

A review of the OECD Income Distribution Database

Leonardo Gasparini & Leopoldo Tornarolli

The Journal of Economic Inequality

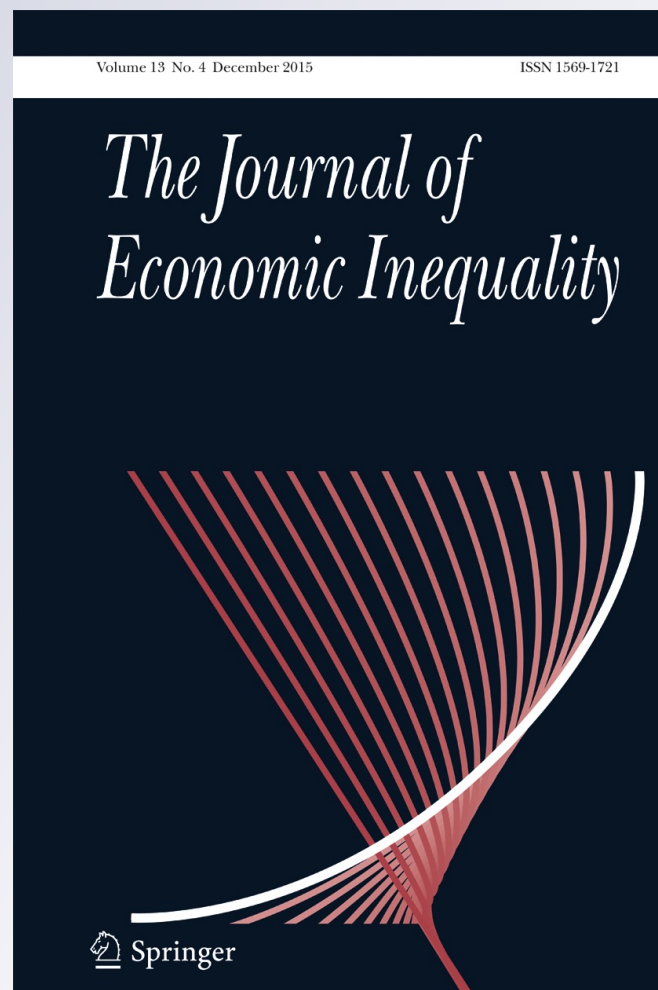
ISSN 1569-1721

Volume 13

Number 4

J Econ Inequal (2015) 13:579-602

DOI 10.1007/s10888-015-9299-x



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media New York. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

A review of the OECD Income Distribution Database

Leonardo Gasparini · Leopoldo Tornarolli

Received: 6 January 2015 / Accepted: 16 February 2015 / Published online: 18 March 2015
© Springer Science+Business Media New York 2015

Keyword Income distribution inequality OECD

1 Introduction

One of the major advances in the field of income distribution in the last two decades has been the increasing availability of large international summary inequality databases. By reporting indicators collected from many countries over time, usually after a process of harmonization, these datasets allow users to monitor and analyze inequality with a scope and accuracy that was unreachable just two decades ago. These sources of information are being used by researchers with increasing frequency, particularly in analyses that involve the comparison of levels and trends of inequality across several countries. Understandably, due to their recent inception and the formidable challenges of the undertakings, they still have some drawbacks and limitations.

One of these valuable initiatives is the OECD Income Distribution Database (henceforth, IDD), a dataset of inequality and poverty indicators for the countries that are part of the organisation for economic co-operation and development (OECD) and for the Russian Federation. The IDD was created in the late 1990s with the goal of improving the comparative assessment of distributive statistics for the member countries of the OECD. From an initial database of a few variables covering 13 economies, the IDD has grown to extend its coverage to all 34 OECD members and the Russian Federation, adding new indicators and

L. Gasparini (✉) · L. Tornarolli
CEDLAS, Universidad Nacional de La Plata, La Plata, Argentina
e-mail: gasparinilc@gmail.com

L. Tornarolli
e-mail: tornarolli@gmail.com

various breakdowns of the information. This database, available online,¹ is extensively used for OECD reports, and it is also a useful input for researchers.²

One key feature of the IDD is that the reported indicators are not computed in-house from the original microdata sources, but instead are collected through an identical questionnaire delivered to consultants in each country, typically from national statistical offices or ministries. This procedure may be seen as lying between that of the Luxembourg Income Study (LIS), which produces standardized microdata, and that of the UNU-WIDER World Income Inequality Database (WIID), where the results are reported from primary sources without any harmonization process.

The procedure for data collection taken by the OECD has the advantage of producing statistics with a standardized methodology, drawing on the experience of country experts who know the specificities of the national surveys.³ On the other hand, this process of collecting information has limitations; as statistics are produced with delays, the scope of indicators is relatively small, and the flexibility to generate new analysis is limited. Some of these concerns are currently being tackled by the OECD.

In this review we expose the OECD Income Distribution Database to critical scrutiny, identifying its strengths and weaknesses. As it is almost inevitable in any critical assessment, we could not avoid being somewhat biased toward highlighting the limitations, without being similarly emphatic about the virtues of the database. To partly compensate for this asymmetry, we should make clear from the outset that this database is a remarkable undertaking that greatly contributes to the study of income inequality in the OECD countries and that deserves full praise for allowing researchers free and easy access to the data. The following comments should then be read bearing always in mind this positive assessment.

The rest of this review is organized as follows. In Section 2 we describe the main features of the database, including the geographical coverage and the time frequency of the data reported. Section 3 discusses the procedures for data collection and the underlying data sources. Section 4 tackles various methodological issues, while Section 5 discusses the comparability of the reported indicators, and Section 6 comments on the accessibility of the data and the quality of the documentation. The relevance of the IDD in terms of its use in order to monitor and analyze inequality by both OECD and external users is reviewed in Section 7. In Section 8 we present some comparisons of inequality patterns using IDD and alternative sources. Finally, in Section 9 we conclude with some remarks.

2 The database

The OECD IDD was created in the late 1990s with the aim of allowing a better comparative assessment of income inequality and poverty levels and trends among the country

¹<http://stats.oecd.org; www.oecd.org/els/soc/income-distribution-database.htm>

²Differently from other undertakings in this field, the primary goal of IDD is not that of allowing researchers access to the data, but rather to provide policy-makers and policy-analysts with a trusted and up-to-date basis for their deliberations.

³Although countries do not formally provide official endorsements, they have the opportunity to comment on data and indicators before the release.

members of the OECD.⁴ The main antecedents of the database include early efforts by Sawyer (1976) in the 1976 OECD economic outlook, who reported income distribution statistics for 12 countries in the 1960s and 1970s drawn from national sources, Förster (1994), who discussed methodological alternatives for international comparisons in the OECD context, and Atkinson et al. (1995), who in a paper for the OECD Social Policy Studies series, documented inequality in the 1980s using data from LIS.

The IDD formally started with the first wave of data collection carried out in 1997/8.⁵ This first initiative included distributional data for 13 countries for two data points, the mid-1980s and mid-1990s. At the time of writing there have been seven waves of data collection: 1997/8, 1999/2000, 2004/5, 2006/7, 2009/11, 2012 and 2014.⁶ The coverage of the database has gradually increased from the original 13 countries in the first round to 35 nations, as data collection efforts were extended to other members, and new countries joined the OECD. Since 2014, data are collected and published on an annual basis. Also, the questionnaires have become more ambitious, adding new indicators, various breakdowns of the information, and standard errors for key variables in some countries.

The main features and results drawn from the IDD have been thoroughly discussed in OECD publications (Förster and Pellizari, 2000; Förster and Pearson, 2002, Förster and Mira d'Ercole, 2005), and used for several reports, including a number of OECD flagship publications such as *Growing Unequal?* (2008) and *Divided We Stand – Why Inequality Keeps Rising* (2011).

The IDD contains information on 70 indicators classified in four categories: income levels, inequality, poverty, and population. Each measure is presented for three different population groups: (i) the entire population, (ii) the population of working age (18 to 65) adults, and (iii) those of retirement age (66 and over). Regarding income inequality, the database reports eight measures: the Gini coefficient for the distribution of equivalized household disposable income (post taxes and transfers), the standard error of that Gini coefficient, the Gini coefficient before taxes and transfers, the P90/P10, P90/P50 and P50/P10 income decile ratios, the S80/S20 income quintile share ratio, and the S90/S10 income decile share ratio.⁷ The information can be explored through OECD.Stat, the statistical online platform where all the OECD's statistical databases can be accessed, or directly from the website of the project.

The IDD reports distributive information on the 34 country members of the OECD, and the Russian Federation (Table 1). While most of them are European nations, the OECD also includes two economies in the Middle East (Israel and Turkey), two in North America (Canada and United States), two in Latin America (Chile and Mexico), two in Asia (Japan and South Korea), and two in Oceania (Australia and New Zealand). The IDD is obviously relevant for those users interested in OECD issues, but

⁴The OECD also implements the earnings distribution database (EDD), which includes information on earnings of full-time dependent employees. The main indicators drawn from EDD are reported in the OECD Employment Outlook. In this review we focus on the IDD, although many of our comments apply to the EDD, as well.

⁵The first wave was a joint effort of the OECD social policy division and the OECD economics department. The following waves were undertaken under the sole responsibility of the OECD Social Policy Division. Since 2012 the management of the database is shared with the OECD statistics directorate.

⁶The timing of this review was somewhat unfortunate. The review took place between October 2013 and February 2014, missing the June 2014 update of IDD indicators to income year 2011/2012.

⁷See the IDD web page for precise definitions. For instance, the S80/S20 ratio refers to the share of all incomes received by the top quintile divided by the share of the first.

Table 1 Coverage of the IDD

Country	Source of data year	Latest		Number of observations				Total
		1970s	1980s	1990s	2000s			
Australia	Survey of Income and Housing	2010	0	0	1	4	5	
Austria	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Belgium	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Canada	Survey of Labour and Income Dynamics	2010	4	10	10	11	35	
Chile	CASEN household survey	2011	0	0	0	2	2	
Czech	Republic EU Survey of Income and Living Conditions	2010	0	0	2	8	10	
Denmark	Danish Law Model System	2010	0	5	10	7	22	
Estonia	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Finland	EU Survey of Income and Living Conditions	2010	0	1	1	5	7	
France	Enquete Revenus Fiscaux et Sociaux	2010	0	0	4	11	15	
Germany	German Socio Economic Panel	2010	0	1	9	8	18	
Greece	EU Survey of Income and Living Conditions	2010	1	1	2	7	11	
Hungary	Hungarian Household Panel/Household Monitor	2009	0	0	9	5	14	
Iceland	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Ireland	EU Survey of Income and Living Conditions	2009	0	0	0	6	6	
Israel	CBS household expenditure survey	2010	0	1	2	5	8	
Italy	EU Survey of Income and Living Conditions	2010	0	1	2	5	8	
Japan	Comprehensive Survey of Living Conditions	2009	0	1	1	4	6	
Korea	Household Income and Expenditure Survey	2011	0	0	0	5	5	
Luxembourg	EU Survey of Income and Living Conditions	2010	0	1	1	8	10	
Mexico	Survey of Household Income and Expenditure (ENIGH)	2010	0	1	1	4	6	
Netherlands	Netherlands Income Panel Survey	2010	1	1	2	7	11	

Table 1 (continued)

Country	Source of data year	Latest				Number of observations		
		1970s	1980s	1990s	2000s	Total		
New Zealand	Household Economic Survey	2009	0	1	2	4	7	
Norway	Income Statistics for Household	2010	0	1	1	5	7	
Poland	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Portugal	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Russian Federation	Household Budget Survey	2010	0	0	0	2	2	
Slovak Republic	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Slovenia	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Spain	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Sweden	Income Distribution Survey	2010	1	1	2	5	9	
Switzerland	EU Survey of Income and Living Conditions	2009	0	0	0	1	1	
Turkey	Household Income and Living Condition Survey	2009	0	1	1	3	5	
United Kingdom	Family Resources Survey	2010	1	1	3	11	16	
United States	Current Population Survey	2010	1	7	10	10	28	
Japan	Comprehensive Survey of Living Conditions	2009	0	1	1	4	6	
Korea	Household Income and Expenditure Survey	2011	0	0	0	5	5	
Luxembourg	EU Survey of Income and Living Conditions	2010	0	1	1	8	10	
Mexico	Survey of Household Income and Expenditure (ENIGH)	2010	0	1	1	4	6	
Netherlands	NetherlandsIncome Panel Survey	2010	1	1	2	7	11	
New Zealand	Household Economic Survey	2009	0	1	2	4	7	
Norway	Income Statistics for Household	2010	0	1	1	5	7	
Poland	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Portugal	EU Survey of Income and Living Conditions	2010	0	0	0	7	7	
Russian Federation	Household Budget Survey	2010	0	0	0	2	2	

Table 1 (continued)

Country	Source of data year	Latest 1970s	Number of observations				Total
			1980s	1990s	2000s		
Slovak Republic	EU Survey of Income and Living Conditions	2010	0	0	0	7	
Slovenia	EU Survey of Income and Living Conditions	2010	0	0	0	7	
Spain	EU Survey of Income and Living Conditions	2010	0	0	0	7	
Sweden	Income Distribution Survey	2010	1	1	2	9	
Switzerland	EU Survey of Income and Living Conditions	2009	0	0	0	1	
Turkey	Household Income and Living Condition Survey	2009	0	1	1	5	
United Kingdom	Family Resources Survey	2010	1	1	3	16	
United States	Current Population Survey	2010	1	7	10	28	

Note: Information extracted from the IDD website on 15 Jan 2014. Note: 1970s=1974-1979, 1980s=1983 to 1989, 1990s=1990 to 1999, 2000s=2000 to 2010

it is also a prime source of information for those studying inequality in high-income economies. The inclusion of some middle-income countries (Chile, Mexico, and Turkey) extends the usefulness of the database, as it allows some comparisons with the developing world.⁸

The IDD reports information starting in 1974, although there are only a few observations for the 1970s and early 1980s. Considering a window of five years around 1985, the database provides information for 17 countries. That number falls to just 12 countries around 1990, increases to 21 around 1995, and grows to 33 around 2005 and 35 around 2010. There are substantial differences in the coverage by country; while the database includes more than 20 observations in Canada, Denmark and the United States, and between 11 and 20 in France, Germany, Greece, Hungary, Netherlands, and the United Kingdom, the number of observations for most countries (23) ranges between 5 and 10. Table 1 displays the number of observations in each country by decade. In some nations, the information needed to trace inequality patterns since the 1980s is sufficient, but in several cases data for the 1980s and 1990s is either very scarce or inexistent. With the exception of some few countries, the IDD allows a close monitoring of inequality patterns only from the mid-1990s through the early 2010s.

Considering the period 1983–2011 for which the customized information is displayed in the database, the IDD includes 327 observations out of the 1015 possible country-year combinations. Although in many cases the missing observations are due to the absence of a data resource, it is technically possible to add observations in some countries when a relevant survey is available. For instance, Atkinson and Morelli (henceforth AM; 2012), in their *Chartbook of Economic Inequality*, construct inequality series for most OECD countries with more observations than those available in the IDD (see Section 8). Since the Atkinson and Morelli dataset is drawn mainly from papers, national reports and official statistics, the cross-country comparability in that database is limited. The OECD has the potential to expand its panel of inequality statistics with better perspectives in terms of cross-country comparability. Such expansion would contribute to enhance the richness of the database and its usefulness for researchers, in particular those in need of large panel datasets.

At the time of our review (early 2014) the latest observations in the IDD correspond to income earned in the year 2010, with just two countries with data for 2011 (Chile and Korea). That delay has been usual in the past: data is usually published with a gap of around three years. Partly, this delay is due to sluggishness in the publication of statistics by the national offices, but the process of collecting the data for the IDD database adds at least an additional year. For instance, for the fourth wave of data collection, the median total response time by the national experts was 16 months (OECD, 2012a).

The OECD is taking action to alleviate this drawback by shortening the questionnaire in order to speed up the responses, reinforcing management of the project by involving the OECD Statistics Department, and increasing the frequency of data collection (to annual collection). Additionally, the project started to calculate indicators in-house on the basis of some available microdata sources, in particular the EU-SILC surveys, and sending the results to national experts and statistical offices for verification. It is still early to assess whether these efforts will be successful in significantly reducing the delay in the publication of statistics.

⁸The OECD envisages to include a larger number of emerging economies in the coming round of updates.

3 Data sources and collection

Data in the IDD is collected through an identical questionnaire delivered to national experts in each country. Typically, the national consultants selected for the project are experts in a government agency in charge of carrying out the household survey and/or producing national distributive statistics. The questionnaire collects summary statistics calculated from microdata from the main household survey (or other source) of each country. These calculations should be carried out by the consultant in accordance with a given protocol. The answered questionnaires, that include tabulations along with metadata with the characteristics of the underlying surveys, are then checked by the OECD for omissions, errors and consistency.

The procedure for data collection taken by the OECD produces statistics with a standardized methodology, taking advantage of the expertise of national consultants. Contrary to other databases, in most cases the IDD estimates are computed from the internal files of the national statistical offices, rather than from public-use files, hence, they are not affected by censoring at the top and other features that may bias the analysis.

On the other hand, this process of collecting information has limitations since statistics are produced with delays, the scope of indicators is relatively small, and the flexibility to generate new analysis is limited. In addition, although a common questionnaire and instructions are delivered to all national experts, the decentralization of the production of statistics may generate spurious heterogeneity in the results, as instructions may be interpreted and carried out in different ways, without the possibility for the external user to control for the quality of the data received.

The questionnaires are filled out by experts, typically in national statistical offices, on a voluntary basis, as no binding legal framework applies. Therefore, since completing the OECD questionnaire is not part of their regular work agenda, delays in the response are frequent, and the possibilities to extend the questionnaires for more ambitious analyses are constrained. Aware of this limitation, the OECD is trying to transform the data collection process into a recognized, more official, and more regular data request with its member countries (OECD, 2012a).

Atkinson et al. (2010) propose a useful classification of databases according to the process of data collection and standardization: (1) common survey instruments, (2) *ex ante* harmonized frameworks, (3) *ex post* standardized microdata, (4) *ex post* customized results, and (5) meta-analyses of results. The OECD IDD belongs to group (4) in which efforts are made to produce harmonized results from the existing set of surveys (or other data instruments). According to this classification LIS would belong to group (3), EU-SILC to group (2) and WIID to group (5). However, as Atkinson et al. (2010) point out, the order does not necessarily imply a quality ranking, as tighter requirements of standardization may have a cost in terms of reduced accuracy in the statistical outcomes.

Table 1 provides a list of the main primary source of information used in each country to estimate the statistics included in the latest round of the IDD.⁹ The data sources are chosen in agreement with officials from member countries and national consultants. In countries where there is more than one survey collecting income information, the rule is

⁹The list of the surveys used in each country in previous waves can be accessed at the IDD website. <http://www.oecd.org/els/soc/IDD-Metadata.pdf>

to choose the survey that better preserves consistency over time and comparability across countries.

The European Union has made strong movements toward a unified system of statistics, a process that had a significant impact on the IDD. In particular, since 2004 for most OECD members belonging to the European Union the statistics on income inequality in the IDD are estimated based on EU-SILC surveys. The fact that the microdata from these surveys are processed in-house reduces the delays, as well as the biases arising from the way in which different national experts interpret the terms of reference. Also, the use of these surveys enhances comparability among countries and over time. Although using the EU-SILC surveys seems a sensible decision, it comes with some inevitable drawbacks. First, it implies a major break in the series, which introduces noise in the comparisons over time.¹⁰ Second, in seven EU countries with national surveys with a long tradition the IDD is still based on a national survey different from EU-SILC, which introduces an asymmetry that hinders the cross-country comparisons.¹¹

4 Welfare variable

Distributive statistics in the IDD are restricted to the income dimension. In particular, the proxy for individual welfare used in the database is *equivalized household disposable income*, which is constructed by dividing *household disposable income* by the square root of household size. In turn, *household disposable income* is obtained through the addition of *cash disposable income* for each household member. This variable is constructed in several steps. The first one is to obtain *factor income* by adding gross wages and salaries, income from self-employment and realized property income. Then, occupational pensions and factor income are summed to get *market income*. Adding cash transfers, both from private and public sources, gives *gross income*. Finally, subtracting personal income taxes and employees' social security contributions from gross income produces *cash disposable income*.¹² In this section we review some of the methodological decisions taken by the OECD in constructing these income variables.¹³

The specific income definitions used in the IDD are based on the recommendations of the Canberra Group Handbook on Household Income Statistics (UN, 2011). The Canberra Group (CG) proposes a *conceptual* ("ideal") income definition, but recommends a narrower *operational* definition, due to practical measurement issues. This definition departs from the conceptual one by excluding some income components: value of unpaid domestic services, value of services from household consumer durables and social transfers in kind. Additionally, for purposes of comparison across countries, the CG prescribes an even more restricted (*practical*) definition, which differs from the operational one in several aspects: it only includes wages and salaries in cash and it excludes employer's social insurance contributions, current transfers received from non-profit institutions, current transfers received

¹⁰The break is typically not in the IDD but in the countries themselves as, following the introduction of EU-SILC, they discontinued the previous surveys.

¹¹The OECD considers that for these seven countries the national sources provide a superior base for analysis.

¹²Some few countries (Mexico, Turkey and Hungary) do not have data on income taxes in the household surveys, and hence not all income concepts can be computed. However, as incomes are reported on a net-income basis, estimates of cash disposable income are available for all countries.

¹³See <http://www.oecd.org/els/soc/IDD-ToR.pdf>

from other households in kind, and compulsory fees and fines (paid). The income definition in the IDD is based on this *practical* income definition, although it does not exactly coincide with it. The main difference comes from the fact that the OECD does not include the net value of owner-occupied housing services (imputed rent).¹⁴ This exclusion may introduce biases in comparative distributive analysis, both within and between countries, since (i) some groups enjoy higher rates of outright ownership and live in larger and better dwellings (e.g., the elderly in European countries), and (ii) there are large differences in the proportion of housing owners across economies (e.g., around 50 % in Germany and Austria, and more than 80 % in Eastern Europe countries (Törmälehto and Sauli, 2010)) and in the share of owners with outright ownership or with mortgage debt.¹⁵

Imputed rents are currently included in the official income definition of several OECD countries, while others are moving in the direction of collecting more comprehensive information on this item. The OECD should consider the possibility of broadening the income definition to include imputed rents in the near future. Naturally, this is a difficult issue that deserves serious analysis; rent imputations should be made in a consistent manner to avoid compromising cross-country comparability.

Even though it does not imply a departure from the recommendations of the CG, the exclusion of some non-cash income components, such as the value of home-produced consumption goods, can also influence comparative distributive judgments. While the exclusion of home production may not significantly affect inequality measures in high-income countries, in some of the middle-income economies included in the last waves of the IDD (Chile, Mexico, and Turkey) home production is an important income component, particularly in poor rural households. We understand that, aware of this problem, the OECD has changed its income definition to include non-cash components in the published updates from 2015.

The OECD requires national experts to compute statistics based on the distribution of *after-tax* incomes. For most countries included in the IDD, income components are reported in surveys before deduction of direct taxes and social security contributions paid by households. Hence, an additional step is required to identify these deductions. This step may involve some comparability problems, since some countries use tax records (e.g., Denmark, France, the Netherlands, Norway, and Sweden), others rely on self-reported data (e.g., Japan), and others use micro-simulation models, with different methodologies and assumptions, to impute taxes (e.g., Italy, New Zealand, and the United States). Additionally, there are some countries where incomes are reported net of taxes and contributions: Austria (data prior to the mid-2000s), Belgium (data prior to the mid-2000s), Greece, Hungary, Mexico, Portugal (data prior to the mid-2000s), Spain (data prior to the mid-2000s), and Turkey.

In most household surveys in the OECD countries the reference period is the year preceding the interview or the previous calendar year. However, there are countries for which the household survey collects incomes over a shorter reference period (Australia, Chile, Israel, Mexico, and the United Kingdom), and then incomes in the IDD are converted to

¹⁴The current income definition adopted in the IDD is more similar to the “old” Canberra Group practical income definition, which was proposed in the Canberra Group Handbook 2001. In that definition, imputed rents were excluded from household disposable income.

¹⁵At the same time, including imputed rent will introduce another bias when countries measure it in different and non-comparable ways as is currently the case across OECD countries.

an annual basis. Given that income is expected to have wider fluctuations over shorter periods, inequality in those countries may be overestimated compared to the rest of the OECD countries, raising comparability concerns.

The inequality measures presented in the IDD are estimated over the distribution of *equivalized* household disposable income, constructed by dividing household income by the square root of household size. This adjustment is made to reflect the fact that the needs do not grow proportionally with household size, due to economies of scale in consumption. Since there is no general agreement on the issue of the equivalence scales in the literature, it would be advisable to report some key results using alternative scales. In particular, (i) the use of the so-called *modified OECD equivalence scale* would allow better comparisons with statistics from Eurostat, while (ii) the report of statistics on household *per capita* income will increase the comparability of the results with other databases including emerging economies and developing countries where reporting on a per-capita basis is the norm.¹⁶

5 Comparability

International databases of inequality statistics are useful when the reported indicators are comparable across countries and over time. In fact, the creation of the IDD in the late 1990s was motivated by several studies pointing out the difficulties of assessing inequality in the OECD, as countries used different methodologies, including national-specific income definitions (Sawyer, 1976; Förster, 1994, Atkinson et al., 1995).

The IDD project makes concrete efforts to compute standardized income distribution statistics by asking the data providers to comply with a set of methodological choices including the income definition, the unit of analysis, the adjustment for needs, and the reporting period. In that sense, the IDD implies a major step toward the provision of comparable income inequality and poverty statistics in the OECD.

While the use of a common protocol to compute inequality is the basis for cross-country comparability, the IDD project is also concerned in promoting comparability *within* countries over time. National consultants are chosen due to their expertise with the country data, which they use focusing on the comparability of the national statistics over time. Also, as mentioned above, the IDD uses a national household survey different from the standardized EU-SILC framework in several European economies. OECD (2012b) justifies that decision on the grounds of (i) having longer time series at the national level and (ii) drawing inequality statistics from surveys that are more representative and more frequently used in the national social debates, two reasons that give priority to the within-country over the cross-country comparisons.¹⁷

Major challenges to comparability are discontinuities caused by changes in the choice of survey used as source of information, or by changes in survey design, weighting or other methodological matters. The best way to address these issues is by reporting statistics for the same year with the old and new survey or methodology. This procedure allows users to

¹⁶Sensitivity analyses comparing key indicators based on the square root scale with those obtained with the modified OECD scale show only very minor differences. Differences with results based on per-capita income are more sizeable.

¹⁷OECD argues that in some countries the source used in the IDD (e.g. GSOEP in Germany) provides estimates that are more comparable to those for other countries than those that could be obtained based on EU-SILC.

assess the impact of the change on the inequality index and to construct better time series by chain-linking the indicators. The OECD provides the original data from old sources with an overlap year with the current data source, which facilitates the assessment of the impact of methodological changes. The OECD also provides (when information for overlapping year is available) break-adjusted series.

Cross-country comparability in the IDD is enhanced by the fact that information is drawn from the standardized EU-SILC framework in almost half of the countries. EU-SILC seeks cross-national comparability through the *ex ante* adoption of common definitions and concepts, although it does not require members to adopt a common survey. In fact, the EU-SILC income surveys have significant differences across countries (Atkinson and Marlier, 2010, Iacovou et al., 2012). For instance, in terms of sampling design, while some nations use administrative records supplemented with interviews (e.g., Finland, the Netherlands, Norway, Slovenia, and Sweden), most of the countries use rotational panel household surveys, with variations in the number of rotation groups and length of time in the panel. Wolff et al. (2010) point out differences in fieldwork periods, in the method of data collection, in interview duration, and in non-response rates. The questionnaires from which the income variables are derived also have some differences across the EU-SILC surveys, as countries adjust them to idiosyncratic factors.

The cross-national comparability in IDD is lower when including those European countries for which the EU-SILC survey is not used,¹⁸ and the rest of the non-European OECD economies for which a common framework does not exist. For these countries the IDD involves *ex post* harmonization, but without an *ex ante* standard framework, and hence comparability is limited by the constraints imposed by the survey designs.

The use of standardized definitions and concepts in the IDD is aimed at producing comparable income inequality indicators. However, it is practically impossible to obtain full comparability, since the processing of a household survey to construct an income variable requires solving a long list of small issues that countries tackle in various ways. These issues include the treatment of missing information, extreme values, inconsistent answers, zero income, underreporting, and others. For instance, some countries impute missing values and/or recode extremely small and high income values, while others do not perform any kind of adjustment. Even among countries that impute and recode, there are significant differences in the way that those adjustments are performed.

Although the comparability across countries is compromised by the issues discussed in this and previous sections, the degree of comparability in the IDD is still relatively high, at least relative to world inequality databases that include countries in the developing world. Two facts are mainly responsible for this positive assessment: (i) the method of data collection including a standard questionnaire and the requirement to estimate inequality over a well-specified common welfare variable and (ii) the fact that the underlying data sources are similar across most countries, and in particular across countries in the EU.

We believe that the community of users of the IDD would benefit from an additional effort to increase the cross-country comparability of the inequality indicators, by keeping the current methodology while at the same time providing a second set of statistics constructed with the aim of maximizing cross-country comparability. This set could be constructed in-house using a common methodology across countries and data sources that are as comparable as possible.

¹⁸This is a debatable issue: see previous footnote.

Distributive indicators in IDD are difficult to compare to those in other databases for the developing world (e.g., PovcalNet and WIID) due to several methodological differences.¹⁹ The presence of some countries in both sources may serve as a (yet fragile) bridge between them. For instance, for year 2010 the IDD reports a Gini coefficient over the distribution of equivalized household income of 0.466 in Mexico, by far one of the two highest values among all OECD countries (with Chile). The Gini coefficient for the distribution of household per capita income in Mexico 2010 calculated in SEDLAC (2014) with a standardized methodology is 0.475.²⁰ Ginis for Latin American countries in 2010 for that variable range from 0.440 to 0.567, suggesting that all Latin American countries have income distributions that are substantially more unequal than in any country of the OECD. Although this comparative result is commonly asserted in the literature, it is in fact grounded on this kind of extrapolation, for which databases such as the IDD, with data for developed and developing countries from a common framework, are highly valuable.

The OECD has made some efforts to include information from non-OECD countries. For instance, in the fifth round of the project data on some emerging economies were included (Russia, South Africa). If the OECD managed to collect reliable distributive statistics from some developing countries in different regions of the world following the same protocol of the IDD project, that could be a valuable source of information for international studies of inequality and poverty.

6 Accessibility and documentation

Data from the IDD is available online at the webpage of the project,²¹ and can also be accessed through OECD.Stat, the statistical online platform of the OECD. Specifically, the IDD can be found browsing the *Data by theme* panel and clicking in the “Social Protection and Well-being” sub-theme. Finding the IDD in the huge OECD website is not a very easy task, in particular since the database is not included under a heading that clearly refers to inequality or poverty. We believe the visibility of the IDD should be enhanced with a more direct access to the data. In addition, once in the project’s website, there is not any introduction to the database: the first information that the user finds at the top of the website is a short report. At least for the point of view of researchers it would be desirable to include an introduction presenting the database and the website.

With the benefit of the experience of the OECD in producing and disseminating statistical information, the database is very well-organized and user friendly.²² The user can customize what information to see and/or download. For instance, (s)he can get the entire database by downloading individual datasets (one for each country) containing information on all inequality measures, for the whole period. Alternatively, the same result can be reached by downloading 70 individual datasets (one for each inequality measure), including information for all country-year combinations. Those datasets can be downloaded in different formats: Excel, CSV, PC-axis, or XML.

¹⁹For instance, the fact that in many developing countries inequality is computed over the distribution of consumption, not income.

²⁰Mexico is not among the most unequal economies in Latin America. For instance, SEDLAC data from 17 Latin American countries places Mexico as the fifth least unequal country.

²¹www.oecd.org/els/soc/income-distribution-database.htm.

²²The website allows presenting the information by age groups, customizing the layout and the selection, and drawing charts.

Although the data reported in the IDD is very valuable, it is still a small fraction of the entire data collection (25 % to 30 % according to OECD, 2012b). Researchers would greatly benefit from an expansion of the database that can be accessed online.

The IDD is frequently revised and updated, a praiseworthy practice that however raises problems for replication. It is advisable to number the different releases of the dataset and keep available all the versions of the online data, in order to facilitate replications.

Regarding the documentation of the database, the website contains several methodological files including (i) the metadata of the IDD with information on the underlying household surveys; (ii) the Terms of Reference that guide the consultants in each country in the process of data collection and in the calculation of income components and indicators; (iii) a short note on equivalence scales; (iv) documents presented in meetings with data providers covering a variety of topics (e.g., income definition, classification of income components, adjustments with National Accounts information, treatment of negative income, top and bottom coding, correction for item-non-response, computation of standard errors, and household definitions); (v) a quality review of different dimensions of the IDD; and (vi) country reviews.

While most of the information needed to understand the IDD is provided in the documents listed above, the information should be better organized, maybe in a more comprehensive methodological document. In addition, some issues, particularly those related to specific decisions taken by data producers that could potentially affect comparability, need to be explained in more detail: income components collected in each survey, treatment of bottom and top coding, elimination of extreme values, treatment of non-response and under-reporting (including details of imputation procedures if apply) and methodologies used to estimate taxes and social contributions.

7 Uses

Since its inception, the IDD was used in various OECD publications and working papers. The data collected in the first wave was used by Burniaux et al. (1998) and Oxley et al. (1999) to trace the evolution of the income distribution over two decades ending in the mid-1990s. Even when the first wave of the IDD was an important step in the direction of better comparable income distribution data, Burniaux et al. (1998) recognized that "... the lack of consistent cross country definitions for components of income, population coverage and methods of treatment of certain observations makes cross-country comparisons less reliable...".

Förster and Pellizari (2000) and Förster and Pearson (2002) used the second wave of the database to produce a detailed analysis of income distribution changes in the OECD countries. Even when they focused on trends within countries, they also carried out comparisons across economies, initiating the use of the IDD for cross-country analysis. Förster and Mira d'Ercole (2005) compiled the key results of the third wave of the IDD. They found that inequality in the distribution of household disposable income had a slight increase over the second half of the 1990s.

The fourth wave of the IDD included, for the first time, information on the distribution of household disposable income for all OECD members (30 countries at that moment). This wave provided evidence on income distribution and poverty from the mid-1980s to the mid-2000s and was used as a key input in a major OECD report (*Growing Unequal?*) published in 2008. The report, which had a big impact in terms of dissemination and debate, presented a detailed analysis of the trends and driving factors of the income

distribution in OECD countries using information from the IDD, as well as an evaluation of the distribution of other economic resources, such as in-kind public services, consumption patterns and household wealth.²³ The main finding of the report was that income inequality rose in a majority of countries (around 3/4 of them) over the period under analysis.

The fifth wave of the IDD, which was collected during the period 2009–2011, was used as an input to produce a follow-up to *Growing Unequal?*. This report, *Divided We Stand: Why Inequality Keeps Rising* (2011), studies “whether and how trends in globalization, technological change and institutions and policies translated into wage and earnings inequality”, assessing also the role played by other factors such as changes in family structures, tax and benefit systems and public services. The results based on the IDD confirmed those found on the 2008 report: although following different time patterns, income inequality significantly rose in the period mid 1980s–late 2000s in 17 out of 22 OECD members for which there are available information to construct a long data series.²⁴

Besides the aforementioned works on income inequality, the information in the IDD is used in several other OECD publications,²⁵ and features in the bi-annual publications *Society at a Glance–Social Indicators*, and *How's Life? Measuring Well-Being*. Also, selected indicators from IDD are included in other OECD databases, such as OECD Health Data or the OECD Family Database.

The external use of the IDD has been more limited, although it is significantly growing. The database is quoted in journals and books as one of the main sources of information on inequality in high-income countries. Just to mention two relevant examples, the *Oxford Handbook of Economic Inequality* recognizes the IDD as a major contribution that helps overcome the shortcomings of earlier studies (Salverda et al., 2009), and the recent *Handbook of Income Distribution* uses the IDD as a central source of information to trace inequality patterns in high-income economies (Morelli et al. 2015).

8 Comparison with other sources

There are other international databases that provide inequality estimates for OECD countries. In this section we compare coverage and results of the IDD with LIS, the EU-SILC database and the Chartbook of Economic Inequality by Atkinson and Morelli.²⁶

²³It is important to note that even when the information on the dimensions besides income was drawn from the same household surveys used in the IDD, that information was not incorporated in the database.

²⁴More recently, the 2012 OECD flagship publication *Going for Growth* includes a chapter based on IDD data, discussing policy reforms that could yield both increases in GDP per capita and reductions in income inequality.

²⁵These publications include the OECD Economic Department Working Papers, OECD Labour Market and Social Policy Occasional Papers, OECD Economic Studies and OECD Social Employment and Migration Working Papers.

²⁶Other databases include information for OECD countries, such as the UNU-WIDER World Income Inequality Database (WIID), and the dataset assembled for the World Bank's World Development Report 2006. However, the reported inequality estimates in these sources are mostly drawn from (or coincide with) one of the databases reviewed in this section. Inequality estimates for the developing countries in the OECD are also computed and reported in the World Bank's PovcalNet and in some regional initiatives (e.g., SEDLAC and CEPALSTAT).

The Luxembourg Income Study, reviewed in another paper in this volume, is a standardized database that applies common definitions to micro records from different national surveys. LIS covers all OECD countries except Chile, Iceland, New Zealand, Portugal and Turkey, while it also includes statistics for a small set of middle-income non-OECD countries.²⁷

As discussed above, EU-SILC is a project launched in 2003 by Eurostat, which provides annual comparable income survey statistics through an *ex ante* harmonized framework (Atkinson et al., 2010).²⁸ All EU member states are obliged to implement EU-SILC, which requires common procedures, concepts and classifications, and the construction of harmonized variables, but allows countries some degree of flexibility in the underlying sources and in the definitions. The overlap in terms of coverage between IDD and EU-SILC is large, since the latter includes information for all European OECD countries. Only 9 out of the 34 countries in the OECD are not included in the EU-SILC database.²⁹ As discussed above, the IDD project has been using information from EU-SILC to compute inequality and poverty indicators for 16 of the countries included in EU-SILC.

The Chartbook of Economic Inequality by Atkinson and Marlier (2010) (AM) is an effort to track inequality in income, earnings, and wealth in a set of rich and developing countries. The AM database is concentrated on comparability over time within countries, making use of data series from different sources, mainly from national reports and academic studies. AM includes information for half of the countries in the OECD.

Table 2 compares the four databases by summarizing their coverage in the OECD countries and the Russian Federation.³⁰ Compared to LIS, the IDD has similar country coverage, but provides more data points, and it is more updated. On the other hand, LIS allows a longer historical perspective of changes in inequality.³¹ The IDD has some obvious advantages when compared with EU-SILC, since the latter covers only European countries and statistics start in 2003. However, the coverage since that year is more complete and updated in the EU-SILC database. AM provides longer and more complete inequality series than IDD, but the country coverage is smaller, and the comparability of the inequality statistics across countries is lower, since indicators are taken from a number of different sources without any cross-country harmonization.

In summary, the four databases have some pros and some cons in terms of coverage: none of them dominates the others in all dimensions. It should be possible to combine the four sources to construct a larger database of inequality statistics for the OECD countries. For instance, the IDD could be complemented with the results in LIS and AM to extend the historical coverage and with EU-SILC for a more complete and updated assessment of inequality in the latest years. Naturally, for this merge to be possible statistics should be comparable. In the rest of this section we tackle the issue of consistency of inequality results across data sources.

²⁷Indicators for Brazil, China, Colombia, Guatemala, India, Peru, Romania, Russia, South Africa, Taiwan and Uruguay are included, although in most of these cases the number of observations is small.

²⁸EU-SILC is focused on income, but it also covers housing, labor, health, demography, education and deprivation issues.

²⁹Australia, Canada, Chile, Israel, Japan, Korea, Mexico, New Zealand and United States. EU-SILC also covers some non-OECD European countries such as Bulgaria, Croatia, Cyprus, Latvia, Lithuania, Malta and Romania.

³⁰All the databases in the comparison cover other countries as well, beyond those listed in the table.

³¹OECD argues that in some countries there are serious problems of comparability between older and current surveys (e.g. Australia).

Table 2 Initial and latest year and number of observations in alternative databases

Country	OECD IDD			LIS			EU-SILC			Atkinson & Morelli		
	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys
Australia	1995	2010	5	1981	2003	6				1981	2009	14
Austria	2004	2010	7	1987	2004	6	2003	2011	9			
Belgium	2004	2010	7	1985	2000	6	2003	2011	9			
Canada	1976	2010	35	1971	2007	11				1976	2009	34
Chile	2006	2011	2									
Czech Republic	1992	2010	10	1992	2004	3	2005	2011	7			
Denmark	1985	2010	22	1987	2004	5	2003	2011	9			
Estonia	2004	2010	7	2000	2010	4	2004	2011	8			
Finland	1986	2010	7	1987	2010	7	2004	2011	8	1966	2009	27
France	1996	2010	15	1978	2005	6	2004	2011	8	1956	2009	22
Germany	1985	2010	18	1973	2010	11	2005	2011	7	1962	2009	31
Greece	1974	2010	11	1995	2010	5	2003	2011	9			
Hungary	1991	2009	14	1991	2005	4	2005	2011	7			
Iceland	2004	2010	7				2004	2011	8	2003	2011	9
Ireland	2004	2009	6	1987	2010	8	2003	2011	9			
Israel	1985	2010	8	1979	2010	8						
Italy	1984	2010	8	1986	2010	11	2004	2011	8	1911	2008	33
Japan	1985	2009	6	2008	2008	1				1962	2005	30
Korea	2006	2011	5	2006	2006	1						
Luxembourg	1986	2010	10	1985	2010	8	2003	2011	9			
Mexico	1984	2010	6	1984	2004	9						
Netherlands	1977	2010	11	1983	2010	8	2005	2011	7	1977	2008	23
New Zealand	1985	2009	7			1982	2009	2011	13			

Table 2 (continued)

Country	OECD IDD			LIS			EU-SILC			Atkinson & Morelli		
	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys
Norway	1986	2010	7	1979	2004	6	2003	2011	9	1986	2010	24
Poland	2004	2010	7	1986	2010	7	2005	2011	7			
Portugal	2004	2010	7				2004	2011	8	1967	2011	21
Russian Federation	2008	2010	2	2000	2010	4						
Slovak Republic	2004	2010	7	1992	2010	5	2005	2011	7			
Slovenia	2004	2010	7	1997	2010	5	2005	2011	7			
Spain	2004	2010	7	1980	2010	7	2004	2011	8	1964	2011	19
Sweden	1975	2010	9	1967	2005	8	2004	2011	8	1975	2010	33
Switzerland	2009	2009	1	1982	2004	5	2007	2011	5	1982	2004	5
Turkey	1987	2009	5				2006	2011	6			
United Kingdom	1975	2010	16	1969	2010	11	2005	2011	7	1961	2011	51
United States	1974	2010	28	1969	2010	10				1944	2009	64
Mean	1992.3	2009.9	9.6	1985.5	2007.4	6.5	2004.2	2011.0	7.8	1967.9	2008.9	26.6
Japan	1985	2009	6	2008	2008	1				1962	2005	30
Korea	2006	2011	5	2006	2006	1						
Luxembourg	1986	2010	10	1985	2010	8	2003	2011	9			
Mexico	1984	2010	6	1984	2004	9						
Netherlands	1977	2010	11	1983	2010	8	2005	2011	7	1977	2008	23
New Zealand	1985	2009	7				2009	13				
Norway	1986	2010	7	1979	2004	6	2003	2011	9	1986	2010	24
Poland	2004	2010	7	1986	2010	7	2005	2011	7			
Portugal	2004	2010	7				2004	2011	8	1967	2011	21
Russian Federation	2008	2010	2	2000	2010	4						

Table 2 (continued)

Country	OECD IDD			LIS			EU-SILC			Atkinson & Morelli		
	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys	Initial year	Latest year	# surveys
Slovak Republic	2004	2010	7	1992	2010	5	2005	2011	7			
Slovenia	2004	2010	7	1997	2010	5	2005	2011	7			
Spain	2004	2010	7	1980	2010	7	2004	2011	8	1964	2011	19
Sweden	1975	2010	9	1967	2005	8	2004	2011	8	1975	2010	33
Switzerland	2009	2009	1	1982	2004	5	2007	2011	5	1982	2004	5
Turkey	1987	2009	5				2006	2011	6			
United Kingdom	1975	2010	16	1969	2010	11	2005	2011	7	1961	2011	51
United States	1974	2010	28	1969	2010	10				1944	2009	64
Mean	1992.3	2009.9	9.6	1985.5	2007.4	6.5	2004.2	2011.0	7.8	1967.9	2008.9	26.6

Source: own calculations based on IDD, EU-SILC, LIS and Atkinson and Morelli (2012). Note: Information extracted on 15 Jan 2014.

In principle, the results should be quite similar among databases, since the differences in methodology are small. For instance, in IDD, LIS and EU-SILC the concept of disposable income is quasi-identical. In fact, consistent with this expectation, researchers typically find that the IDD data compares well with the LIS and EU-SILC data (OECD, 2008, Atkinson et al. 2010; Morelli et al. 2015). Table 3 shows the linear and rank correlation coefficients between the Gini coefficients in the IDD and those reported in alternative data sources. In all cases estimates refer to the Gini coefficient for the distribution of disposable household income, with some variations in terms of the adjustment for household demographics.³² Since the years covered in each database do not necessarily coincide, we assemble a dataset with observations centered at years 1985, 1990, 1995, 2000, 2005 and 2010.³³ The correlations reported in Table 3 are all positive and statistically significant at 1 % level.³⁴ The global picture of inequality in the OECD countries is highly consistent across the different databases. The least unequal economies are those in Northern Europe (Belgium, Denmark, Finland, Iceland, Norway, and Sweden), as well as some Eastern European nations (Slovenia, Czech Republic, and Slovak Republic). The United States is the most unequal economy among the set of rich nations; in the OECD it is only less unequal than Turkey, Mexico, and Chile. The range of the values of the Gini coefficient is wide: from around 25 in the least unequal economies of Northern Europe to around 50 in the Latin American members of the OECD.

Although the general picture is quite consistent across databases, there remain a few differences in some countries. The most significant ones are accounted for by the use of different household surveys, as in some European countries where the IDD is based on a different survey from EU-SILC. For instance, in Germany the IDD statistics are computed from the German Socio-Economic Panel, which is used by most official national reports on the subject, while in Italy (up to 2006) the microdata of the Survey of Household Income and Wealth is complemented with estimates of household taxes from a micro-simulation model run by the national statistical office.³⁵

Table 4 reports the Gini coefficients for equalized household disposable income in IDD, LIS and EU-SILC for 2010. In general the differences are small, typically lower than one Gini point, but in some cases the gap is wider. More worrying, the differences are not always of the same sign. In some cases the divergences are difficult to explain. Take the case of Denmark and Netherlands; while according to the IDD, the Gini coefficient is 3.6 percentage points higher in the Netherlands, according to EU-SILC it is 2 percentage points lower. Differences with LIS are also generally minor, but worrying in some cases: in Ireland the Gini in LIS is 3.7 percentage points lower than in the IDD, while in UK it is 1.6 percentage points higher.

The correlation coefficients are not so high when considering changes, but still they suggest a broad consistent picture of inequality trends across databases (Table 3). Figure 1 extends the analysis by showing Gini coefficients in a set of 15 countries with enough observations in the IDD and in at least an additional data source. The general patterns that

³²The EU-SILC uses the modified OECD equivalence scale, while the IDD and LIS use a square root equivalence scale.

³³For a given year we take the Gini coefficient of that year, and if an estimate is not available we look for the nearest estimate in a 5-year window.

³⁴The only exception is the correlation for the changes in the 2000s with LIS, where coefficients are only significant at 10 %.

³⁵ See OECD (2012b) for an assessment of the comparison between the IDD methodology and alternative sources for each country.

Table 3 Correlations of Gini coefficients from IDD with other databases

	Levels			Changes	
	Linear correlation	Rank correlation		Linear correlation	Rank correlation
With EU-SILC			With EU-SILC		
2005	0.957	0.931	2005–2010	0.725	0.808
2010	0.936	0.911			
With LIS			With LIS		
1985	0.965	0.923	1985–2000	0.722	0.747
1990	0.934	0.891	2000–2010	0.676	0.548
1995	0.968	0.957	1985–2010	0.745	0.767
2000	0.966	0.918			
2005	0.985	0.973			
2010	0.926	0.921			
With AM			With AM		
1985	0.873	0.903	1985–2000	0.825	0.897
1990	0.823	0.810	2000–2010	0.707	0.790
1995	0.893	0.951	1985–2010	0.886	0.829
2000	0.777	0.811			
2005	0.854	0.890			
2010	0.869	0.864			

Source: own calculations based on IDD, EU-SILC, LIS and Atkinson and Morelli (2012).

emerge from the figures are highly consistent. The distribution of income has become more unequal in most of the rich OECD countries in the last three decades, with few exceptions. The increases in inequality were larger in the 1980s and 1990s, and smaller in the 2000s. While the Gini coefficient increased between 1985 and 2000 in all OECD countries with available information, in the 2000s half of the countries experienced reductions (although some of them may not be statistically significant). The mean Gini in the IDD increased 3 percentage points between 1985 and 2000, while it went up just 0.23 percentage points between 2000 and 2010.

Although the broad patterns are similar when considering other data sources,³⁶ there are some differences for some countries. Take the case of Germany between 2004 and 2010: while the IDD reports almost no change in the Gini coefficient (28.5 in 2004 and 28.6 in 2010), LIS and AM report an increase of around one percentage point, and EU-SILC records a substantial hike of 3 percentage points.³⁷

The inequality patterns reported in IDD for the two Latin American countries in the OECD – Chile and Mexico – are consistent with those estimated in the two databases specialized in Latin American data: SEDLAC and CEPALSTAT. In particular, the IDD records an increase in inequality in Mexico until the late-1990s and a robust fall since then

³⁶ The contrast between decades is slightly more marked when considering AM data and less dramatic when using LIS data.

³⁷ Frick and Krell (2010) find the increase reported in EU-SILC difficult to explain.

Table 4 Gini coefficient, equivalized disposable income, year 2010

Country	IDD	LIS	EU-SILC	Country	IDD	LIS	EU-SILC
Austria	26.7		26.1	Italy	31.9	32.7	31.9
Belgium	26.2		26.3	Luxembourg	27.0	26.9	27.2
Czech Republic	25.6		25.2	Netherlands	28.8	25.7	25.8
Denmark	25.2		27.8	Norway	24.9		22.9
Estonia	31.9	32.5	31.9	Poland	30.5	31.0	31.1
Finland	26.0	26.3	25.8	Portugal	34.4		34.2
France	30.3		30.8	Slovak Republic	26.1	26.3	25.7
Germany	28.6	28.6	29	Slovenia	24.6	25.2	23.8
Greece	33.7	33.3	33.5	Spain	33.8	33.3	34.5
Hungary	27.2		24.1	Sweden	26.9		24.4
Iceland	24.4		23.6	Switzerland	29.8		29.6
Ireland	33.1	29.4	30.7	United Kingdom	34.1	35.7	33.0

Source: own calculations based on IDD, EU-SILC and LIS.

Note: Gini coefficient is multiplied by 100. All values refer to income year 2010, except for Hungary, Ireland and Switzerland that refer to 2009.

(see Fig. 1). Data for Chile, only available in the IDD since the mid-2000s, also reveals a decreasing pattern in inequality.

9 Concluding remarks

The OECD Income Distribution Database is a valuable data resource that greatly contributes to the study of income inequality and poverty in the OECD countries. In this review we expose this database to critical scrutiny from the standpoint of the potential user, identifying their strengths and weaknesses.

The IDD is a major contribution to the cross-country analysis of income distribution in high-income economies. Statistics reported by the OECD are extensively used by researchers and policy analysts to monitor and analyze inequality and poverty in those countries. A characteristic feature of the IDD – the data collection process through identical questionnaires delivered to consultants in each country – allows an *ex post* standardization that increases the comparability of the inequality statistics, in relation to those released by official national sources. At the same time, this process of data collection has some limitations, as statistics are produced with greater delays than those that would be possible through direct access to the underlying microdata, the scope of indicators is relatively small, and the flexibility to check the robustness of the results and generate new analysis for external users is limited.

We believe the database could be extended and improved in some dimensions, as discussed throughout the paper.³⁸ However, the current scheme of data collection, dependent

³⁸The OECD carried out a comprehensive quality self-assessment of the database in 2010 (OECD, 2012a, also available from the website) that included country-reviews, in which the OECD benchmark data series are compared with other data sources. While some of the recommendations have been already implemented (e.g. a more frequent data collection), some others are still pending.

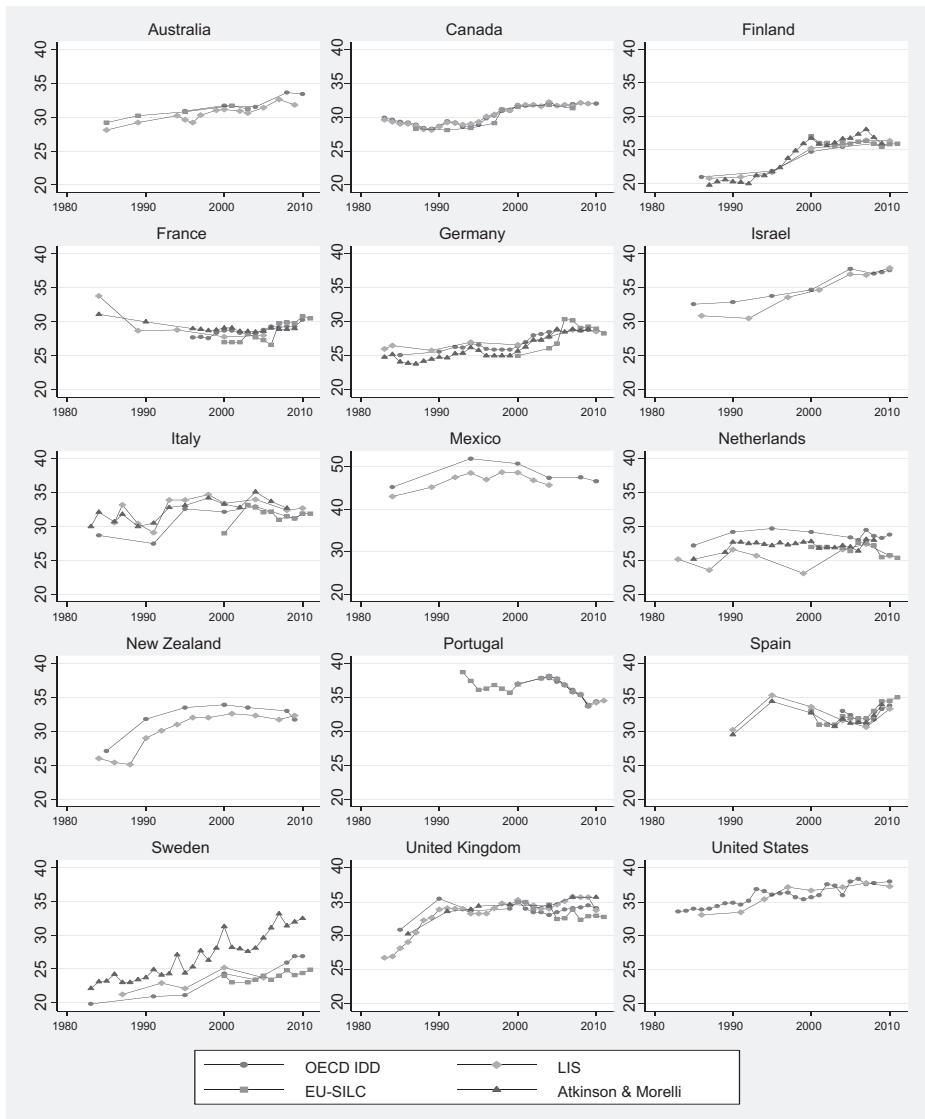


Fig. 1 Gini coefficient. Disposable household income adjusted for demographics. Alternative sources. Source: own calculations based on IDD, EU-SILC, LIS and Atkinson and Morelli (2012). Note: Gini coefficient in scale [0,100]

on the goodwill of governments answering to the requirements of the OECD without a binding legal basis, may not be the ideal environment for a project upgrade. A more ambitious database would likely imply a move toward more in-house work, based on microdata provided by the national governments. In such a framework, it would be easier to improve the dataset by adding other inequality indicators, measures of other distributive dimensions (e.g., polarization, mobility, absolute inequality, and aggregate welfare), estimates of the distributive impact of taxes and transfers (as done in Morelli et al. (2015) using IDD

data), and statistics on inequality in other non-monetary variables (e.g., education). Also, it would be easier to compute confidence intervals for all the indicators, and perform a regular sensitivity analysis to issues such as equivalence scales, weighting and income definitions.

Acknowledgments This paper was prepared for the Journal of Economic Inequality special issue “A Review of International Datasets of Inequality Indicators”. We are grateful to Nora Lustig, Francisco Ferreira, Dan Teles and Michael Förster and colleagues for valuable comments and suggestions, and María Isabel Egas Salgado, Noemí Katzkowicz, Pilar Castrosin and Germán Reyes for excellent research assistance. All errors and omissions are our responsibility. Leonardo Gasparini is at CEDLAS-UNLP and Conicet; Leopoldo Tornarolli is at CEDLAS-UNLP. CEDLAS is the Center for Distributive, Labor and Social Studies at Universidad Nacional de La Plata (Instituto de Investigaciones Económicas, Facultad de Ciencias Económicas).

References

- Atkinson, A., Rainwater, L., Smeeding, T.: Income Distribution in OECD Countries. OECD Social Policy Studies 18, Paris (1995)
- Atkinson, A., Marlier, E.: Income and Living Conditions in Europe. Eurostat Statistical Books. <http://ec.europa.eu/eurostat/documents/3217494/5722557/KS-31-10-555-EN.PDF/e8c0a679-be01-461c-a08b-7eb08a272767?version=1.0> (2010)
- Atkinson, A., Morelli, S.: Chartbook of economic inequality: 25 countries, 1911–2010. <http://www.chartbookofeconomicinequality.com/> (2012)
- Atkinson, A., Marlier, E., Montaigne, F., Reinstadler, A.: Income poverty and income inequality. In Atkinson and Marlier (eds.) (2010), Income and Living Conditions in Europe, Eurostat Statistical Books (2010)
- Burniaux, J., Dang, T., Fore, D., Förster, M., Mira d’Ercole, M., Oxley, H.: Income distribution and poverty in selected OECD countries. OECD Economics Department Working Paper, 189, Paris (1998)
- Förster, M.: Measurement of low incomes and poverty in a perspective of international comparisons. OECD Labour Market and Social Policy Occasional Paper, 14, Paris (1994)
- Förster, M., Mira d’Ercole, M.: Income distribution and poverty in OECD countries in the second half of the 1990s. OECD Social, Employment and Migration Working Paper, 22, Paris (2005)
- Förster, M., Pearson, M.: Income distribution and poverty in the OECD area: trends and driving forces. OECD Economic Studies 34, Paris (2002)
- Förster, M., Pellizzari, M.: Trends and driving factors in income inequality and poverty in the OECD area. OECD Labour Market and Social Policy Occasional Paper 42, Paris (2000)
- Frick, J., Krell, K.: Measuring income in household panel surveys for Germany: a comparison of EU-SILC and SOEP. SOEP Papers on Multidisciplinary Panel Data Research 265, DIW, Berlin (2010)
- Iacovou, M., Kaminska, O., Levy, H.: Using EU-SILC data for cross-national analysis: strengths, problems and recommendations. ISER Working Paper Series 2012-03, Institute for Social and Economic Research (2012)
- Morelli, S., Smeeding, T., Thompson, J.: Post-1970 trends in within-country inequality and poverty: rich and middle-income countries. In: Atkinson and Bourguignon (eds.) Handbook of Income Distribution, vol. 2, chapter 8, pp. 593–696. Elsevier (2015)
- OECD: Growing Unequal? Income Distribution and Poverty in OECD Countries. OECD, Paris (2008)
- OECD: Divided We Stand – Why Inequality Keeps Rising. OECD, Paris (2011)
- OECD: Quality review of the OECD database on household incomes and poverty and the OECD earnings database. OECD, Social Policy Division (2012a)
- OECD: Economic Policy Reforms 2012: Going for Growth. OECD Publishing (2012b)
- Oxley, H., Burniaux, J., Dang, T., Mira d’Ercole, M.: Income distribution and poverty in 13 OECD countries. OECD Economic Studies 29, Paris (1999)
- Salverda, W., Nolan, B., Smeeding, T.: Introduction. In: Salverda, Nolan and Smeeding (eds.): The Oxford Handbook of Economic Inequality. Oxford University Press, Oxford (2009)
- Sawyer, M.: Income Distribution in OECD Countries. OECD Economic Outlook, OECD, Paris (1976)
- Törmälehto, V.-M., Sauli, H.: The distributional impact of imputed rent in EU-SILC. Eurostat, Methodologies and working papers (2010)
- United Nations Handbook on Household Income Statistics. Canberra Group, 2nd edition, Geneva (2011)
- Wolff, P., Montaigne, F., Rojas González, G.: Investing in statistics: EU-SILC. In Atkinson and Marlier (eds.) (2010), Income and Living Conditions in Europe, Eurostat Statistical Books (2010)