

13 Indicators to present distributional results

Step 5 concerns the derivation of indicators to present the distributional results. This chapter first discusses how to ensure comparability across household groups of different size and composition when publishing the results. It then highlights some examples of indicators that may be used to present the distributional results.

13.1. Introduction

At the end of the process, when the data have been aligned to the national accounts totals and households have been clustered into relevant household groups, results can be presented for these household groups and indicators can be derived to show the degree of disparity between them. To arrive at comparable results across household groups, the results are often recalculated on the basis of the number of households or the number of consumption units per household group. It depends on the (policy) use what type of results will be preferred. This chapter discusses these two concepts in more detail and provides some examples of indicators that may be used to present the distributional results.

Section 13.2 first explains how to derive results in terms of consumption units and in terms of households. Section 13.3 then describes disparity ratios that are currently used in the DNA work. This is followed by additional overviews that may be used to present the results in Section 13.4. Finally, Section 13.5 presents some additional indicators that are often used in inequality analyses which are based on underlying micro data. Depending on the level of detail available from the calculations, these may also be used by compilers to show distributional results.

13.2. Presentation of data on “per household” or “per consumption unit” basis

The various breakdowns as targeted in the DNA work provide information on results for various household groups. To arrive at comparable results across household groups and to be able to conduct comparisons over time and across countries, the results are often presented per household and per consumption unit, which take into account differences in the number of households and their composition across household groups.

Per household results can be derived by dividing the amounts for the specific household group by the number of households in that group. This would show the average value per household in the group. For a given household group and component, the average measure (\bar{X}) per household (hh) is computed as follows:

$$\bar{X}_i^{NA_adj}_{hh} = \frac{X_i^{NA_adj}}{n_i}$$

using the notation:

X : income/consumption component

i : {1, 2, ..., l} to identify household groups

n_i : total number of households in group i

$X_i^{NA_adj}$: adjusted national accounts subtotal for group i

Per consumption unit results can be derived by dividing these amounts by the number of consumption units (cu), showing the (equivalized) value for a person in that household group, on the basis of the following formula:

$$\bar{X}_i^{NA_adj}_{cu} = \frac{X_i^{NA_adj}}{cu_i}$$

using the notation:

X : income/consumption component

i : {1, 2, ..., l} to identify household groups

cu_i : total number of consumption units in group i

$X_i^{NA_adj}$: adjusted national accounts subtotal for group i

As was explained in Box 2.1 in Chapter 2 the number of consumption units reflects the consumption needs for households of different size, taking into account that consumption needs of a household will increase with each additional household member, but not in a proportional way due to economies of scale. For the purpose of the DNA work, the OECD-modified equivalence scale has been chosen as reference method which assigns a value of 1 to the first adult in the household, a value of 0.5 to each additional person aged 14 and over, and 0.3 to all children under 14. Depending on country specific situations compilers may also decide to use a different equivalence scale. In any case, in presenting the results, one always has to specify which equivalence scale has been used.

Results may also be calculated on a “per capita” basis. This may be considered as a specific application of the per consumption unit calculation, applying a value of 1.0 to all household members.

13.3. Measures of disparity

Disparities across households can be analysed on the basis of three main ratios. All of these concern relative ratios, which helps in cross-country comparisons and to analyse trends over time.

The first ratio is the *ratio to the average*, which shows the value of income and consumption for each household group relative to the average household value. It is computed as follows for household group i:

$$\text{Ratio to average}_i = \frac{\bar{X}_i^{NA_adj}}{\bar{X}^{NA_adj}}$$

using the notation:

X : income/consumption component

z : {*EDI, MSI, HT*}: identifies the household classification variable, i.e. equivalised disposable income, main source of income and household type

i : {1, 2, ..., l} to identify household groups

n_i : total number of households in group i

N : total number of households in the population

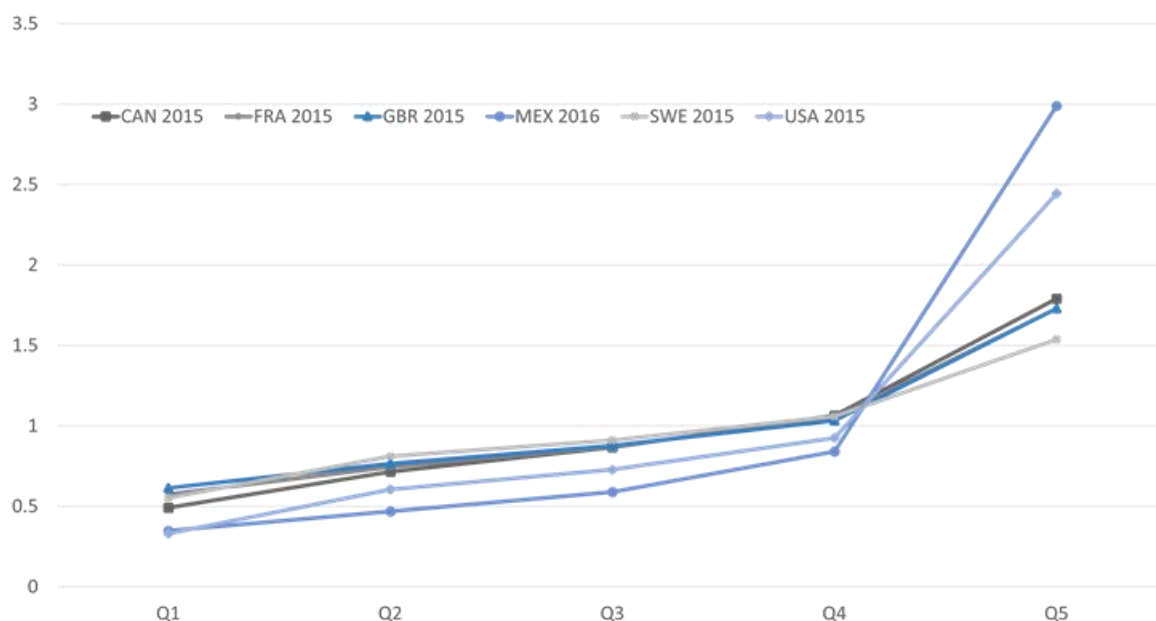
$\bar{X}_i^{NA_adj}$: per household or per consumption unit adjusted national accounts for group i

\bar{X}^{NA_adj} : per household or per consumption unit adjusted national accounts

When the ratio to the average is compiled on the basis of ordinal scales (e.g. a breakdown according to income quintile) the results can be presented as a line connecting the ratios for the various household groups (see Figure 13.1). If the ratio is compiled for non-ordinal scales (e.g. into main source of income) the results are usually presented in the form of bar charts.

Figure 13.1. Example of presentation of ratio to the average for an ordinal breakdown into household groups

Relative position of each household group compared to the average, by equivalized disposable income quintile



Note: The results show adjusted disposable income per consumption unit for each group to the average adjusted disposable income per consumption unit for the private household sector as a whole.

Source: Zwijnenburg et al. (2021^[1]).

The ratio of the *highest to lowest* shows the value of income and consumption for the highest household group to the lowest household group value. It is computed as follows for a given classification of household z (e.g. equivalized disposable income quintile; main source of income; and household type):

$$\text{Ratio highest to lowest}_z = \frac{\text{Max}_{i \in z} \{\bar{X}_i^{NA_{adj}}\}}{\text{Min}_{i \in z} \{\bar{X}_i^{NA_{adj}}\}}$$

using the notation:

X : income/consumption component

z : $\{MSI, EDI, HT\}$: identifies the household classification variable

i : $\{1, 2, \dots, l\}$ to identify household groups

$\bar{X}_i^{NA_{adj}}$: per household or per consumption unit adjusted national accounts for group i

This ratio is often used to make cross-country comparisons in which case the results are presented in a bar chart. However, it may also be used to monitor changes over time within a country, in which case results could be presented in the form of a line chart.

As a third measure, the coefficient of variation is taken as a disparity index that shows the variation from the average. For a given classification of households (e.g. equivalised disposable income quintile; main source of income; and household type), the coefficient of variation (CV) is the ratio of the standard deviation to the mean calculated as follows:

$$\text{Coefficient of variation}_z = \frac{\sqrt{\frac{1}{N} \times \sum_{i \in z} [n_i \times (\bar{X}_i^{NA,adj} - \bar{X}^{NA,adj})^2]}}{\bar{X}^{NA,adj}} \times 100$$

using the notation:

X : income/consumption component

z : {*MSI, EDI, HT*}: identifies the household classification variable

i : {1, 2, ..., l} to identify household groups

n_i : total number of households in group i

N : total number of households in the population

$\bar{X}_i^{NA,adj}$: per household or per consumption unit adjusted national accounts for group i

$\bar{X}^{NA,adj} = \frac{1}{N} \sum_{i \in z} n_i \times \bar{X}_i^{NA,adj}$: per household or per consumption unit adjusted national accounts.

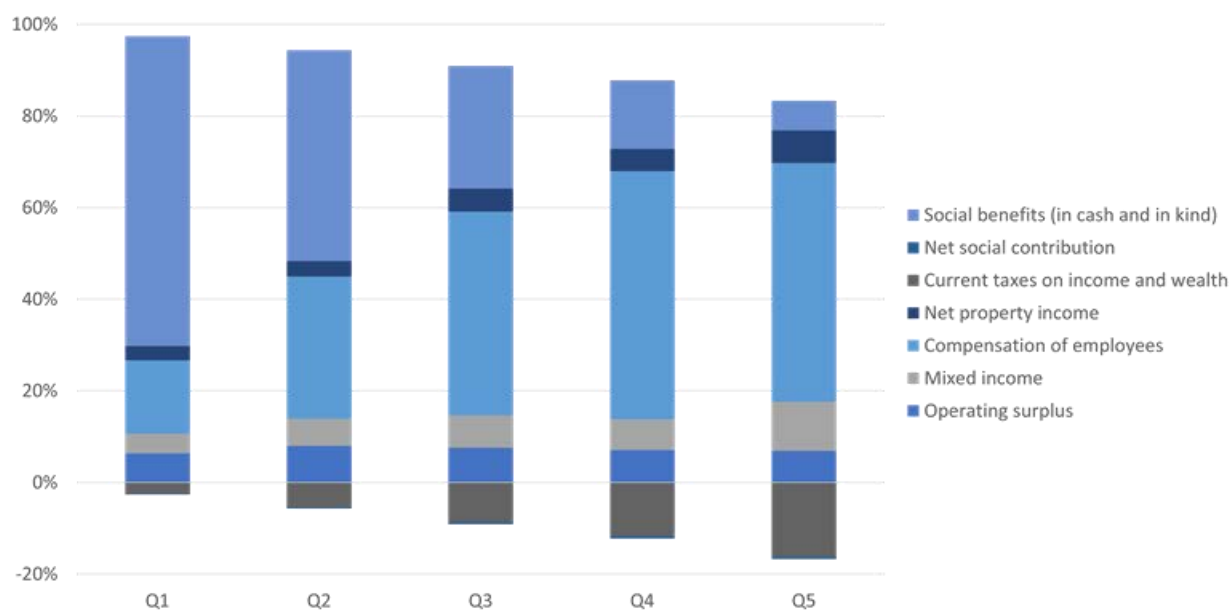
As was the case with the ratio highest to lowest, this ratio may be used to make cross-country comparisons as well as to monitor changes over time within a country. When analysing the disparity index, two properties need to be taken into account. First, the above calculation assumes that each household receives (or spends) the average income (or expenditures) of his group, i.e. disparity within a household group is supposed to be zero, implying that the disparity index underestimates household disparities. This is particularly important for household groups that do not rely on income level as the disparities in income and consumption within the group may be quite large. Second, the results for the above disparity index depend on the household structure in each country. Consequently, divergences in coefficient of variations between two countries may be explained by two factors: differences across countries in the extent to which one given household group departs from the average; and cross-country differences in the share of the household groups in the total household population. This has to be borne in mind when conducting cross-country comparisons.

13.4. Composition of household income and consumption

In addition to focusing on disparities in income and consumption levels between groups of households, the results may also be used to assess differences in composition of income and consumption. It may for example provide information on the main sources of income for different household groups and on their main consumption categories, which may be of interest to assess how vulnerable specific groups may be to changes in certain types of income (for example due to a change in interest rate) or to changes in prices for specific consumption categories.

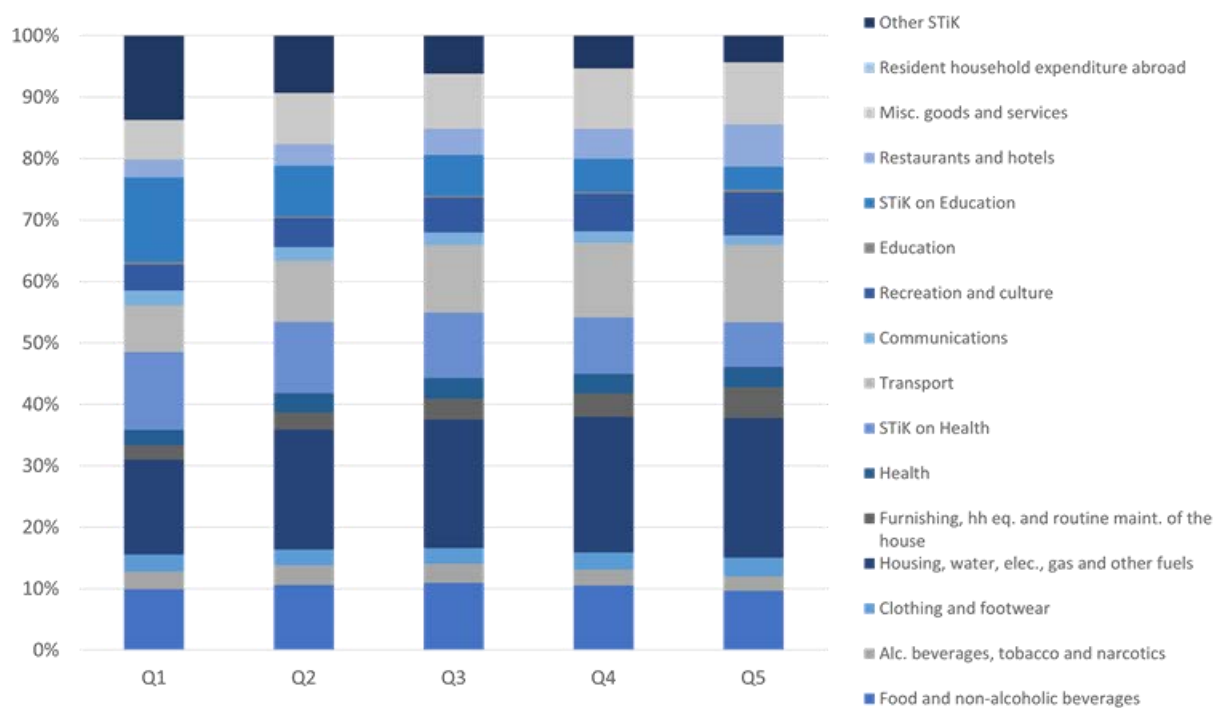
The most common way to present these results is to look at the share of the various items in total disposable or adjusted disposable income, and in total final consumption expenditure or actual final consumption. It will depend on the available underlying information what level of detail could be provided. Figure 13.2 and Figure 13.3 show examples on how to present this information on the basis of DNA results collected in 2020 (Zwijnenburg et al., 2021^[11]).

Figure 13.2. Example of composition of adjusted household disposable income per quintile



Source: Zwijnenburg et al. (2021^[1]).

Figure 13.3. Example of composition of actual final consumption per quintile



Source: Zwijnenburg et al. (2021^[1]).

13.5. Indicators based on underlying micro data

Several indicators used in distributional analyses focus on the underlying micro data instead of on aggregated results (see for example Cowell (2011^[2])). For example, some aim to provide insight into the share of households that meet a certain criterion (e.g. disposable income below a certain threshold to assess the number of households in poverty) to have a better understanding of how many households are in a certain situation or may be affected by a specific event. Others may focus on very granular levels of detail to derive inequality measures such as the Gini coefficient. As these indicators are derived on the basis of very granular data, their reliability will very much depend on the level of detail and the quality of the underlying data.

As the aim of the work is to derive distributional results for aggregated groups of households, it will often not be possible nor desirable to derive these types of indicators for the DNA results. In that regard, it has to be borne in mind that the process to compile distributional results in line with national accounts totals often involves several assumptions to allocate imputations for missing elements to the relevant households and to align micro data to the national accounts totals. Whereas usually sufficient information may be available to properly allocate these amounts at an aggregated level, this will often become more complicated at more granular levels of detail. In those cases, one should carefully assess whether it would still be opportune to publish at these more granular levels or to derive indicators on the basis of these detailed results, as the results may be highly sensitive to specific assumptions in the compilation process. Only in the case that input data is available at a very granular level of detail and the impact of assumptions is deemed to be relatively small, compilers may decide to publish on the basis of these detailed results. Below, some examples are presented of results that may be derived in that case.

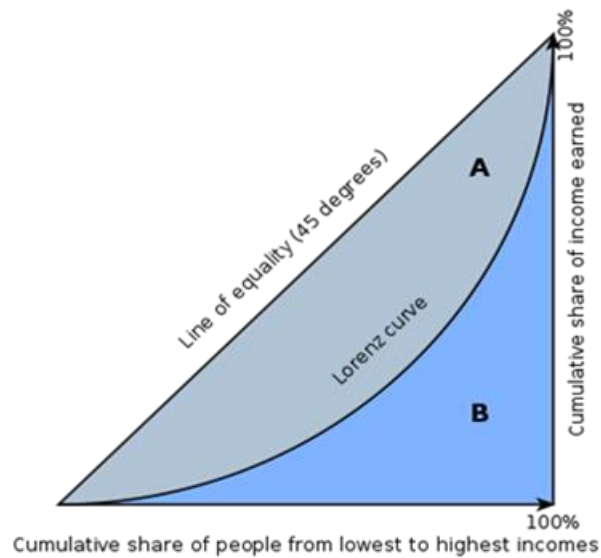
A first indicator that could be derived from underlying micro data is the household participation rate. This ratio provides insight into the share of households that report a value for a specific item. For example, the participation rate for distributed income of corporations shows how many households benefit from this type of income. It can be derived by simply dividing the number of households that report a specific item by the total number of households in a specific household group. This type of information would not only provide interesting information for users, but also for compilers to check the plausibility of the results, analysing the changes in the participation ratio for various items over time.

A second indicator type that could be derived on the basis of the underlying micro data is a measure that specifies the share of households above or below a certain threshold. An example of such a measure is the poverty measure which looks at the number of people with an income below a certain threshold. If data aligned to national accounts totals are available at a sufficient level of detail, percentages for these thresholds may be derived from the underlying micro data.

A third measure is the median which looks at the value of the household that is in the middle of the distribution (or of a specific household group). The benefit of looking at the median value is that it is less skewed in relation to extreme values, whereas the mean value may be more affected by long tails at either end of the distribution. If it is possible to derive a plausible median value, this would also provide the opportunity to calculate the ratio of the mean to the median, which will provide more insight into the skewedness of the distribution. The ratio will be higher than 1 when the mean is higher than the median, reflecting that most households have an income below the mean.

Finally, on the basis of micro data, it would also be possible to derive a Lorenz curve as well as the corresponding Gini coefficient. A Lorenz curve for income is created by ranking households (or individuals) from the poorest to the richest and plotting the cumulative share of household income and the cumulative share of the number of households, as proportions of the total household income and the total number of households, respectively. When every household group has the same income the Lorenz curve would be a 45-degree line. Figure 13.4 illustrates an example of a Lorenz curve as applied to a given income item.

Figure 13.4. Example of Lorenz curve

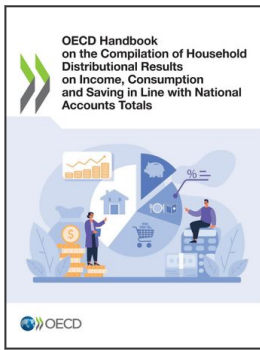


Source: Wikipedia (2023_[3]).

The Gini coefficient is a summary measure of income (or wealth) dispersion in the population that is derived from the Lorenz curve. Gini coefficients are scaled from 0 to 100 per cent, with a value of 0 indicating perfect equality and a value of 100 indicating that one household or individual has all the income. In the graph above it is equal to area A divided by area A plus B.

References

- Cowell, F. (2011), *Measuring Inequality*, Oxford University Press, [2]
<https://doi.org/10.1093/ACPROF:OSOBL/9780199594030.001.0001>.
- Wikipedia (2023), *Lorenz curve*, https://en.wikipedia.org/wiki/Lorenz_curve (accessed on [3]
21 November 2023).
- Zwijnenburg, J. et al. (2021), "Distribution of household income, consumption and saving in line [1]
with national accounts: Methodology and results from the 2020 collection round", *OECD
Statistics Working Papers*, No. 2021/01, OECD, Paris, [https://www.oecd-
ilibrary.org/docserver/615c9eec-
en.pdf?expires=1669202191&id=id&accname=ocid84004878&checksum=689A9023E81B32
F43D91EF34FBDE7A7A](https://www.oecd-ilibrary.org/docserver/615c9eec-en.pdf?expires=1669202191&id=id&accname=ocid84004878&checksum=689A9023E81B32F43D91EF34FBDE7A7A) (accessed on 23 November 2022).



From:

OECD Handbook on the Compilation of Household Distributional Results on Income, Consumption and Saving in Line with National Accounts Totals

Access the complete publication at:

<https://doi.org/10.1787/5a3b9119-en>

Please cite this chapter as:

OECD (2024), "Indicators to present distributional results", in *OECD Handbook on the Compilation of Household Distributional Results on Income, Consumption and Saving in Line with National Accounts Totals*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/7df7fa3e-en>

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.