarizing the steps to replicate the SSGD v2.0.

**Figure 14: Steps to replicate the SSGD v2.0**

A screenshot of a computer screen

Description automatically generated

Source: Own elaboration.

1. **Further improvements**

To enhance the utility and precision of the SSGD v2.0, it is recommended to transition from wave notation to a continuous time series model. This change would allow for annual or more frequent data updates, providing users with timely insights and enabling more detailed trend analysis. This method will facilitate a deeper understanding of how social indicators evolve, which is crucial for evaluating the effects of policies or global events.

Further development could also include expanding the subnational decomposition of indicators. Extending this feature to cover more granular regional data would enable precise identification of disparities at local levels. With more detailed regional analysis, policymakers and researchers could better target interventions and resources to areas where they are most needed, thus enhancing the efficacy of development programs and policies.

Additionally, revising the ethnicity harmonization framework to a frequentist approach would improve the database's analytical depth. By categorizing data according to the top three or five most populous ethnic groups and including a category for all remaining groups, the database would offer a clearer picture of ethnic dynamics. This structure would support more targeted studies of ethnic disparities and inform policies designed to address specific community needs.

1. **References**

Cuesta Leiva, J. A., López-Nova, B., & Niño-Zarazúa, M. (2022). Social Exclusion: Concepts, Measurement, and a Global Estimate (No. 10097). The World Bank.

Barron, P., Cord, L., Cuesta, J., Espinoza, S., & Woolcock, M. (2023). Social sustainability in development: Meeting the challenges of the 21st century. World Bank Publications.

1. **Appendix**
2. **Supplementary material for the SSGD v2.0**

The supplementary material is freely accessible and available in the SSGD v2.0 repository on GitHub. Below you can find the source links to the document containing the extended version of Table 2 present in section 2 of this document and the SSGD v2.0 codebook.

[Link to the extended version of Table 2](https://github.com/Paola-Ballon/Social-Sustainability-Global-Database-v2.0/blob/main/documentation/supplementary_material_SSGD_v2.0.xlsx)

[Link to the SSGD v2.0 codebook](https://github.com/Paola-Ballon/Social-Sustainability-Global-Database-v2.0/blob/main/documentation/SSGD%20v2.0%20codebook.xlsx)

1. **Instructions to manually download selected databases**

Here are instructions for downloading raw data and associated documentation for selected data sources used in SSGD v2.0 that are not able to be downloaded by automatic means such as web scrapping: ASB, LB, ACLED, WVS, and EVS.

Asianbarometer (ASB): The user must fill out a data request form by clicking on the following link: <https://www.asianbarometer.org/data?page=d10>. After that, the user must select all the databases required. For the case of ASB wave four, the user must choose the "All Country" option. For ASB wave 5, the user must select the data files for the following countries: Philippines, Mongolia, Australia, India, Indonesia, Japan, South Korea, Malaysia, Myanmar, Taiwan, Thailand, and Vietnam. After that, a download portal will open with a list of the selected data files, and the user must download and unzip all the files listed. Finally, the user must add the uncompressed data to the folders "asianbarometer4" and "asianbarometer5" appropriately. These folders are inside a folder called "raw\_data" which, in turn, is located inside the "SSGD v2.0" folder.

Latinobarometro (LB): The user can access the data download portal for the LB by clicking on the link: <https://www.latinobarometro.org/lat.jsp>. There are no register or data form requirements to access the data. The user must download the files in the Stata format for 2017, 2018, and 2020. The user must uncompress and copy the files to the folders "latinobarometro2017", "latinobarometro2018", and "latinobarometro2020" as appropriate. The user can find these folders within the "raw\_data" folder mentioned before.

ACLED: To access ACLED data, registering on the portal website is mandatory. The link to register is as follows: <https://developer.acleddata.com/>. Once registered, the user will be able to access his profile in the so-called "ACLED Access Portal" where he will have at his disposal an "access key" (for instance, "Q6CrThSB-G1UUcTL9ylZ") that will allow him to access the data from the following link: <https://acleddata.com/data-export-tool/>.

For each download request, the user must provide the "access key", and the email address used for registration. Also, the user must enable the compatibility mode as a type of data export. The user must request two downloads and therefore obtain two data files. The first corresponds to the period between 01/01/2018 and 31/12/2020, while the second corresponds to the period between 01/01/2021 and 31/12/202. Finally, the user must copy the data files to the "acled" folder inside the "raw\_data" folder.

WVS: The user must complete a request form on the WVS portal website to access the data. The link to the form is as follows: <https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp>. The user must download a file called "WVS TimeSeries 1981 2022 Stata v4 0.zip", in the "Longitudinal Data Files" section at the end of the web portal. After that, the user must uncompress and copy the data files to the "wvs6\_7" folder within the "raw\_data" folder.

EVS: The user must complete an online registration form at GESIS (Leibniz Institute for the Social Sciences). Here you can find the link to the form: <https://login.gesis.org/realms/gesis/login-actions/registration?client_id=gesis-gws-client&tab_id=zKStYMozrOQ>. Once you log into the GESIS portal you must access the EVS2017 Integrated Dataset (ZA7500) at <https://europeanvaluesstudy.eu/methodology-data-documentation/survey-2017/full-release-evs2017/documentation-survey-2017/>. The user must download the “ZA7500\_v5-0-0.dta.zip” file, it contains a Stata .dta file with the microdata. After uncompressing the data file, the user must copy the data file to the “evs” folder within the “raw\_data” folder.

To summarize these steps, Table B.1 lists the files the user must download and copy to the "raw\_data" folder. Five sources and 21 files are involved in this manual data-downloading process. Finally, it is worth noting that three databases are already "included" in the “SSGD v2.0” folder (as shown in Table 6); these databases are the following: CIVICUS, EIU, and EQOSOGI. These data sets have been included by default since it is impossible to download the data from a web portal. Instead, all the information within these sources comes from reports. Table B.2 lists the databases that are included by default in the "SSGD v2.0" folder.

**Table B.1**: List of data files to be manually downloaded and copied

|  |  |  |  |
| --- | --- | --- | --- |
| **Source** | **Access** | **Files** | **Folder** |
| ASB | <https://www.asianbarometer.org/data?page=d10> | W4\_v15\_merged20181211\_release.dta | asianbarometer4 |
| 20230504\_W5\_merge\_15.dta | asianbarometer5 |
|  |  | Latinobarometro2017Esp\_v20180117.dta | latinobarometro2017 |
| LB | <https://www.latinobarometro.org/latContents.jsp> | Latinobarometro\_2018\_Esp\_Spss\_v20190303.sav | latinobarometro2018 |
|  |  | Latinobarometro\_2020\_Esp\_Stata\_v1\_0.dta | latinobarometro2020 |
| ACLED | <https://acleddata.com/data-export-tool/> | 2018-01-01-2020-12-31.csv | acled |
| 2021-01-01-2022-12-31.csv |
| WVS | <https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp> | WVS\_TimeSeries\_4\_0.dta | wvs6\_7 |
| EVS | <https://europeanvaluesstudy.eu/methodology-data-documentation/survey-2017/full-release-evs2017/documentation-survey-2017/> | ZA7500\_v5-0-0.dta | evs |

**Table B.2**: List of databases included by default

|  |  |  |
| --- | --- | --- |
| Source | Account | Access |
| CIVICUS | People Power Under Attack 2022 | <https://civicus.contentfiles.net/media/assets/file/2022GlobalFindingsEmbargoed16March.pdf> |
| EIU | Democracy Index Reports 2015-2022 | <https://www.eiu.com/n/campaigns/democracy-index-2022/?utm_source=google&utm_medium=paid-search&utm_campaign=democracy-index-2022&gclid=EAIaIQobChMIhNOPlc-HgAMVgUBIAB3OTwxmEAAYASAAEgJjn_D_BwE> |
| EQOSOGI | Equality of Opportunity for Sexual and Gender Minorities | <https://www.worldbank.org/en/publication/equality-of-opportunity-for-sexual-and-gender-minorities> |

aIn the case of EIU, only the URL for the 2022 report is provided.

1. Barometers typically involve participants aged 18 and older, with some exceptions where the starting age is lower, such as 16 in Brazil and Nicaragua, and 17 in Palestine. In some Asian countries, the minimum age for being a participant is 19 or 20. [↑](#footnote-ref-1)
2. Notable exceptions include countries like India or China where the sample size exceeds 4000. [↑](#footnote-ref-2)
3. Some exceptions include Canada, China, and Indonesia with more than 3000 observations in WVS, and Denmark, Switzerland, Germany, Italy, and Netherlands with more than 2000 interviewees in EVS. [↑](#footnote-ref-3)
4. The terms “indicators” and “variables” are used interchangeably. [↑](#footnote-ref-4)
5. Since the category column can take country-specific values, i.e. the names of subnational units, the resulting variable column could be too large in many cases. To solve this issue, each subnational unit is assigned with a correlative number. For instance, the ‘Amazonas’ department in Peru would be referred as ADM1\_PER1. [↑](#footnote-ref-5)
6. It would not be possible to upload all the databases to the "SSGD v2.0" folder since it would imply a violation of the use-of-data policies. [↑](#footnote-ref-6)