



Income Distribution Database (<http://oe.cd/idd>)

## **TERMS OF REFERENCE**

### **OECD PROJECT ON THE DISTRIBUTION OF HOUSEHOLD INCOMES**

The OECD income distribution questionnaire aims at collecting each year a basic set of indicators on income inequality and poverty to support OECD comparative analysis. The present version of the Terms of Reference retains all the key assumptions used in 2015/16 for the computation of the OECD indicators.

## 1. Main Definitions

### Reference units, equivalence scale and adjusted income

<b>Observation Unit</b>	<p>The unit of observation of the survey is the <b>household</b>.</p> <p>A household is either an individual person or a group of persons who live together under the same housing arrangement and who combine to provide themselves with food and possibly other essentials of living. [This is the definition recommended in the 2011 <i>Canberra Handbook</i>: countries departing from this definition are asked to indicate so in the metadata sheet].</p>
<b>Reference unit for income distribution indicators</b>	<p>All income distribution indicators refer to <b>persons</b>.</p> <p>In the distribution, each household is weighted by the number of individuals who belong to this household. For instance, a household of four people has a weight equal to four; this is equivalent to considering a distribution in which this household is represented by four individuals with the same level of income.</p>
<b>Equivalence scale</b>	<p>All the tables specified in this document should be calculated using an <b>equivalence elasticity of 0.5</b>.</p> <p>This means that the all income components of each household are adjusted by the square root of the household size. For instance, the income of a household with four persons should be divided by two and then attributed to the four members of the household (see <a href="http://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf">http://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf</a>).</p> <p>The equivalence elasticity (<math>\epsilon</math>) characterises the amount of scale economies that households can achieve. An equivalence elasticity lower than unity implies the existence of economies of scale in household needs, i.e. any additional household member needs a less than proportionate increase of household income in order to maintain a given level of welfare. Under this assumption, the sum (across the <math>j</math> members of the same household <math>i</math>) of individual "adjusted" incomes <math>DI_{ij}</math> will exceed the total household disposable income by the amount of scale economies.</p>
<b>Adjusted disposable income</b>	<p>Individuals are ranked according to the value of the "adjusted" disposable income per equivalent household member of the household to which they belong.</p> <p>For instance, if <math>Y_i</math> denotes the total disposable income of household <math>i</math>, the "adjusted" income of each member <math>j</math> of household <math>i</math> (<math>DI_{ij}</math>) is calculated as following:</p> $DI_{ij} = Y_i / S_i^\epsilon$ <p>where <math>S_i</math> is the number of members in household <math>i</math> and <math>\epsilon</math> is the equivalence elasticity.</p>

### *Income components, disposable, market and primary income*

Income distributions refer to a particular year, which should be indicated in the Excel spreadsheet "Metadata". All income components should be reported on an *annual basis and in nominal prices*. Five main components of household disposable income are identified in the OECD questionnaire:

- **E:** employee income, including wages and salaries, cash bonuses and gratuities, commissions and tips, directors' fees, profit sharing bonuses and other forms of profit-related pay, shares offered as part of employee remuneration, free and subsidised goods and services from an employer, severance and termination pay.<sup>1</sup> Sick pay paid by social security should also be included.

---

<sup>1</sup> The elements detailed (for each of the five income variables) are those included in the conceptual definition of household disposable income of the 2011 *Canberra Handbook* and that most OECD countries appear to collect in their micro-sources. See <http://www.unece.org/stats/groups/cgh.html> Countries that do not cover some of these detailed components in their source should indicate so in the metadata sheet of the OECD questionnaire.

- **KI<sup>2</sup>**: capital and property income, including income from financial assets (net of expenses), income from non-financial assets (net of expenses) and royalties. Regular receipts from voluntary individual private pension plans and life insurance schemes should also be included in this income component. In line with the *2011 Canberra Handbook*, capital gains should not be included in KI.
- **SEI<sup>3</sup>**: income from self-employment, including profits and losses from unincorporated enterprises, as well as goods produced for own consumption (net of the costs of inputs). [The inclusion of this latter variable aims to adjust the OECD income concept to the realities of middle-income countries (such as Brazil, South Africa and others), where subsistence agriculture represents a significant income source for people at the bottom of the distribution. Countries that do not collect information on this income item should indicate so in the metadata sheet of the OECD questionnaire].
- **TRR**: current transfers received, including transfers from social security (including accident and disability benefits, old-age and survivor cash benefits, unemployment benefits, maternity allowances, child and/or family allowances, all income-tested and means-tested benefits that are part of social assistance, including quasi-cash transfers given for a specific purpose such as food stamps); transfers from employment related social insurance; as well as cash transfers from both non-profit institutions and other households.
- **TRP**: current transfers paid, including direct taxes on income and wealth, social security contributions paid by households, contributions to employment-related social insurance, current transfers paid to both other households and non-profit institutions. Taxes on realised capital gains should be excluded from wealth taxes when possible. [Please report values for transfers paid in the OECD questionnaire with a negative sign in order to reflect that, differently to the other income components, they should be deducted].

For four of these components, a more detailed breakdown is also requested:

- In the case of employee income (**E**):
  - **EH**: the wage and salary income of the household head, excluding employers' contributions to social security, but including sick pay paid by social security.
  - **ES**: the wage and salary income of the household head spouse or partner, excluding employers' contributions to social security, but including sick pay paid by social security.
  - **EO**: the wage and salary income from other household members, excluding employers' contributions to social security, but including sick pay paid by social security.
- In the case of self-employment income (**SEI**):

---

<sup>2</sup> Please note that this definition of capital and property income differs from the definition used in former ToR ("K") insofar as it does no longer include transfers received from compulsory employment-related occupational pension schemes.

<sup>3</sup> Please note that this definition of self-employment income differs from the definition used in former ToR("SE") insofar as it adds the value of goods produced for own consumption.

- **SE**: Profits and losses from unincorporated enterprises.
- **OC**: income from goods produced for own consumption.
- In the case of current transfers received (**TRR**):
  - **TRRSS**: current transfers received from social security.
  - **TRRER**: current transfers received from employment-related social insurance schemes (e.g. occupational pensions), where such schemes meet at least one of the following conditions: i) participation is obligatory; ii) the scheme is collective; and iii) the employer makes a contribution on behalf of an employee.<sup>4</sup>
  - **TRROT**: current transfers received from non-profit institutions and other private households, e.g. alimonies.
- In the case of current transfers paid (**TRP**):
  - **TA**: direct taxes on income and wealth paid by households (net of refunds), as well as contributions paid by households to public social security schemes.
  - **TRPER**: contributions paid by households to employment-related social insurance schemes (as defined above).
  - **TRPOT**: current transfers paid by households to non-profit institutions and other households, e.g. alimonies.

While relevance and data availability for the sub-components of current transfers will vary across countries (depending on the structure of their social protection system and on features of their micro-data), this more detailed breakdown allows better reflecting the situation of countries with an important employment-related pension pillar.

All household income components are expressed in terms of equivalent household member, dividing the component by  $S_i^\varepsilon$ , i.e. the number of household members to the power of the equivalence elasticity  $\varepsilon$ .

The income components defined above can be aggregated into various concepts of equivalised household income. Individual **primary, market, gross and disposable** income per equivalent household member, for each member  $j$  of household  $i$ , can be expressed as follows:

---

<sup>4</sup> Note that regular receipts from voluntary individual private pension plans (i.e. non employment-related) should be included in KI (i.e. treated as capital income).

$$[1] \text{ Equivalised primary income: } PI_{ij} = E_{ij} + KI_{ij} + SEI_{ij} + (TRROT_{ij} - TRPOT_{ij})$$

$$[2] \text{ Equivalised market income: } MI_{ij} = PI_{ij} + TRRER_{ij}$$

$$[3] \text{ Equivalised gross income: } GI_{ij} = MI_{ij} + TRRSS_{ij} - TRPER_{ij}$$

$$[4] \text{ Equivalised disposable income: } DI_{ij} = GI_{ij} - TA_{ij}$$

$$DI_{ij} = E_{ij} + KI_{ij} + SEI_{ij} + TRR_{ij} - TRP_{ij} =$$

$$= (EH_{ij} + ES_{ij} + EO_{ij}) + KI_{ij} + (SE_{ij} + OC_{ij}) + (TRRSS_{ij} + TRRER_{ij} + TRROT_{ij}) - (TA_{ij} + TRPER_{ij} + TRPOT_{ij})$$

In equation [1], **primary income** includes earnings, self-employed income, capital income and the balance between the transfers *received* by each household from non-profit institutions and other households (TRROT<sub>ij</sub>) and the transfers *paid* by each household to non-profit institutions and other households (TRPOT<sub>ij</sub>). In equation [2], **market income** adds to primary income the value of employment-related social insurance transfers received by households (TRRER<sub>ij</sub>). In equation [3], **gross income** adds to market income the current transfers received from social security schemes (TRRSS<sub>ij</sub>) and deducts the value of transfers paid to employment-related social insurance schemes (TRPER<sub>ij</sub>).<sup>5</sup> In equation [4], **disposable income** deducts from gross income the value of taxes on income and wealth paid and of contributions paid by households to public social security schemes (TA<sub>ij</sub>).

The main income concept used in tables 1 through 3 is [4], disposable income (DI). Tables 1 and 3 also ask for Gini coefficients and poverty rates according to concept [2], market income (MI), or “before taxes and transfers”. In addition to these two main income concepts, table 1 also asks for Gini coefficients according to concept [1], primary income (PI) and according to gross income (GI) [3], i.e. disposable income “before tax”.

The income concepts described above provide the basis for computing the Gini coefficients to be reported in Table 1: in each case, individuals should be ranked in increasing order of the relevant income concept before computing Gini coefficients. Individuals should thus be ranked by *DI* in row 9, by *MI* in row 11, by *GI* in row 12 and by *PI* in row 13.

### ***Treatment of negative income***

Special treatments should be applied to ensure that values of household income, for each household and individual, are non-negative. All the treatments described below should be applied once the adjustment to equalise household income (using the equivalence elasticity of 0.5) has been applied.

<sup>5</sup> The variable TRPEPR is deduced when computing gross income, rather than when defining market income, in order to preserve comparability with the definition of market income used in previous versions of the Terms of Reference (which did not consider current transfers paid by households other than taxes).

First, household-level values of EH, ES, EO, KI, SE, OC, TRRSS, TRRER and TRROT showing negative values should be set to zero. For instance, any negative value of self-employment income should be set equal to zero before computing the income of each household.<sup>6</sup>

Second, in those cases where, even after setting to zero EH, ES, EO, KI, SE, OC, TRRSS, TRRER and TRROT the household-level income in equations [1], [2] and [4] shows negative values, current transfers paid should be scaled down so that the underlying income concept is set to zero. In the case of primary income, the following adjustment should be applied to expression [1]:

If  $TRPOT > \text{sum}(E, KI, SEI, TRROT)$  then

$TRPOT' = \text{sum}(E, KI, SEI, TRROT)$

[1] Equivalised primary income =  $\text{sum}(E, KI, SEI, TRROT) - TRPOT' = 0$

Then market income is calculated according to expression [2].

In the case of negative values of gross income, the following adjustment should be applied to expression [3]:

If  $TRPER > \text{sum}(MI, TRRSS)$  then

$TRPER' = \text{sum}(MI, TRRSS)$

[3] Equivalised gross income =  $\text{sum}(MI, TRRSS) - TRPER' = 0$

Disposable income is then computed as in [4].

In the case where taxes are larger than gross income, making disposable income negative, the following adjustment should be applied to expression [4]:

If  $TA > GI$  then

$TA' = GI$

[4] Equivalised disposable income =  $GI - TA' = 0$

Finally, the following will be applied to TRP:

$TRP' = TA' + TRPER' + TRPOT'$

The ranking of individuals should be done on the basis of these new values of disposable income. Finally, mean values of market income and disposable income are computed over all incomes, i.e. both zero and positive incomes.

---

<sup>6</sup> As in the case of taxes (TA), current transfers paid to employment-related social insurance schemes (TRPER) and to non-profit institutions and other households (TRPOT) should be recorded as negative values.

### ***Top and bottom coding***

OECD indicators should be computed based on micro-data that are not top or bottom coded. However, any obvious data-entry error should be eliminated.

### ***Income poverty***

Poverty is defined using both relative thresholds and a more “absolute” threshold (computed from a relative threshold anchored in time):

- *Relative poverty*: the relative poverty threshold is expressed as a given percentage of the median disposable income, expressed in nominal terms (current prices). Therefore, this threshold changes over time, as the median income changes over time. Two relative poverty thresholds are used: the first one is set at 50% of the median equivalised disposable income of the entire population, the second one is set at 60% of that income.
- *“Absolute” poverty*: the “absolute” poverty threshold is set at 50% of the median income observed in a given reference year in the past. Only one reference year is used for this “absolute” threshold: 2005 (or the closest available year). This threshold should be inflation-adjusted each year so as to remain constant, in real terms, over time. The value of the poverty line, the reference year (in case it is not 2005) and the consumer price index used to adjust it for inflation should be reported in the sheet “Metadata” (see below)

Two types of indicators are used to characterise poverty:

- The *headcount ratio*, calculated as the number of individuals in the group considered with disposable household income per equivalent member lower or equal to the poverty threshold, as a percentage of the total number of individuals in the group considered.
- The mean *poverty gap ratio* (income gap expressed as % of the poverty threshold). This is calculated as the difference between the poverty threshold and the mean disposable income of the poor, expressed as a percentage of the poverty threshold.

Note: the poverty threshold is calculated based on the entire population. In other words, poverty rates for the working-age and the retirement-age population are computed based on the median income for the entire population.

## **2. Inequality and poverty indicators (Table 1)**

Table 1 provides a set of aggregate indicators on disposable income, income inequalities and poverty for three different population groups: the entire population, the population of working age (individuals aged 18-65) and the population of retirement age (individuals aged 66 and over). Children (persons aged below 18) should be included only in the entire population.

Individuals are ranked according with their *household disposable income per equivalent household member* as described in equation [1], except for the indicators:

- “Gini market income” (i.e. before taxes and public transfers), where individuals are ranked according with their market income per equivalent household member, including cases with zero market incomes;

- “Gini before taxes”, where individuals are ranked according to their pre-tax income, including cases with zero income; and
- “Gini primary income” (i.e. income before taxes, public transfers and flows associated to employment-related social insurance schemes), where individuals are ranked according to their primary income, including cases with zero income.

#### Indicators formula

Indicator	Formula	Comments
<b>Gini index</b>	$Gini = \left( \frac{2}{\mu \cdot n^2} \cdot \sum_{k=1}^n k \cdot W_k \right) - \frac{n+1}{n} = \frac{2 \operatorname{cov}\left(W_k, \frac{k}{n}\right)}{\mu}$ $= \frac{\frac{2}{n} \sum_{k=1}^n (W_k - \mu) \cdot \left( \frac{k}{n} - \frac{1}{n^2} \sum_{k=1}^n k \right)}{\mu}$	<p>Household incomes per equivalent household members (<math>W_k</math>) are ranked in ascending order (such as <math>k = 1, 2, \dots, n</math>).</p> <p>Individuals falling in each of the three population groups (entire population, population of working age and population of retirement age) should be ranked separately.</p> <p><math>n</math> is the total number of individuals;  <math>\mu</math> is the arithmetic mean of disposable incomes: <math>\mu = \frac{\sum_k W_k}{n}</math>.</p>
<b>Mean poverty gap</b>	$\frac{(z - \mu_p)}{z} = \frac{\left( \frac{1}{P} \sum_{i=1}^p \sum_j (z - W_{ij}) \right)}{z}$	<p><math>z</math> is the poverty threshold;  <math>p</math> is the number of poor;  <math>\mu_p</math> is the mean income of the poor.</p>

#### Poverty indicators “before taxes and transfers”

While poverty indicators “after taxes and transfers” are based on the equivalised disposable income of each person, poverty indicators “before taxes and social security transfers” are based on the equivalised **market** income of the individual. However, both types of poverty indicators are based on a poverty threshold set in terms of equivalised **disposable** income. In other terms, people are counted as poor “before taxes and social security transfers” when their **market** income is lower or equal to 50% (or 60%) of the **median disposable** income (i.e. the poverty thresholds are the *same* as those used for poverty indicators “after taxes and social security transfers”).

### 3. Disposable income per deciles (Table 2)

Table 2 describes the structure and composition of household disposable incomes across deciles. The income sources considered are those specified in identity [1] above. This table indicates the distribution across deciles of the different income sources, for three population groups: the entire population; the population of working-age (individuals aged 18-65) and the population of retirement-age (individuals aged 66 and above). Children (persons aged below 18) should be included among the entire population.

Individual observations are ranked by *ascending values of household disposable income per equivalent household member* ( $DI_{ij}$ ). For each of the two panels, income estimates are ranked separately; i.e. upper bound values should be specific to the two population groups, and each decile should contain 10% of the respective reference population.



The upper bound value is the income value at the upper breaking point of the corresponding decile. Therefore, the upper bound value of decile 1 corresponds to the income of the 10% up from the bottom individual; that of decile 9, to the income of the 90% up from the bottom individual and that of decile 10, to the highest (possibly top coded) income value.

For each income decile, the sum of all income components should be equal to the mean (equivalised) disposable income value reported for that decile in the second column of Table 2. Therefore, taxes should be entered with a negative sign.

#### 4. Disposable income per household groups (Table 3)

Table 3 provides information on which types of households are at risk of low incomes, and how some particular sub-groups contribute to shape the overall pattern of inequality and income poverty. It shows, for various population sub-groups, the following variables:

- the percentage share of people in the total population<sup>7</sup>;
- the mean disposable income (in nominal prices);
- the poverty rate, before and after accounting for net transfers (taxes and public transfers), expressed in terms of the headcount ratio. The poverty threshold is equal to the first relative threshold used to calculate poverty indicators reported in Table 1, i.e. 50% of the current median equivalised disposable income of the *entire* population.

#### *Definition of household types, by household structure and work attachment*

Individuals should be classified by household type according to the characteristics of the household reference person (or household head). In line with the *2011 Canberra Handbook*, it is recommended that the household reference person be identified by going through (sequentially) the criteria listed below, until a person is identified:

- one of the partners in a registered or de facto marriage, with dependent children;
- one of the partners in a registered or de facto marriage, without dependent children;
- a lone parent with dependent children;
- the person with the highest income; and
- the oldest person.

These criteria imply that, in the case of households composed by two or more adults, the household reference person (or head) is the one with the highest income or (in the unlikely case where two adults have identical income) the oldest person.

The basic criteria to be used to classify people by household type is the age of the household reference person (non-retirement-age head, i.e. under 66 years old ; and retirement age head, i.e. aged 66 and over), leading to two major groups. This version of the OECD questionnaire includes breakdowns for both households with a non-retirement-age head and for household with a retirement age head.

---

<sup>7</sup> This implies that the sum of all shares by household type (no matter the age of the household head) should equal 100%.

The first group corresponds to individuals belonging to a household with a head of non-retirement-age (under 66). Therefore, all individuals belonging to a household with a head above 66 years old are excluded from the sample for the purposes of filling the upper section of Table 3. For the purposes of completing Table 3, household heads below 18 years old should be considered as “non-retirement age head”; this will ensure that the populations shares reported in Table 3 add up to 100%. Then, within this reference population, individuals are cross-classified according to each of the following criteria:

- the number of adults in the household they belong to: single adult vs. two adults or more. An adult is any individual aged 18 and above;
- the number of children in the household they belong to: with children vs. without children. A child is defined as any individual aged 17 or less;
- the number of household members in employment: no worker, one worker, two workers. A worker is an adult with non-zero annual earnings or self-employment income.

This classification for households with a head of non-retirement age results in ten household types:

- 1) single adult, no children, working;
- 2) single adult, no children, non working;
- 3) single adult, with children, working;
- 4) single adult, with children, non working;
- 5) two or more adults, no children, two or more working;
- 6) two or more adults, no children, one working;
- 7) two or more adults, no children, non working;
- 8) two or more adults, children, two or more working;
- 9) two or more adults, children, one worker;
- 10) two or more adults, children, no workers.

In the (rare) case of households headed by a person aged less than 17, it is recommended that, for the purposes of completing Table 3, these household reference persons should be also considered as “adult”.

The second group corresponds to individuals belonging to a household with a head of retirement age (i.e. 66 and over). Therefore, all individuals belonging to a household with a head below 18 years old or between 18 and 65 years old are excluded from the sample for the purposes of filling the lower section of Table 3. Within this reference population, individuals are cross-classified according to a more simplified set of criteria than the one used for households with a working age head and the one used in past TORs.

- First, as very few members of these households are likely to be aged less than 18, no distinction is made according to the presence or absence of children (i.e. household types are based on the number of person in the households, rather than distinguishing between adults and children).
- Second, the classification only distinguishes between household with at least one working member and those where no member has a paid job.

The classification of household types for households with a head of retirement age hence results in 4 categories:

- 11) single person, working;
- 12) single person, not working;

- 13) two or more persons, at least one working; and
- 14) two or more persons, none working.

### ***Definition of age and gender groups***

The reference population is the entire population, and individuals are grouped according to their age into seven age ranges: 1) 0 to 17 years old; 2) 18 to 25 years old; 3) 26 to 40 years old; 4) 41 to 50 years old; 5) 51 to 65 years old; 6) 66 to 75 years old; 7) over 75.

Table 3 also asks for information on the share of the population, mean income and the poverty headcount by gender of each individual, across the seven age groups detailed above. Population shares should be computed with reference to the total population (i.e. the sum of population across all men and women should add to 100%).

## **5. Metadata**

The questionnaire contains two “metadata” sheets:

- The “Metadata” sheet should be used to report the definitions used and the assumption made to calculate the various indicators, notably in cases where the national data used do not allow following strictly the recommendations made in this Terms of Reference. In such case, national contact-points should provide the alternative definitions or assumptions that they have adopted. Relative to the questionnaire used for wave 6, the present metadata sheet includes additional rows referring to: i) the value of 2005 anchored poverty line (in annual national currency and current prices); ii) the CPI used for deflating incomes; iii) the per capita mean disposable income (non-equivalised, in nominal current prices); and iv) information on standard errors for Gini coefficients (methods used, features of sampling design considered).
- The “Canberra table” is intended to assess the availability of different income components in national sources, and the scope for better adhering to the Canberra 2011 conceptual definition. This new version of the Canberra table also asks for information on the income component “TRPER”. The questionnaire is formulated as mainly closed-loop questions in order to make it easier to respond and collect homogeneous information across member countries. Blank cells are available to add important information that the questionnaire may miss, and to report cases where the definitions used deviate from the OECD template.