Inequality Transparency Index Update

Francois Burq
Lucas Chancel

November 2020

World Inequality Lab

Inequality transparency index update
Technical Note[1]

François Burq[2] and Lucas Chancel [3]

November 2020

1 Introduction

The past decades have been witness to important developments in the measurement of income and wealth inequality. However, available information on income and wealth distributions remains particularly scarce across the world. The opacity of the financial system, the types of tools used by statistical administrations to track inequality and sometimes the reluctance of governments to publish data they have in hand still make it particularly difficult to know which groups of the population benefit from economic progress.

   The World Inequality Database combines the best available sources, namely household surveys, tax data, national accounts (and, when available information from financial leaks, such as the "Panama papers") in a systematic and transparent way in order to publish inequality series for most countries (See

[2]François Burq is research assistant at the World Inequality Lab (francois.burq@psemail.eu).
[3]Lucas Chancel is co-director of the World Inequality Lab (lucas.chancel@sciencespo.fr).

the World Inequality Report 2018 for a longer discussion on data quality issues). Inequality series published on WID.world constitute an improvement from official statistics as they include more information (particularly at the top of the distribution). They however remain imperfect.

In order to improve existing series, more data will have to be released by statistical agencies and governments. In order to give a sense of the road ahead, WID.world publishes an inequality transparency index — an evolutive and collaborative tool describing the availability and quality of information on income and wealth inequality in a given country.

This note describes the construction of the index.

2   Objectives

The objectives for the index are twofold. First, it produces an assessment of the state of inequality data throughout the world. Secondly, it should also work as an incentive for future improvements in data production and publication. We can see that numerous countries still produce very little data (tax data, in particular is extremely rare). Even in countries which produce yearly data on wealth and income (sometimes including tax data), we must insist on the fact that the quality of this data is far from the ideal 20/20 situation. In particular new automatic transmission of banking information in OECD countries has not yet proven to significantly enhance wealth data production. All in all, we are still waiting for OECD countries to produce distributional accounts (see DINA guidelines). Finally, we also hope that this index will help countries to take step in publishing transparent data and allowing easier access for researchers. For the moment, we encounter numerous situations where the data exists but is so hard to access that it is almost unavailable.

3   Calculating the index

The index is constructed around two dimensions:

- In the first dimension we differentiate between four different sources of data : income surveys, income tax data, wealth surveys and wealth tax data.

- In the second dimension, we evaluate various components for each of these sources : quality, frequency of publication and access to the data.

The total grade out of 20 can be decomposed into two grades out of 10 for income and wealth data respectively. The maximum grade of 20/20 corresponds to an ideal situation where countries would publish yearly distributional accounts on wealth and income. As of 2020, we are still very far from this situation and the maximum grade that we find is 16.5/20.

3.1 Type of data

We distinguish between four types of data: income surveys, income tax data, wealth surveys and wealth tax data. While each of these sources have their interests we stress that a proper measurement of basic inequality indicators (Top shares, Ginis, Quantile ratios) requires particularly good income and wealth tax data (or a good integration of income and wealth surveys with administrative data). We therefore give a slightly stronger weight on administrative tax data than on survey data.

We also place slightly more weight on wealth tax data because it is essential to properly measure capital incomes and hence to properly track the dynamics of income inequality. Indeed, in several countries, income inequality dynamics have been driven by (non-fiscal) capital incomes over the past decades, which are often missing from income and wealth surveys, as well as from income tax data.

For simplicity, we grade each of these data sources out of 10 points and apply weights corresponding to the importance that we attribute to a given data source. Income and wealth are weighted by 0.4, income tax data is weighted by 0.5 and wealth tax data is weighted by 0.7. We detail the attribution of these 10 points below.

3.2 Frequency, precision, quality and access

3.2.1 Frequency

Frequency captures the number of surveys or years of tax data that were produced and made available in the last ten years. It is given a grade out of 2 :

- 0.5 point if there is one survey

- 1 point if there are 2 or 3 surveys

- 1.5 points if there are between 4 and 7 surveys

- 2 points if there are more than 8 surveys

In an ideal situation, we would be able to access data from 2019 when updating our index in 2020. The ten years limit for data availability therefore takes into account the data that has been produced since 2009.[4]

3.2.2 Precision

Microdata is very important for the production of distributional accounts as it opens a variety of options for statistical analyses (making it possible to check the series' consistency, match the datasets with others, etc.). We give up to

---

[4]In order to avoid penalizing the few countries that have published wealth tax data 15-12 years ago.

2 points for microdata availability. These 2 points measure the availability of microdata and not its quality which is graded separately (see below). Full points are awarded when microdata is available whenever data is published. If, for some reason, microdata is not available for some years, we give 1 or 0.5 points. Consequently, when the number of surveys in the last ten years is below three, we only give up to 1 point as we cannot check that microdata is systematically available.

3.2.3 Quality

We evaluate the quality of each data source out of 5 points. For each country, the quality score is attributed by the World Inequality Lab regional coordinators. When considering data for which only tabulations are available, we give a maximum grade of 2.5. When microdata is considered, the maximum grade is 5. Microdata quality can be given a grade lower than 2.5 because points were already given to take into account microdata availability. In certain cases, standard grades are applied and are presented below:

- PovcalNet consumption tabulations are given a grade of 1

- PovcalNet income tabulations are given a grade of 1.5

- EU-SILC income surveys are given a grade of 3

- HFCS wealth surveys are given a grade of 3

Users should refer to the methodology of these specific datasets in order to obtain more precise information on the quality of these sources. These quality grades are meant to be updated gradually while we get feedback from researchers working with the data.

3.2.4 Access

This component captures additional difficulties to access the data when it is available. This is particularly relevant for income and wealth tax microdata for which on site presence or nationality restrictions can be in place. When only tabulations are available, we automatically give 0 points for access. For microdata, we give 0 if there are specific restrictions to access the data and 1 otherwise. This score is most of the time 0 when considering tax microdata and almost always 1 when considering survey microdata. When data is so hard to access that it can almost be considered unavailable, we reduce the score given in the "Microdata" component as we can assume that microdata is not de facto available for each year.

4 Main results

Among the 181 countries that we include in the index, we can very broadly distinguish between 5 groups of countries :

- A group of 28 countries with a score of 0. These countries have not produced any data in the last 10 years. This is the case for 5 African countries, 11 Latin-American countries and 6 Middle East/Asian countrie and Papua New Guinea

- A second group of 73 countries with scores between 0.5 and 2.5. These countries typically only produce income survey tabulations (with varying quality and frequency) that we access through PovcalNet. In this group, we find most of the remaining African countries (47 countries), some Middle East/Asian countries (20)and a few Eastern European countries (6).

- 34 countries with scores between 3 and 5. These countries have more detailed income data then the aforementioned countries, but still obtain a score of 0 for wealth. Improvements in income data can include access to income survey microdata and the publication of tax data (often tabulations). In this group, we can find 3 African countries (Egypt, Central African Republic and Cote d'Ivoire), 9 American countries, 2 eastern European countries, the bulk of Asian countries (18) and 2 countries from Oceania (Fiji and Solomon Islands).

- A fourth group includes 38 countries with scores ranging from 5.5 to 12.5. These countries publish a wealth survey on top of some income data, however they do not obtain scores higher than 12 because some data is still missing (administrative wealth data in most cases). Those are the remaining Asian countries (4), 7 American countries, 2 countries from Oceania (Australia and New Zealand), South Africa, and 24 European eastern and western countries.

- 8 countries have scores between 13 and 16.5. These countries publish all four types of data (income survey, income tax, wealth survey and wealth tax data). Differences in scores as well as the 16.5 maximum grade are due to some issues in the quality of this data. These countries are Denmark, Italy, Sweden, Uruguay, France, the United-States,the United Kingdom and Norway. We stress that even for these countries, there is still a fair amount of progress to be made in order to reach transparent inequality statistics.

5 Improving inequality data transparency

Because the index encompasses such a wide range of countries, these results can be complex to interpret. Countries with very low scores could improve their

net income bracket ($)	number of individuals	total income	labor income	capital income	incl. housing asset income	incl. equity asset and net interest income	incl. pension and life insurance asset income	total income taxes	incl. personal income tax	incl. corporate income tax	incl. capital gains tax	total wealth taxes	incl. wealth and property tax	incl. inheritance and estate tax
0–10k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
10k–20k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
20k–30k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
30k–40k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
40k–50k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
50k–70k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
70k–100k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
100k–150k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
150k–200k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
200k–400k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
400k–600k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
600k–800k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
800k–1m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1m–10m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
10m–100m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
>100m	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Figure 1: Income inequality data to be reported by authorities by income bracket

Source: DINA Guidelines 2020

data quality by producing more data. Specifically, improved transparency of tax data in conjunction with a systematic conduction of wealth surveys would drastically improve data quality for numerous countries. On the other hand, as the score gets higher, important areas of improvement move to the quality component : even when every type of data is available, numerous improvements can be expected. Among others, we expect improvements in quality such as an increased frequency of data publication, more transparency in tax data access, a systematic estimation of wealth levels (and in the cases where wealth tax is non existent, a proper administrative wealth tax estimation). In general, inequality data remains completely insufficient as of 2020. Even in countries which obtain the higher grades, there is an urgent need for improvement if we are to obtain precise distributional national accounts.

We also stress that data on the automatic exchange of bank information (under the aegis of the OECD) is still unavailable to researchers or the general public at this stage, making it impossible to properly and systematically assess any progress made in the context of the fight against tax evasion and illicit financial transactions. As a result, there are still many unknowns about the exact inequality levels (in particular at the top of the distribution, where evasion if found to be the highest, see World Inequality Report 2018, section 5.2) in all countries, including those with seemingly high grades. In that regard, all countries, including those with high grades in 2020 are still lagging behind basic transparency standards.

In order to obtain the maximum grade, countries would need to publish the following tables (and in particular Table 3)

net wealth bracket ($)	number of individuals	total wealth	incl. currency and deposits	incl. bonds and loans	incl. equities and fun shares	incl. pension funds and life insurance	incl. real estate	incl. business and other non-financial assets	incl. debt	incl. total domestic assets	incl. total foreign assets	total income	incl. capital income	incl. labor income
<0	…	…	…	…	…	…	…	…	…	…	…	…	…	…
0–10k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
10k–100k	…	…	…	…	…	…	…	…	…	…	…	…	…	…
100k–1m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1m–10m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
10m–100m	…	…	…	…	…	…	…	…	…	…	…	…	…	…
100m–1bn	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1bn–5bn	…	…	…	…	…	…	…	…	…	…	…	…	…	…
5bn–10bn	…	…	…	…	…	…	…	…	…	…	…	…	…	…
>10bn	…	…	…	…	…	…	…	…	…	…	…	…	…	…

Figure 2: Wealth inequality data to be reported by authorities by wealth bracket

Source: DINA Guidelines 2020

net wealth bracket ($)	number if individuals	incl. number of residents	incl. number of non-residents	total net wealth	incl. residents	incl. non-residents	total wealth taxes	incl. wealth and property tax	incl. capital gains tax	incl. inheritance and estate tax	total income taxes	incl. personal income tax	incl. corporate income tax
<0	…	…	…	…	…	…	…	…	…	…	…	…	…
0–10k	…	…	…	…	…	…	…	…	…	…	…	…	…
10k–100k	…	…	…	…	…	…	…	…	…	…	…	…	…
100k–1m	…	…	…	…	…	…	…	…	…	…	…	…	…
1m–10m	…	…	…	…	…	…	…	…	…	…	…	…	…
10m–100m	…	…	…	…	…	…	…	…	…	…	…	…	…
100m–1bn	…	…	…	…	…	…	…	…	…	…	…	…	…
1bn–5bn	…	…	…	…	…	…	…	…	…	…	…	…	…
5bn–10bn	…	…	…	…	…	…	…	…	…	…	…	…	…
>10bn	…	…	…	…	…	…	…	…	…	…	…	…	…

Figure 3: Total taxes paid by wealth bracket to be reported by authorities

Source: DINA Guidelines 2020

6   Future developments

We think that the current index can help assess the relative quality of inequality data across the globe. The index itself can be improved in the future. We discuss below some possible developments.

6.1   National Accounts

The 2020 update does not yet take into account the quality of national accounts. This is a very important component that should be included in next year's update. To include these elements into the score system, we could take into account the frequency of publication and the level of aggregation of national accounts.

6.2   Quality grade

On top of updating the quality grades with new information, a nice improvement would be to define a precise scale for the criteria of this component. For now, we give global quality grades out of 5 which take into account numerous questions, but we have not yet defined precisely the weight that they should have within the quality component. This scale could include questions such as :

- How comprehensive is the definition of income ? Are all type of income taken into account ?

- Is the study representative at the national level ?

- Is there a link between survey and fiscal data ?

- Is the data comparable across years for this country ?

Such a detailed scale requires extremely precise knowledge of each specific survey and can only be implemented gradually.

Overall, this 2020 updates aims to set the framework for a comprehensive index allowing for a precise assessment of the state of data on inequality. We insist on the fact that all the grades are not final and will be updated when better data is available or when we are made aware of specific issues for some countries. In the years to come we hope to see a general increase in the scores of this index, reflecting improvement in the quality and availability of data that we can use.