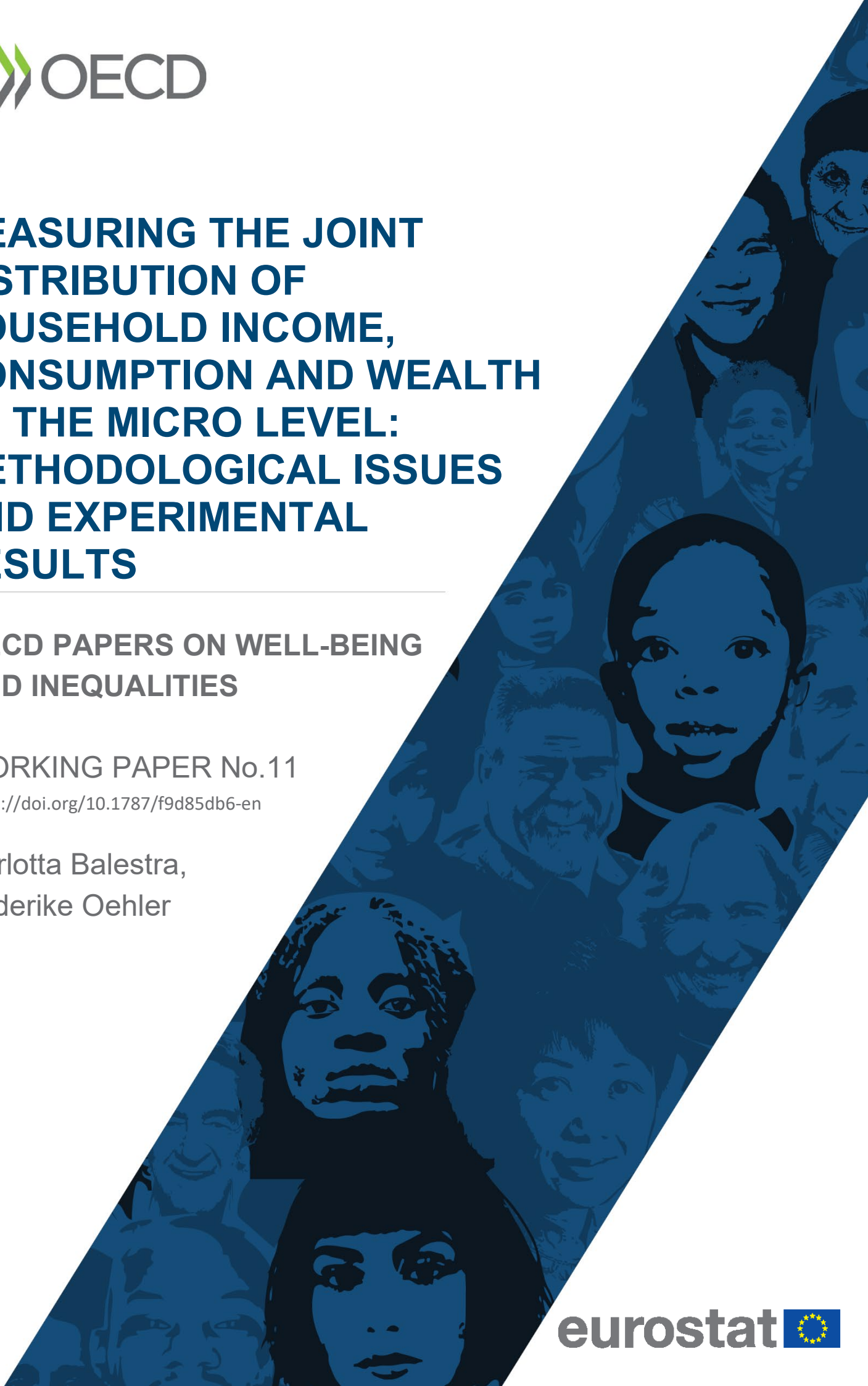
# MEASURING THE JOINT DISTRIBUTION OF HOUSEHOLD INCOME, CONSUMPTION AND WEALTH AT THE MICRO LEVEL: METHODOLOGICAL ISSUES AND EXPERIMENTAL RESULTS



## OECD PAPERS ON WELL-BEING AND INEQUALITIES

WORKING PAPER No.11
https://doi.org/10.1787/f9d85db6-en

Carlotta Balestra,
Friderike Oehler

eurostat

Organisation for Economic Co-operation and Development

WISE(2023)1

English text only

WELL-BEING, INCLUSION, SUSTAINABILITY AND EQUAL OPPORTUNITY CENTRE

# Measuring the joint distribution of household income, consumption and wealth at the micro level

## Methodological issues and experimental results

JEL Classification: C81, D12, D31, E21, I32
Keywords: income inequality, consumption inequality, wealth inequality, experimental statistics

All OECD Working Papers on Well-being and Inequalities are available at https://www.oecd.org/wise/papersandbriefs/.



Carlotta Balestra: Carlotta.BALESTRA@oecd.org;
Friderike Oehler: Friderike.OEHLER@ec.europa.eu.

# OECD Papers on Well-being and Inequalities

**www.oecd.org/wise/papersandbriefs**

The OECD Papers on Well-being and Inequalities Series – managed by the OECD Centre on Well-being, Inclusion, Sustainability and Equal Opportunity (WISE) – features Working Papers on the measurement agenda for well-being, inclusion, sustainability and equal opportunity as well as papers seeking to deepen the understanding of the drivers of these issues, the ways in which they interact and how they evolve. These papers are prepared by OECD staff, external experts or by outside consultants working on OECD projects.

This Working Paper should not be reported as representing the official views of the OECD or of its Member countries or of the European Union or its Member States. The opinions expressed and arguments employed are those of the authors.

Working Papers describe preliminary results or research in progress by the authors and are published to stimulate discussion on a broad range of issues on which the OECD and Eurostat work. Comments on this Working Paper are welcomed, and may be sent to the OECD Centre on Well-being, Inclusion, Sustainability and Equal Opportunity (wellbeing@oecd.org) and to EUROSTAT Unit F1 (ESTAT-ICW@ec.europa.eu).



This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The release of this Working Paper has been authorised by Romina Boarini, Director of the OECD Centre on Well-being, Inclusion, Sustainability and Equal Opportunity (WISE).

Eurostat ISBN: 978-92-76-49146-0 DOI: 10.2785/489635

# Acknowledgements

This work has only been possible thanks to the co-operation of national experts who participated in the activities of the Eurostat-OECD Expert Group on the Joint Distribution of Income, Consumption and Wealth at Micro Level (EG ICW) and, where possible, provided national estimates according to common guidelines.

The following experts have been involved in the work of the Expert Group.[1]

## National authorities

| | |
|---|---|
| Australia | Vanessa Nelson (ABS) |
| Austria | Richard Heuberger, Gregor de Cillia (both Statistik Austria), Peter Lindner and Martin Schürz (both Austrian Central Bank) |
| Belgium | Laurent Van Belle (Banque Nationale de Belgique) |
| Canada | Brian Murphy and Xuelin Zhang (both Statistics Canada) |
| Chile | Javiera Vasquez (Banco Central de Chile) |
| Czech Republic | Jiří Vopravil (Czech Statistical Office) |
| Finland | Tarja Hatakka, Tara Junes (both Statistics Finland) and Veli-Matti Törmälehto (Statistics Finland/European Central Bank) |
| France | Jérôme Accardo (INSEE) |
| Germany | Holger Breiholz, Johannes Köckeis (both Destatis), Julia Le Blanc and Tobias Schmidt (both Deutsche Bundesbank) |
| Greece | Giorgos Ntouros (Statistics Greece) |
| Hungary | Éva Zsuzsanna Ménesi (Hungarian Central Statistical Office) |
| Ireland | Edel Flannery and Brian Cahill (both CSO) |
| Israel | Yafit Alfandari (CBS) |
| Italy | Gabriella Donatiello (ISTAT) and Giuseppe Ilardi (Bank of Italy) |
| Japan | Shunichi Shimakita and Makiko Ochiai (both Statistics Bureau of Japanese Government) |
| Korea | Hae Ryun Kim (OECD/Statistics Korea) |

[1]Listed affiliations reflect the expert's affiliation at the time the work was done.

| | |
|---|---|
| Luxembourg | Jérôme Hury and Guillaume Osier (both STATEC) |
| Mexico | Edgar Vielma Orozco (INEGI) |
| Netherlands | Jeroen Nieuweboer (CBS) |
| Norway | Tor Morten Normann (SSB) |
| Poland | Łysoń Piotr and Żyra Małgorzata (both Statistics Poland) |
| Portugal | Luis Correia (INE) |
| Romania | Roxana Adam (INS) |
| Slovenia | Brigita Vrabic Kek (SURS) |
| Switzerland | Stephan Häni (BFS) |
| United Kingdom | Richard Tonkin, Dominic Webber and Ozer Beha (all ONS) |
| United States | Jonathan David Fisher (Stanford University), Liana Fox (Census Bureau) and Melissa Gentry (Federal Reserve Bank of Boston) |

## International agencies

| | |
|---|---|
| European Central Bank | Juha Honkkila and Veli-Matti Törmälehto |
| Luxembourg Income Study | Teresa Munzi, Jörg Neugschwender and Piotr Paradowski |

Eurostat (Jean Louis Mercy and Friderike Oehler) and the OECD (Marco Mira d'Ercole and Carlotta Balestra) acted as Secretariat to the Expert Group. Pierre Lamarche (INSEE, formerly Eurostat) and Richard Tonkin (United Nations, formerly ONS and OECD) helped steer the work of the Expert Group by contributing to the development of the methodology, operational guidelines and the quality assessment framework that underlie the estimates presented in this paper. In addition, the authors would like to thank Florian Pallaro (Sogeti) for helping produce and validate the data for EU Member States, Irene Riobóo (Eurostat) for her contributions to the methodology and quality assessment of Eurostat's centralised exercise, as well as Kamil Kouhen, Luiz Hermida and Christine Le Thi (all OECD) for statistical assistance. The OECD gratefully acknowledges financial support from Statistics Canada. This paper uses data from the Eurosystem Household Finance and Consumption Survey (HFCS), which is run by the national central banks of the euro area and co-ordinated by the European Central Bank. The results published in this report and the related observations and analysis may not correspond to results or analysis of the data producers. The authors wish to thank the members of the Expert Group: Marco Mira d'Ercole (OECD Centre on Well-Being, Inclusion, Sustainability and Equal Opportunity), Sebastian Königs, Vanda Almeida (both OECD Directorate for Employment, Labour and Social Affairs) and Jorrit Zwijnenburg (OECD Statistics and Data Directorate) for their comments on previous versions of this paper.

# Country abbreviations

| Country name | Two-letter code | Three-letter code |
|---|---|---|
| Australia | AU*** | AUS*** |
| Austria | AT | AUT |
| Belgium | BE | BEL |
| Bulgaria | BG** | BGR** |
| Canada | CA*** | CAN*** |
| Croatia | HR** | HRV** |
| Cyprus | CY†,** | CYP†,** |
| Czech Republic | CZ | CZE |
| Denmark | DK | DNK |
| Estonia | EE | EST |
| Finland | FI | FIN |
| France | FR | FRA |
| Germany | DE | DEU |
| Greece | EL | GRC |
| Hungary | HU | HUN |
| Ireland | IE | IRL |
| Israel | IL††,*** | ISR††,*** |
| Japan | JP*** | JPN*** |
| Korea | KR*** | KOR*** |
| Latvia | LV | LVA |
| Lithuania | LT | LTU |
| Luxembourg | LU | LUX |
| Malta | MT** | MLT** |
| Mexico | MX*** | MEX*** |
| Netherlands | NL | NLD |
| Norway | NO*** | NOR*** |
| Poland | PL | POL |
| Portugal | PT | PRT |
| Romania | RO** | ROU** |
| Slovak Republic | SK | SVK |
| Slovenia | SI | SVN |
| Spain | ES | KOR |
| Sweden | SE | SWE |
| Switzerland | CH*** | CHE*** |
| United Kingdom | UK*** | GBR*** |
| United States | US*** | USA*** |

Notes: ** denotes non-OECD countries; *** denotes non-EU countries.

† Note by the Republic of Türkiye: The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the "Cyprus issue". Note by all the European Union Member States of the OECD and the European Union: The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

†† The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

# Abstract

This paper provides an overview of the work of the Expert Group on the Joint Distribution of Income, Consumption and Wealth at Micro Level (EG ICW) set up by Eurostat and the OECD. It discusses the challenges of producing joint income, consumption and wealth estimates, assesses their quality, and presents selected experimental results. Although the analysis reveals large differences between countries, a number of general patterns emerge. First, income, consumption and wealth are partially correlated, with the association being stronger in the tails of the joint distribution than around its middle. Second, risk of poverty goes beyond income, with asset and consumption risk of poverty being widespread, especially among some population groups. Third, a large share of households spend more than they earn. This is corroborated by negative median saving rates for households in the bottom income quintile. Fourth, inequalities are significantly higher when using a comprehensive measure of material living standards than a distributional analysis of disposable income would suggest. Looking ahead, this paper calls for further efforts to improve the robustness of the results.

# Résumé

Ce document donne une vue d'ensemble du travail du Groupe d'Experts sur la distribution conjointe du revenu, de la consommation et du patrimoine au niveau micro créé par Eurostat et l'OCDE. Il discute des défis liés à la production d'estimations conjointes du revenu, de la consommation et du patrimoine, évalue leur qualité et présente une sélection de résultats expérimentaux. Bien que l'analyse révèle de grandes différences entre les pays, un certain nombre de tendances générales se dégagent. Premièrement, le revenu, la consommation et le patrimoine sont partiellement corrélés, l'association étant plus forte dans les queues de la distribution conjointe qu'en son milieu. Deuxièmement, le risque de pauvreté dépasse la question du revenu, le risque de pauvreté lié aux actifs et à la consommation étant aussi très répandu, en particulier au sein de certains groupes de la population. Troisièmement, une grande partie des ménages dépensent plus qu'ils ne gagnent. Ce constat est corroboré par des taux d'épargne médians négatifs pour les ménages du quintile de revenu inférieur. Quatrièmement, les inégalités sont nettement plus élevées lorsque l'on utilise une mesure globale du niveau de vie matériel que ne le suggère une analyse de la distribution du revenu disponible. Ce document plaide pour des efforts supplémentaires afin d'améliorer la robustesse des résultats.

# Table of contents



## FIGURES

## TABLES

# 1 Introduction

Income, consumption and wealth (ICW) are the key components of households' economic well-being. Analysis of the distribution of each component is typically done individually, with each category being considered as a proxy of household economic well-being. While analysis of these components yields important insights, assessing them together can deepen our understanding of economic well-being. In most developed countries, analyses of material living standards have generally used household income data (or components of income, such as earnings). This reflects the higher frequency with which income statistics are produced and the fact that, for many households, income is the most important economic resource for meeting everyday living expenses. However, this approach is limiting in several respects.

For instance, household income can be small, and wealth can be large (or vice versa). Moreover, both household wealth and income may be small, but the recourse to higher debt or wealth revaluations (for example, those associated with changes in housing and stock prices) may meet, partially or completely, people's consumption needs[2]. Therefore, both income and wealth determine consumption possibilities, while low levels of income and wealth may not always imply a low level of consumption. However, households may choose to consume less than their income and wealth would otherwise permit in order to save more to be better prepared for future adverse events or to build up wealth.

There is (limited) evidence to show that analysing economic inequality through income, consumption or wealth alone understates the level and growth in inequality in the other two dimensions. Multi-dimensional inequality in the United States has increased faster than one-dimensional inequality due to an increasing overlap of households at the very top of the distribution of multiple dimensions (Fisher et al., 2021[1]). These patterns emphasise the importance for a multi-dimensional framework to understand material living standards and their sustainability over time better.

The need for an integrated analysis of household economic well-being has been acknowledged in various forums, including the report of the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz, Sen and Fitoussi, 2009[2]) and the Vienna Memorandum issued by the Conference of the Directors Generals of National Statistical Institutes (DGINS) in 2016[3,4]. In response to the growing demand for relevant statistics, in 2017 Eurostat and the OECD launched the joint Expert Group on the Joint Distribution of Income, Consumption and Wealth at Micro Level (EG ICW) to compile measures of

---

2 In this paper, the term 'consumption' is used as an equivalent to 'consumption expenditure'. However, the two concepts are distinct. Consumption includes inter-household in-kind transfers of gifts and services and social transfers in kind, which are not recorded in an expenditure account. Moreover, only the annual value of services from household consumer durables is included in consumption instead of the initial purchase of the capital items. The distinction is important as it may affect some of the findings reviewed in the paper.

3 The DGINS memorandum can be found at: https://ec.europa.eu/eurostat/documents/13019146/13237859/DGINS+Memorandum+2016.pdf/4ebdf162-1b20-4d9e-a8c7-ae880eca9afd?t=1501752851000.

4 Recommendation 2 of the Stiglitz-Sen-Fitoussi report argues that 'the most pertinent measures of the distribution of material living standards are probably based on *jointly* considering the income, consumption, and wealth position of households or individuals' (Stiglitz, Sen and Fitoussi, 2009, p. 39[2]).

the joint distribution of economic well-being across household groups. The *2013 OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth* (OECD, 2013[3]), the experimental estimates first released by Eurostat in 2017[5], and national experiences in developing measures of the joint distribution of all types of economic resources provided a point of departure for the expert group's work. This work was the first attempt to develop comparable experimental statistics on the joint distribution of ICW beyond the EU.

As part of the expert group, national experts from EU Member States and OECD countries computed distributional results on household ICW using detailed micro national information, following a standard template and methodology. In total, the exercise covered 36 countries; for most of them, information on the three economic dimensions is only available from separate surveys of different household samples with no common identifier. In addition, for several countries there are no data on the distribution of household wealth. In the absence of robust information on the joint distribution of ICW compiled through a single survey, model-based approaches (for example, statistical matching based on different microdata sources) were used to compile a fused ICW (or income and consumption (IC)) dataset. For most EU Member States, estimates were produced by Eurostat using the information available in the EU Statistics on Income and Living Conditions, the Household Budget Survey, and the Household Finance and Consumption Survey (Lamarche, Oehler and Rioboo, 2020[4]).

From a policy perspective, a multi-dimensional approach to household economic well-being can prove important in a number of ways, including those described below.

- **Better understanding the differences in household behaviour** (for instance, in terms of consumption and saving patterns) and focusing on specific population groups. From a risk-of-poverty perspective, this could mean accounting for the existence of multiple deprivations and **identifying people who lack enough resources** to meet essential material needs (Social Metrics Commission, 2018[5]). Understanding these patterns could also have important implications when analysing **intergenerational mobility and the transmission of economic status** since there is evidence that children are overrepresented in the bottom half of all three distributions (Smeeding, 2016[6]).
- **Anticipating and reacting to imbalances in household accounting,** such as those that led many households, including low- and middle-income ones, to accumulate unsustainable levels of debt in the period before the global financial crisis, spurred by stagnant earnings, rising house prices and more access to credit.
- Evaluating the **overall redistributive effects of tax and benefit systems** and assessing jointly the impact of both direct taxes (for example, income and wealth taxes) and indirect taxes (for example, VAT), which require information on both household income and wealth/expenditure (Kuypers, Figari and Verbist, 2021[7]; Eurostat, 2020[8]).
- Providing an insight into the **effects of household economic well-being on aggregate economic performance**. For example, the coincident measurement of household income, consumption and wealth at the micro level can provide information concerning how changes in the level and composition of taxation and government spending can influence aggregate demand in the economy to achieve the goals of full employment and inclusive growth.

This paper has two aims. The first is to discuss methodological challenges that had to be addressed to produce reliable estimates of the joint distribution of income, consumption and wealth at the micro level. The second is to present key results and showcase how these data could be used in policy making.

---

5 Eurostat's experimental estimates on the joint distribution of income, consumption and wealth: https://ec.europa.eu/eurostat/web/experimental-statistics/income-consumption-and-wealth.

The evidence presented in this paper is experimental and should be interpreted with caution. The paper does not intend to provide an accurate picture of income, consumption and wealth inequality levels in EU Member States and OECD countries but rather to explore **joint** distributional aspects. In this respect, despite the experimental nature of the estimates, the analysis in this paper can help improve our understanding of economic inequalities in ways that would not be possible with a univariate approach relying on data from different sources.

Although the analysis reveals big differences between countries, a number of general patterns emerge:

- In general, the distributions of consumption and income are more closely linked than the distributions of consumption and wealth, with the association of the wealth and income distributions lying in between.
- The three distributions overlap more strongly at the tails than around the middle, meaning that households at the top (or bottom) of one distribution tend to belong to the top (or bottom) of the other two distributions as well. In general, the association is higher at the top of the joint distribution than at the bottom.
- The overlap in risk-of-poverty measures is highest between income and wealth: on average, 80% of income-poor individuals are also asset poor. Unemployed people, people who rent and people without at least an upper secondary level of education are systematically at a higher risk of being poor in more than one economic dimension. By contrast, while both young people and older people are worse off than other adults in terms of income, only young people are disproportionately represented in the lower reaches of the wealth and consumption distributions.
- Overall, one in three households appear to spend more than they earn, possibly reflecting measurement errors, limits in the statistical concepts used, and the importance of wealth in financing expenditure. However, saving behaviour varies greatly according to household characteristics. Saving rates are usually lower for households at the bottom of the income distribution and, at a given income level, where the household member with the highest income is younger than 35 years or older than 64 years. This is also the case for single-parent households.
- Inequalities in living standards are generally much higher when looking at discretionary income (in other words, the amount of household disposable income that is left for discretionary spending, investing, or saving after paying for necessities) than at disposable income or consumption expenditure.

While the analysis in this paper is based on data that date back to around 2015, it documents existing fault lines and high levels of household financial vulnerability that are likely to have been exacerbated by the COVID-19 pandemic and the recent cost-of living crisis. These events have demonstrated the importance of using multi-dimensional measures of household economic well-being and assessing households' capacity to maintain their living standards in the face of large shocks. When new data become available, the measurement framework developed as part of the expert group's work will prove appropriate in assessing the overall distributional effect of recent events, including the COVID-19 crisis and inflationary pressures, on household economic well-being and identifying the worst-hit population groups.

This paper is structured as follows:

- Section 2 provides information on the work of the EG ICW, including a review of the guidelines and quality assessment framework used in the exercise.
- Section 3 discusses the main challenges faced when building joint ICW distributions.
- Section 4 presents the experimental results.
- Section 5 concludes by summarising the main lessons learned from the exercise so far and calling for further work to improve the robustness, timeliness and granularity of ICW experimental estimates.

Results presented for Austria, Belgium, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Germany, Greece, Hungary, Ireland, Latvia, Lithuania, Luxembourg, Malta, Poland, Portugal, Romania, the Slovak Republic, Slovenia, Spain and Sweden have been computed in an experimental exercise by Eurostat. Experimental results for the remaining countries are denoted as national or country-based estimates and marked with * in tables and graphs. These estimates have been computed by either national experts (Australia, Finland, France, the Netherlands, Norway, Switzerland, the United Kingdom and the United States) or by the OECD based on microdata shared by national experts (Canada, Japan, Korea, Israel and Mexico). If data are not available, for whatever reason, these are indicated in tables with a colon (:).

Non-OECD countries are marked with '**' in tables and graphs.

# 2 The work of the EG ICW

The EG ICW was launched in 2017, co-chaired by Eurostat and the OECD. It consisted of representatives from national statistical offices, other agencies collecting or producing micro or macro-level data on household economic resources (for example, national central banks), researchers from EU Member States and OECD countries with experience in the field of micro statistics on household economic resources, as well as national accounts experts and representatives from other bodies (for example, the ECB)[6]. The expert group operated mainly via electronic communication, supported by three face-to-face meetings. In parallel with the EG ICW, the OECD-Eurostat Expert Group on Disparities in National Accounts (EG DNA) was continuing previous work on the reconciliation of microeconomic and macroeconomic data on household income, consumption and savings from 2017 to 2020 (see Box 2.1).

### Box 2.1. OECD-Eurostat Expert Group on Disparities in National Accounts

In addition to the work of the EG ICW, the OECD-Eurostat Expert Group on Disparities in National Accounts (EG DNA) is also aiming to derive distributional results, but aligned to macroeconomic aggregates. The frameworks used by the ICW and DNA expert groups are similar, but still differ in some aspects. The EG ICW is looking at the distribution of income, consumption and wealth in the population by joining (bottom up) individual records of different household surveys, whereas the EG DNA applies a (top down) approach, distributing macroeconomic totals for household income, consumption and savings across households based on survey and administrative data. A step-by-step approach is applied to arrive at the distributional results. First, and most importantly, the best possible conceptual correspondence of income and consumption items has to be found, since macro and microdata sources are usually based on different concepts and definitions. Second, an adjustment for differences in the household population is also needed, since microdata are limited to the resident population not living in institutions. Then, the gap between macro and micro aggregates needs to be bridged for each income and consumption item using the most suitable method. To this end, the compilers explore the most likely underlying reasons for the gaps (such as conceptual and classification differences, the impact of underground activities (such as unreported legal or illegal activities), specific household groups that may be missing from the microdata (such as the very rich), and possible measurement and estimation errors) and then allocate them accordingly. As such, for items for which it is assumed that the microdata represent the shape of the true distribution fairly well, a proportional allocation is used. Conversely, for items that are assumed to be particularly underrepresented at the top of the distribution in the microdata, larger shares of the gap are allocated to the upper tail, applying, for instance, a Pareto distribution. In a few cases, other approximations are also used. If possible, allocations are also made for items that are not captured in microdata sources, such as items that are specific to the system of national accounts (for example, social transfers in kind and investment income disbursements). This is usually done on the basis of auxiliary (mostly socio-demographic) information. Finally, households are clustered into

6 Individual members and the organisations to which they are attached are listed in the Acknowledgments section of this paper.

relevant socio-economic groups, for example on the basis of their equivalised income or on the basis of socio-demographic characteristics.

Given the differences in microeconomic and macroeconomic frameworks, the work of the two expert groups builds on different income and consumption concepts. Where the joint distributions of ICW at micro level focus on the actual economic opportunities of households (disposable income, consumption expenditure and wealth), the approach used by the EG DNA integrates items from a broader context, which are only indirectly related to a household budget, such as government services and social security benefits classified as 'social transfers in kind'. The two projects also address different research and policy questions: in the case of the EG ICW, what is the economic reality faced by different household groups?; in the case of the EG DNA, where do different household groups stand on the income ladder within the framework of national accounts?. A total of 14 EU Member States and other OECD countries participated in an EG DNA collection round in 2020. In addition, Eurostat developed distributional national accounts for EU Member States not producing their own estimates[7]. First experimental results, metadata and a methodological note were published in December 2020 by both the OECD and Eurostat[8], while more recent results were published by Eurostat in mid-2022. Further information on the work of the OECD-Eurostat Expert Group on Disparities in National Accounts (EG DNA) is available in Zwijnenburg et al. (2021[9]) and Coli et al. (2022[10]).

In parallel to the EG DNA, the European Central Bank's *Expert Group on Distributional Financial Accounts (EG DFA)* is working on reconciling information on financial and non-financial wealth from microdata sources with macroeconomic information for the household sector in the euro area and EU economies. In the future, more work will be needed to ensure that the data presented under the EG DNA and EG DFA frameworks can be analysed together with joint ICW micro distributions in a meaningful way.

The primary objective of the EG ICW was the development of estimates of the joint distribution of ICW for the latest reference year possible. This new statistical product had to be based on harmonised source data and was to be disseminated with relevant disaggregations. The 2013 *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth* (OECD, 2013[3]), previous Eurostat work (Eurostat, 2013[11]; Leulescu and Agafitei, 2013[12]; Serafino and Tonkin, 2017[13]), the experimental estimates of the joint distribution of income and consumption already released by Eurostat for the reference year 2010 as well as national experiences in developing measures of the joint distribution of different types of economic resources (see, for instance, Törmälehto, Kannas and Säylä (2013[14]), Fisher et al. (2018[15]) and ONS (2020[16])) provided the points of departure for the work of the expert group.

In total, 36 countries were covered by the exercise: 21 countries that are EU Member States and members of OECD, five EU Member States that are not members of OECD, and 10 OECD members that are not members of the EU (see Table 3.1), even though for some of them information was available only for income and consumption. Eurostat produced joint IC or ICW distributions for all EU Member States (except Italy) in a centralised exercise, based on the harmonised data collections available at Eurostat (EU Statistics on Income and Living Conditions and the Household Budget Survey) and the ECB (Household Finance and Consumption Survey)[9]. Due to the limited survey frequency, especially for consumption data,

---

7 The OECD is working on similar results for non-EU OECD countries. These results are expected to become available in the course of 2023.

8 See https://stats.oecd.org/Index.aspx?DataSetCode=EGDNA_PUBLIC and https://ec.europa.eu/eurostat/web/experimental-statistics/ic-social-surveys-and-national-accounts.

9 The second wave of the Household Finance and Consumption Survey (HFCS) includes data for all countries of the euro area as well as Hungary and Poland.

it was decided that the reference period for the exercise would be the 2015 calendar year, in order to align with the most recent Household Budget Survey conducted in that year in most European countries[10]. In practice, and particularly where data on the distribution of ICW were not all collected on an annual basis, the need to have data for all three components of household economic resources meant that a different reference period, as close as possible to 2015, has been considered. An overview of countries and specific years covered by the exercise is provided in Table A.1 in Annex A.

Operational guidelines, prepared by the Eurostat-OECD Secretariat in consultation with the expert group (EG) members, detailed the set of common concepts and consistent treatments and classifications (see Box 2.2) underlying the estimates[11]. In an integrated ICW framework, it is important to ensure a high degree of consistency between the definitions underlying the different concepts, so that data for one dimension can be related to those for another dimension. Such a perspective is particularly important when considering a data item such as saving, a residual whose value is usually derived from information on other economic measures in the chosen framework[12].

Where possible, EG members were asked to share their microdata directly with Eurostat and the OECD, in order to maximise flexibility of use. However, a set of standardised templates was also developed to report semi-aggregated data in those cases where supplying anonymised microdata was not a viable option[13]. These templates collected information on a set of core variables and disaggregations (for example, by housing tenure, by household size, by age and by the education level of the household head), to provide a more granular picture of how economic resources are jointly distributed in society.

### Box 2.2. Definition of key concepts used in the ICW exercise

**Population** – all private households and their current members residing in the territory of a country. Collective and institutional households are excluded. In general, the most vulnerable parts of the population (for example, ethnic minorities and persons who are homeless) are not well covered in household surveys. Likewise, households at the very top of the wealth, income or consumption distributions are often underrepresented in the surveys, although national statistical offices have been taking steps to correct this issue by oversampling or reweighting 'rich' households.

**Unit of analysis and equivalisation** – the unit of analysis is the household, looked at as a social unit sharing income, expenses and assets among household members. Micro-datasets are thus based on the concept of non-equivalised income, consumption and wealth variables. Equivalisation (using both

10 The following wave of the EU-wide Household Budget Survey was conducted in 2020, but income and consumption data for 2020 were not available at the time of the drafting of this document.

11 The EG ICW's guidelines included: *i)* the general methodological framework for producing estimates of the joint ICW distributions (which details the income, consumption and wealth concepts to be used, as well as the selection of data sources and methodological approaches); *ii)* guidance on statistical matching and reporting using the associated quality assessment framework; and *iii)* details on the microdata to be produced and (ideally) submitted to the Eurostat-OECD Secretariat, alongside the associated metadata.

12 Only a few countries collect direct information on savings in household surveys and such information is not available in a comparative way at international level.

13 A set of SAS and R codes to perform statistical matching and automatically produce the tables in the templates was also provided to participating countries. Canada, Israel, Korea and Mexico were the only (non-EU member) countries for which anonymised microdata on the joint distribution of ICW were made available to the OECD.

the 'modified OECD scale' and 'square root scale') is applied at a later stage to produce certain risk-of-poverty and inequality indicators for individuals.

**Household reference person** – following the recommendation of the Canberra Group *Handbook on household income statistics* (United Nations Economic Commission for Europe, 2011[17]), the household member with the highest income is generally used as household reference person. In case of equal income of more than one household member, the oldest of these is defined as the household reference person. Given that most ICW statistics broken down by socio-demographic characteristics refer to the household reference person, care must be taken when interpreting them. For instance, the ICW household statistics by age refer to the 'age of the reference person', except for the indicators of risk of poverty, which refer to the age of each household member and are based on equivalised income, consumption or wealth (see Box 4.1 for more details).

**Reference period** – in line with the *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth* (OECD, 2013[3]), the recommended length of reference period is one year, with income and consumption measures ideally reflecting flows across that 12-month period. Information on wealth may refer to the stock of assets and liabilities at either the beginning or end of the reference period, or another point during the year, such as the time of interviews.

**Income concept** – the primary measure of income used is *annual household disposable income*, which includes income from employment, income from self-employment, property income and the balance between current transfers received and paid. Net owner-occupied housing services (in other words, imputed rents) are excluded from this primary measure to ensure coherence with existing international data collections. In order to increase the analytical uses of the data, it was recommended that EG members include the following variables in their microdata where available: income from employment; income from self-employment; property income; market income (the sum of income from employment, income from self-employment and property income) and gross income (the sum of market income and net current transfers received). It was also requested that the net value of owner-occupied housing services be reported as a separate variable, where available. Income from self-employment excludes the value of goods produced for own consumption in Eurostat's centralised exercise (except for Romania) and for most national data submissions. In most countries, the value of goods produced for own consumption is negligible, but there are a few exceptions (for example, Mexico and Romania) for which this value is included. If not specified otherwise, in the remainder of this paper the term 'income' refers to annual household disposable income as defined above.

**Consumption concept** – since the consumption of households cannot be measured directly, consumption expenditure is used as the nearest proxy[14]. If not otherwise specified, in the following analysis, the term 'consumption' refers to annual household consumption expenditure, defined as follows. Annual household consumption expenditures include all expenditures that are not direct investments into pensions, life insurance policies, real estate or other forms of gross capital formation. Imputed rents are excluded from consumption expenditures to ensure consistency with the income concept. Thus, only *actual* housing-related consumption expenditure (for example, actual rentals and maintenance and repair of the dwelling) is included. In addition, consumption expenditures do include consumer durables, although strictly speaking only the annual service from such durables should be

[14] It should be noted that expenditure is an imperfect measure of consumption as the amount spent by a household in a given month may differ from consumption, due to households making use of goods purchased previously or the purchase of consumer durables. In addition, consumption also includes inter-household in-kind transfers of gifts and services and social transfers in kind. However, these aspects of consumption are typically not recorded in surveys, due to the challenges of collecting this type of information (see Attanasio and Pistaferri (2016[31]) and Serafino and Tonkin (2017[26])).

regarded as 'consumption'. Non-monetary expenditures have been subtracted for all countries for which detailed data were available in Eurostat's centralised exercise[15,16].

In addition to total annual household consumption expenditure, EG members were encouraged to provide estimates for the following variables: total current expenditure (including transfers paid and interest payments relating to consumer credits); rent/mortgage interest payments for primary residence; rent/total mortgage payment (the sum of interest and any capital repayment) for primary residence; expenditure on utilities; other actual housing-related consumption expenditure; expenditure on food/beverages at home; expenditure on food/beverages outside home.

**Wealth concept** – household net wealth includes financial and non-financial assets net of liabilities. One variation from the *OECD Guidelines for Micro Statistics on Household Wealth* (OECD, 2013[3]) is the exclusion of occupational pension schemes from the measure of financial wealth (and therefore also from net wealth), in line with the concept used in the OECD Wealth Distribution Database (see Balestra and Tonkin (2018[18])). The rationale for this exclusion is that comparable data on occupational pension schemes are not always available for many countries represented in EG ICW[17]. Pension entitlements accruing under government social security schemes are also excluded from the wealth definition used in this exercise[18]. EG members were also encouraged to include the following information where available: liquid financial wealth (the sum of deposits, bonds, mutual funds, stocks, and other non-pension financial assets); non-financial assets; financial assets; liabilities. Information on liquid financial wealth was used to calculate asset-based risk-of-poverty rates, where an individual is defined as asset poor when lacking sufficient liquid financial wealth to meet their basic needs over a specified period of time (see Box 4.1 for more details).

**Savings concept** – in line with the *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth* (OECD, 2013[3]), household savings is defined as the difference

15 Bulgaria, Estonia, Greece, Spain, Finland, Croatia, Ireland, Lithuania, Latvia and Slovenia.

16 The exception is Romania, where the share of non-monetary expenditures amounts to 10% of total consumption expenditure on average, but with much larger shares at the bottom of the distribution. Since detailed data on these non-monetary expenditures are unavailable, this part is included in the consumption concept for Romania. For the calculation of saving rates, the value of goods produced for own consumption has thus been added to the Romanian income.

17 The relative importance and nature of employment-related schemes varies across countries. There are mandatory or quasi-mandatory occupational schemes in some countries (for example, Finland, Denmark, the Netherlands and Australia) while elsewhere (for example, the United States, United Kingdom, Canada and Japan) occupational pensions are voluntary. For most of the countries with available data, the impact of moving from net wealth to an 'extended net wealth' concept that includes occupational pensions is small for both wealth levels and inequality. However, including occupational pension wealth lowers significantly wealth inequality in Denmark and (to a lesser extent) in Canada, the United Kingdom, Chile, Australia and the United States. For a more detailed discussion, see Balestra and Tonkin (2018[18]).

18 The exclusion of pension entitlements in social security schemes is in line with the *OECD Guidelines for Micro Statistics on Household Wealth*. Conceptually, there are arguments both for and against their inclusion in statistics on the distribution of household wealth. On the one hand, it can be argued that all pension entitlements should be covered, irrespective of the form of scheme, as the level and nature of such entitlements will affect savings behaviour and decisions around household debt, given the considerable variation which exists in pension systems across countries, including all forms of pension wealth would arguably provide the most comparable statistics. On the other hand, it could be argued that including estimates of social security pension entitlements would be of limited use governments can always change the basis of future entitlements. It could also be argued that other social security entitlements (such as disability pensions or unemployment benefits), should be included alongside old-age pensions, particularly where they have some form of contributory element.

between *gross* income (in other words, disposable income plus direct taxes on income and social security contributions) and *total current expenditure* (in other words, consumption expenditure plus direct taxes, social security contributions, current transfers paid to other households – which all make up current transfers paid – and interest paid on consumer credit). However, the estimates for EU Member States produced by Eurostat and those for the United States diverge from this concept, defining savings as the difference between annual *disposable* income (in other words, net of direct taxes and social security contributions)[19] and *consumption expenditure* (in other words, excluding expenditures for current transfers paid by households both to the government and to other private households)[20]. The difference between the two definitions of savings boils down to the treatment of current transfers paid to other households and of interest payments relating to consumer credit, which are taken into account in the definition recommended by the OECD Framework (OECD, 2013[3]) but not in that used by Eurostat and the United States. While this should be kept in mind, on average, current transfers paid to other households and interest payments relating to consumer credit account for a low share of household expenditures.

In other words, for OECD members that are not EU Member States as well as for Finland and the Netherlands saving rates have been computed as

$$saving\ rate = \frac{gross\ income\ -\ total\ current\ expenditure}{gross\ income}$$

Instead, Eurostat and the United States calculated saving rates as

$$saving\ rate = \frac{disposable\ income\ -\ consumption\ expenditure}{disposable\ income}$$

For most of the countries covered by the exercise, information on the three economic dimensions is only available from separate surveys of different household samples. In the absence of information on the joint distribution of income, consumption and wealth available in a single instrument or different sources with common identifiers, model-based approaches (for example, statistical matching) were used to compile a fused or synthetic ICW dataset. The resulting joint distribution of income, consumption and wealth crucially depends on the quality, coherence and harmonisation of the underlying surveys, the common variables, and of the statistical process as a whole (see Subsection 3.2). Given the experimental nature of the data on the joint distribution of income, consumption and wealth compiled for this paper, several criteria were used to assess the quality and plausibility of this synthetic dataset (see Subsections 3.2.1 and 3.3.1); the uncertainty of the estimates was accounted for by computing multiple imputations and providing uncertainty intervals (for further details see Subsection 3.3.3). Where available, auxiliary information on the observed joint distribution of income and wealth from an alternative data source was used to test the conditional independence of the two variables of interest (income and consumption) given the set of common variables used for the data fusion (see Subsection 3.3.2).

[19] For Eurostat, EU-SILC variable HY020 (total disposable household income).

[20] For Eurostat, HBS variables HE00 (total consumption expenditure) minus HE042 (imputed rentals for housing) minus other non-monetary consumption.

# 3 The challenges of producing joint distributions of ICW

This section provides an overview of the main challenges faced by the members of the expert group when producing information on the joint distribution of income, consumption and wealth. These include data availability and comparability (Subsection 3.1) as well as methodological choices and statistical approaches adopted (Subsection 3.2). A special focus of this section is on statistical matching, which is the method used by most EG members to perform the exercise. An overview of the different matching techniques used by EG members, for example, parametric (predictive), non-parametric and mixed approaches, is provided alongside information on the matching variables used for each country. The section concludes with an assessment of the overall quality of the ICW estimates (Subsection 3.3).

## 3.1. Data availability and comparability

Irrespective of the method used to produce estimates of the joint distribution of income, consumption and wealth, a first concern relates to the cross-country comparability of the underlying input datasets. A longstanding tradition of measuring economic inequality via household income has led to broadly consistent, harmonised and regularly updated measures of the distribution of household income for most EU Member States and OECD countries (see, for instance, United Nations Economic Commission for Europe (2011[17]))[21]. By contrast, household expenditure surveys are typically undertaken less frequently than income surveys and their degree of harmonisation is more limited, also reflecting the lack of international standards in this field; factors such as the length of the recall period, the methods of data collection (for example, retrospective interviews versus ongoing diaries), the number of consumption items listed in the survey questionnaire, whether survey participants record their consumption themselves or are interviewed, and so on, all crucially affect the comparability of the underlying estimates. Similarly, household wealth data collections suffer from comparability issues; few countries regularly collect such data, and no international standards exist to guide national compilation efforts. Differences in the year when data are collected, in the range of assets and liabilities covered, and in the degree of oversampling of rich households all crucially affect the comparability of wealth distribution estimates across countries[22,23].

---

21 For EU Member States, the quality of income data from EU-SILC is described in national quality reports. See: https://ec.europa.eu/eurostat/web/income-and-living-conditions/quality/eu-and-national-quality-reports. For OECD countries, metadata on income and wealth statistics is available at: oe.cd/idd.

22 For wealth data from the Household Finance and Consumption Survey, quality issues are described in European Central Bank (2016[32]).

23 While no international standards currently exist for the compilation of micro statistics on household wealth, the publication of the *OECD Guidelines for Micro Statistics on Household Wealth* (OECD, 2013[3]) represents an important milestone towards that goal.

Additionally, cross-country comparability might be hampered by the use of slightly different concepts for each of the three economic dimensions (see Box 2.2). Eurostat, Austria, Canada and Finland indicated in their submissions that the concepts used for the target variables income, consumption and wealth in their respective data sources are consistent with the operational guidelines. This means that 'income' refers to annual household disposable income excluding net owner-occupied housing services (and the value of goods produced for own consumption), 'consumption' refers to total annual household consumption expenditure excluding imputed rent and non-monetary consumption, and 'wealth' refers to net wealth excluding occupational pension schemes. Conversely, the United Kingdom includes inheritance income and winnings from gambling in the income concept. The definition of wealth adopted by the United States included occupational pensions while income is net of transfers made to other households. Due to limited data availability, Switzerland reported relying on a somewhat narrower definition of wealth and using property income gross and not net of expenses. Korea adhered to the income concept but reported small divergences for consumption and wealth.

Different approaches were employed by EG members to produce estimates of the joint distributions of household economic resources, thus potentially limiting cross-country comparability. These can be classified into four broad categories:

- **Survey integration**. This refers to the *ex ante* combination of different surveys. This approach involves either getting the same (sub-)sample of the population to participate in a number of different surveys (in other words, an income survey, an expenditure survey and a wealth survey) or having surveys that cover two or more components of ICW in detail. The main potential drawback of such an approach is the sharp increase in the response burden on certain individuals/households, which may affect the response rate and ultimately the reliability of the collected data. For this reason, pure survey-integration type approaches covering all three components of ICW are very rare, though surveys collecting detailed data on two of the three components (for example, income and consumption expenditure) are more common. Among the countries covered by the exercise, only Australia and Japan can rely on a single, official household survey collecting detailed information on income, consumption and wealth for the same household. A number of countries (France, Korea, the Netherlands, Norway, Switzerland, the United Kingdom and the United States) relied, for this exercise, on an integrated data source for income and wealth, in the form of either an integrated household survey or administrative records. These countries only needed to impute consumption data to obtain the full ICW distribution. The Czech Republic, Hungary, Israel and Mexico integrated only income and consumption surveys, thus missing the wealth component.
- **Multi-source approach**. In this approach, instead of collecting all pieces of information through (face-to-face) questionnaires, as many variables as possible (in particular, economic variables) are gathered from non-survey sources, such as tax/social security registers. The integration between the survey sample and the external sources can then be performed on the basis of identifiers present in both sources or through record linkage (in other words, using information that makes it possible to identify individuals present in both data sources). Administrative sources may be considered more accurate than those collected through surveys as they are less affected by non-response or under-reporting. Moreover, this approach alleviates the response burden, although there can often be an imperfect match between the non-survey data and the concepts used in the survey. The Netherlands and Norway used record linking to combine consumption expenditure information with their integrated income and wealth databases. Likewise, Finland used registers to link income with certain wealth variables.
- **Modular approach**. Under such an approach, a single data source contains detailed information on two or more dimensions of ICW, although the selection of variables on the second or third dimension is more limited and collected through a survey module compiled by either the full sample or a subsample of respondents. For example, a survey primarily focused on income may collect

some information about consumption expenditure, for example, limited to a 'one-shot' question or focus on one aspect (such as expenditure on food), which is then integrated with more detailed expenditure data from another survey. The modular approach offers a high degree of flexibility in questionnaire design, so it is possible to reduce the impact of response burden associated with full survey integration. However, the collected data are generally less accurate than data collected using a full survey. Nevertheless, linking to the more detailed data from external surveys can be applied in order to correct data collected from modules, with the resulting imputed values being typically more accurate due to the presence of these 'hook' variables (Browning, Crossley and Weber, 2003[19]). For instance, in the case of France, household consumption expenditure was estimated by taking advantage of the few questions on consumption asked to households participating in the integrated wealth-income survey *Enquête Patrimoine* supplemented with information on consumption collected in the Household Budget Survey, following the technique developed by Browning, Crossley and Weber (2003[19]) and explained in detail in Garbinti and Lamarche (2014[20]).

- **Statistical matching and modelling**. The fourth approach consists of using the variables common to different datasets in order to merge different samples. Statistical matching techniques assume that individuals or households who turn out to be similar with respect to the common variables are also similar with respect to the variables of interest. Statistical matching is a form of model-based imputation, so there is scope to use a number of imputation techniques (such as regression models, see Subsection 3.2.4 below). However, statistical matching should generally be regarded a 'second-best' solution for the joint analysis of ICW, as it does not necessarily capture the full relationships between all the variables of interest using the set of common matching variables (which in practice can often be fairly limited, particularly where maximising the opportunities for statistical matching was not considered at the design stage of the different sources).

As anticipated, most EG members had to utilise statistical matching to a greater or lesser extent to perform the ICW exercise. In considering the potential data sources to be matched, a number of factors were considered, including the availability of potential matching variables (see Subsection 3.2.3 below) and, more generally, the likely quality of the potential matching to be performed. In general, the choice of sources was influenced by the possibility of deriving higher-quality information on the joint distribution rather than on the statistical quality of each distribution. In some cases, this resulted in selecting data sources different from those used for OECD reporting of income and wealth inequality. For example, in the case of the United States, OECD estimates on income (as available in the OECD Income Distribution Database) are based on the Annual Socio-Economic Supplement of the Current Population Survey, while for the purposes of this project the information on income is drawn from the Survey of Consumer Finances, which is also the main data source for information on household wealth.

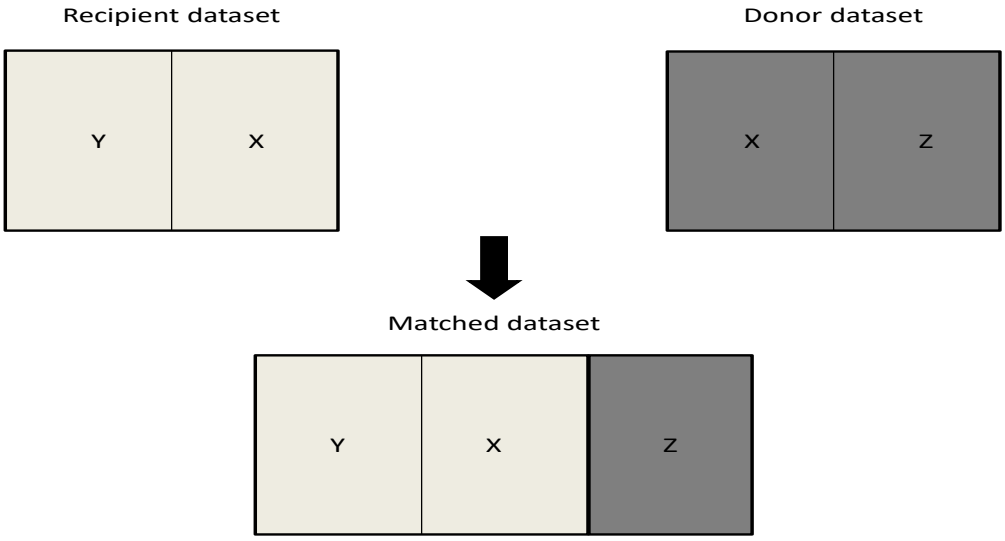
Once the sources to be used for statistical matching have been selected, the next step is to define the recipient and the donor dataset(s)[24]. The recipient dataset contains a variable Y (for example, disposable income) that is not available in the donor dataset, while variable Z (for example, consumption expenditures) is only contained within the donor dataset. The aim is to use the set of matching variables X, to link records from the donor to the recipient datasets and thereby build a complete 'synthetic'[25], matched or fused file. The matched file then contains records where X, Y and Z are jointly present (Figure 3.1).

---

[24] For convenience, it is assumed here that statistical matching is used to combine two datasets. When there are more than two datasets to be combined, one can start by combining two of them and then add one more dataset at a time until all datasets have been combined.

[25] The term 'synthetic' refers to the fact that this file is not the result of a direct observation of all the variables on a set of units belonging to the population of interest, but it is obtained by using information in the distinct files.

**Figure 3.1. Recipient and donor datasets in a statistical matching combing two datasets**




The choice of the recipient dataset is crucial: once filled with the missing variables, it will be the basis of the statistical analyses on the joint distribution of income, consumption and wealth. All members of the expert group who performed statistical matching, including Eurostat, chose the dataset containing detailed information on consumption expenditures as the donor dataset and the income or integrated income-wealth dataset as the recipient (see Table 3.1). This decision was based on the higher accuracy of the income dataset and on the main focus of the analysis presented in this paper. The relative size of the datasets was also important: using the larger dataset as the recipient provided greater precision for the analysis and prevented information loss. In order to ensure a minimum number of donor records for each corresponding group in the recipient dataset, Eurostat implemented a threshold, accepting donor samples only if they were at least one third the size of the recipient sample for each group of households.

For EU Member States, Eurostat produced joint distributions of income and consumption (and wealth, if data were available) through statistical matching in a centralised exercise (referred to as centralised estimates) based on micro-records available for the EU. However, EU members of the EG having better information at their disposal (in other words, Austria, Finland, France and the Netherlands) were encouraged to produce a separate set of estimates (referred to as national estimates). The two sets of estimates may differ from each other due to differences in data sources and methods used to produce the fused dataset underlying the estimates.

Table 3.1 provides an overview of the sources used for each country, years and dimensions covered, and of the matching techniques used.

## Table 3.1. Overview of data sources and methods used for the ICW exercise

| | Year | Dimensions | Data sources | Donor | Recipient(s) | Method used | Matching variables used, if statistical matching performed |
|---|---|---|---|---|---|---|---|
| **Australia** | 2015/16 | ICW | Household Expenditure Survey' | Statistical matching not performed (survey integration) | | | |
| **Austria*** | 2014 | ICW | EU Survey on Income and Living Conditions (I); EU Household Budget Survey (C); Household Finance and Consumption Survey (W) | Consumption data; Wealth data | Income data | Statistical matching, using random forest to select the matching variables | NUTS2 region; Educational level of household reference person; Construction year of dwelling; Size of dwelling; Housing tenure; Municipality size; Household size; Number of children. |
| **Canada*** | 2016 | ICW | Survey of Financial Security (I,W); Survey of Household Spending (C) | Consumption data | Income data; Wealth data | Statistical matching with random hot-deck | Household disposable income quintile; Household type; Housing tenure; Age of household reference person; Sex of household reference person; Region; Household size; Number of children in household; Urbanisation |
| **EU Member States (Eurostat's centralised exercise IC: Bulgaria**, Croatia**, Czech Republic, Denmark, Lithuania, Romania**, Sweden)** | See detailed year in ANNEX A Table A.1 | IC | EU Survey on Income and Living Conditions (I); EU Household Budget Survey (C). | Consumption data | Income data | Statistical matching with random hot-deck | Household type; Housing tenure and rent quintile; Age of household reference person; Educational level of household reference person; Activity status of household reference person; Occupation status of household reference person; Urbanisation; Main source of income; Income ventile |
| **EU Member States (Eurostat's centralised exercise ICW: Austria, Belgium, Cyprus**, Estonia, Finland, Germany, Greece,** | See detailed year in ANNEX A Table A.1 | ICW | EU Survey on Income and Living Conditions (I); EU Household Budget Survey (C); Household Finance and Consumption Survey (W) | Consumption data; Wealth data | Income data; | Statistical matching with random hot-deck (C) and rank hot-deck (W) | Matching variables used in the 2D Eurostat's centralised exercise (see above). Rank hot-deck based on gross income and stratification by food consumption quintile, household type and housing |

| | Year | Dimensions | Data sources | Donor | Recipient(s) | Method used | Matching variables used, if statistical matching performed |
|---|---|---|---|---|---|---|---|
| **Hungary, Ireland, Latvia, Luxembourg, Malta**, Netherlands, Poland, Portugal, Slovenia, Slovak Republic, Spain)** | | | | | | | tenure. |
| **Finland*** | 2016 | ICW | Survey on Income and Living Conditions (I); Household Budget Survey (C); Household Finance and Consumption Survey (W) | Consumption data | Income data; Wealth data | Mixed approach; Statistical matching with predictive mean matching | Ownership of a car; Household has debt; Income decile; Educational level of household reference person; Socio-economic group of household reference person; Region |
| **France*** | 2014 | ICW | *Enquête Patrimoine* (IW); Household Budget Survey (C); | Imputation of consumption based on a short module of the *Enquête Patrimoine* and on an equation built on auxiliary information from HBS data (modular approach) | | | |
| **Israel*** | 2016 | IC | Household Expenditure Survey (I,C) | Statistical matching not performed (survey integration) | | | |
| **Japan*** | 2014 | ICW | National Survey of Family Income and Expenditure (ICW) | Statistical matching not performed (survey integration) | | | |
| **Korea*** | 2015 | ICW | Survey of Household Finances (I,W); Household Income and Expenditure Survey (C) | Consumption data | Income data; Wealth data | Statistical matching with random hot-deck | Employment status of household reference person; Occupation of household reference person; Household size; Disposable income quintile |
| **Mexico*** | 2016 | IC | Survey of Household Income and Expenditure (I,C) | Statistical matching not performed (survey integration) | | | |
| **Netherlands*** | 2015 | ICW | Integral Income and Wealth Survey (I,W); Household Budget Survey (C | Consumption data | Income and wealth data | Record linking (multi-source approach) | (Encrypted) Personal ID |
| **Norway*** | 2015 | ICW | Income Statistics for Households (IW); Micro simulation model LOTTE-consumption (C) | Consumption data | Income and wealth data | Record linking (multi-source approach) | Personal ID |
| **Switzerland*** | 2015 | ICW | Survey on Income and Living Conditions (I,W); Household Budget Survey (C) | Consumption data | Income and wealth data | Statistical matching with random hot-deck | Activity status of household reference person; Age of household reference person; |

| | Year | Dimensions | Data sources | Donor | Recipient(s) | Method used | Matching variables used, if statistical matching performed |
|---|---|---|---|---|---|---|---|
| | | | | | | | Sex of household reference person; Household size; Tenure status; Decile of disposable income; Greater region; Linguistic region |
| **United Kingdom*** | 2015 | ICW | Wealth and Assets Survey (I,W); Living Costs and Food Survey (C) | Consumption data | Income and wealth data | Statistical matching with random hot-deck | Total income; Property income; Wages; Self-employment income; Benefits; Age class; Council tax amount; Pension; Gender; Ethnicity; Marital status; Region |
| **United States*** | 2016 | ICW | Survey of Consumer Finances (I,W); Consumer Expenditure Survey (C) | Consumption data | Income and wealth data | Statistical matching | Race; Education; Urban-rural status; Census division (aggregated states); Rank in the income distribution; Whether the household reported negative income; Whether the household received government transfer income; Whether the household received wage or salary income; Whether the household reported positive capital income such as interest and dividends; Whether the household reported negative capital income |

Note: * denotes national or country-based estimates; ** denotes non-OECD countries. Data on the United Kingdom are limited to Great Britain.
Source: Information compiled based on national contributions and Eurostat's experimental statistics on income, consumption and wealth.

## 3.2. Methodological choices made for the statistical matching

This subsection reviews the main methodological issues that had to be addressed before fusing two or more data sources. These include:

- differences in reference periods across data sources;
- differences in the definitions of the household reference person and members of the household;
- the comparability of common variables in the datasets;
- selection of the matching variables;
- selection of the matching method.

### *3.2.1. Reference years*

For most countries participating in the exercise, no recent **reference year** could be selected for which data on all three dimensions were available. This reflects the fact that income data are collected annually while consumption expenditure and wealth data are collected less frequently. Table A.1 in Annex A shows the selection of reference years chosen for the current exercise with 2015 being most commonly used, and 2014 or 2016 being used when no data for 2015 were available.

When data from different reference years had to be used, the EG guidelines recommended adjusting consumption data for price differences using consumer price indices. In practice, however, such adjustments were rarely performed. Canada was the only country to adjust their consumer prices and volumes for differences in the reference years. For a number of reasons, none of the other countries nor Eurostat did so. First, inflation was low in recent years, meaning that price differences in data from two different reference years are often negligible compared with differences in volumes. Second, income data are usually available every year and, when this is the case, an income reference year equal to the year of consumption data can always be chosen. Third, the reference year of consumption and wealth data are often different; in this situation, adjusting wealth data for price differences is very challenging and makes little sense due to year-to-year differences in asset prices for which it is difficult to account.

### *3.2.2. Comparability of common variables*

Different surveys may measure a number of common variables, in particular regarding demographics, household structure, and so on, that may help perform the matching exercise. However, these common variables are often inconsistent either in their definition or in their categorisation, and will require some harmonisation in terms of codification, level of aggregation, and/or format before they can be used for statistical matching. In this exercise, the common variables were selected carefully to be conceptually comparable and were brought to the same format and level of aggregation.

Since many of the common variables refer to household size and composition as well as to the characteristics of the household reference person, it was essential to first align the definitions of the household reference person and members of the household between data sources. In Eurostat's centralised matching of income (EU-SILC) and consumption (HBS) data, for example, it was necessary to define first the composition of households in a consistent way and then to select a **reference person** based on the same criteria of highest income and age (see also Box 2.2).

### *3.2.3. Selection of matching variables*

After this first harmonisation of common variables, those suitable to stratify households for the matching had to be selected. The matching variables had to meet two essential criteria*: i)* sufficient coherence of the distributions across the two data sources; *ii)* a significant correlation to the variables of interest. The

latter criterion is key, as the statistical matching assumes that the entire correlation between (for example) income and consumption goes through these matching variables (also known as the conditional independence assumption or CIA – see Subsection 3.3.2 below). The two criteria were tested using common statistical tools[26] and the best set of matching variables was selected according to their explanatory power of the variables of interest[27]. This approach was used by Eurostat for matching income and consumption data, while the matching variables uniformly chosen by Eurostat for matching income and wealth were the household type, tenure status and food consumption quintile.

The resulting number and kind of matching variables varied across countries, ranging from three (for Latvia, Lithuania and Luxembourg) to 11 (for the United States). The most common matching variables related to the socio-demographic characteristics of the household reference person and the household type or size (Table A.4 in Annex A shows the final list of matching variables used). The income quintile played a particular role as it served as a 'hook' variable between two datasets, making the conditional independence assumption much more credible.

### *3.2.4. Statistical matching method*

Several statistical matching techniques exist, as reviewed in detail by (D'Orazio, Di Zio and Scanu, 2006[21]). Eurostat and most countries participating in the exercise (see Table 3.1) used a non-parametric hot deck method to impute the missing variable of interest Z (for example, consumption expenditure) of one dataset (the donor) into the (recipient) dataset containing the other variable of interest Y (for example, income), using the set of previously-selected matching variables to narrow down possible matches. There are different types of hot deck procedures that can be used: random hot deck, rank hot deck and distance hot deck. As the name says, random hot deck randomly allocates a donor value to a recipient household after stratifying the data into homogenous subgroups. It was the main method used by countries as well as by Eurostat in the centralised exercise (see Table 3.1). In a second step, Eurostat used the rank hot deck method to join gross income and net wealth. Given that a gross income variable is available in both EU-SILC and HFCS, this variable could be used to rank households in both datasets after stratifying them according to the household type, the tenure status and the food consumption quintile. By contrast, Finland combined the advantages of parametric and non-parametric methods by using a predictive mean matching to join consumption data (HBS) to income and wealth data (integrated EU-SILC/HFCS).

### *3.2.5. Reweighting*

A final step is the modification of survey weights in the fused dataset in order to preserve the marginal distributions and certain indicators of the original surveys. Renssen (1998[22]) suggests to 'force' the weights of the matched dataset to reproduce as much as possible the marginal distribution of a given variable (for instance *Y* or *Z*) observed in the original dataset, and in this way to introduce some auxiliary information known about the joint distribution of (*Y*, *Z*) into the weights of the fused data. Eurostat, the United States and Korea recalibrated the weights of their matched dataset, using the *calib* function in R, whereas the other countries decided against such a reweighting, using the original survey weights of the recipient dataset instead.

---

[26] The similarity of the potential matching variables (first criterion) was assessed using Q-Q plots for continuous variables, Chi-2 tests or the Hellinger distance for categorical variables. For the second criterion, the Spearman's rank or Pearson correlation coefficient were used.

[27] Most countries and Eurostat used a stepwise regression to do so, whereas Austria used a random forest model.

## 3.3. Quality assessment of the matching exercise

The quality of the fused ICW dataset depends on the quality of the input data, the quality of the statistical matching process and the validity of the conditional independence assumption. The EG ICW guidelines provided a quality assessment framework to assess the fulfilment of these three conditions. This subsection first assesses how well the marginal and joint distributions of the target variable in the donor sample were preserved in the matched dataset. It then reviews the correlation structure between the variables of interest in the original and matched datasets. Finally, it provides an indication of the stability of the matching and of the plausibility of the conditional independence assumption, tested by Eurostat on a few countries with available auxiliary information.

The quality of the ICW statistics was only assessed for those countries that used a statistical matching approach to obtain joint ICW (or IC) distributions (see Table 3.1). The quality of the joint dataset for Australia, Japan, the Netherlands and Norway is deemed high, given that these countries do have fully integrated ICW data sources at their disposal or were able to produce a joint ICW dataset using record linking. The same holds true for the integrated IC surveys conducted by Israel and Mexico.

### *3.2.1. Explanatory power of matching variables*

As explained in Subsection 3.3.1, a pre-requisite for successful matching is that the common variables in the different input datasets show homogeneous distributions. The set of final matching variables should also behave as good predictors of the information to be transferred from the donor to the recipient file. The explanatory power of income and consumption by the final set of matching variables is displayed in Table A.4. The R-squared value for predicting consumption ranges from 0.26 for Belgium to 0.76 in the Austrian national exercise. It is much higher for income, ranging from 0.70 for Germany to 0.96 for Estonia, with a median value of 0.84 in Eurostat's centralised exercise. These high values can be explained by the use of the income quantile as a matching variable for all countries[28].

### *3.3.1. Quality of the statistical matching*

Ideally, the quality of the statistical matching should be assessed against several levels of validity (see Rässler (2004[23]) for further details). At a minimum, the marginal and joint distributions of the target variables in the donor dataset should be preserved in the fused file. This means that the analysis of the marginal distribution in the matched data should provide the same valid inference as those based on the original sample. The original marginal distribution should not differ from that based on the joint distribution by more than two random samples drawn from the same underlying population. As Leulescu and Agafitei (2013[12]) point out, the preservation of marginal distributions is not a sufficient quality criterion, nevertheless it is an essential one.

Table 3.2 shows the comparison of original consumption to matched consumption data for the 20th, 50th and 80th percentiles of the respective distributions. The comparison shows that the statistical matching reproduces the original distribution of consumption fairly well for all countries, with differences in marginal distributions between the original and the matched data generally remaining below 3%. Only in a few cases do the differences exceed 3%, mostly in the lower consumption groups. As such, we underestimate consumption in the matched ICW dataset is underestimated by around 4-7% in Slovenia, Malta, Croatia, Ireland and Portugal. For the United Kingdom, by contrast, consumption of the 20th percentile is overestimated by roughly 5%. It should be kept in mind though that low absolute consumption margins at the bottom of the distribution result in high relative differences between the

[28] Results of the statistical matching for Italy are not shown, due to the very low R-squared values of the regression that result from the lack of information on income quintiles in the Italian HBS.

matched and the original distribution in percentage terms even if absolute differences are very minor. For the 80th percentile, divergences of the consumption distribution in the joint ICW dataset from original consumption sources are below 3% in all countries apart from the United Kingdom[29].

## Table 3.2: Comparison between the marginal distribution of consumption in the original and matched datasets

| | | 20th percentile | | | Median | | | 80th percentile | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | original | matched | gap (%) | original | matched | gap (%) | original | matched | gap (%) |
| Eurostat's centralised exercise (euro) | Austria | 15 685 | 15 555 | -0.83 | 27 147 | 27 487 | 1.25 | 45 774 | 45 709 | -0.14 |
| | Belgium | 15 331 | 15 166 | -1.08 | 25 197 | 24 992 | -0.81 | 41 521 | 41 414 | -0.26 |
| | Bulgaria** | 2 241 | 2 212 | -1.29 | 3 969 | 3 976 | 0.18 | 6 739 | 6 779 | 0.60 |
| | Croatia** | 5 179 | 4 909 | -5.23 | 9 513 | 9 363 | -1.58 | 15 243 | 15 159 | -0.55 |
| | Cyprus** | 11 838 | 12 113 | 2.32 | 21 830 | 21 807 | -0.11 | 37 692 | 37 338 | -0.94 |
| | Czech Republic | 5 335 | 5 354 | 0.36 | 8 686 | 8 786 | 1.16 | 12 828 | 12 954 | 0.98 |
| | Denmark | 19 416 | 19 356 | -0.31 | 30 917 | 30 288 | -2.03 | 48 823 | 48 771 | -0.11 |
| | Estonia | 4 076 | 4 131 | 1.34 | 7 730 | 7 836 | 1.38 | 15 463 | 15 689 | 1.46 |
| | Finland | 14 811 | 14 708 | -0.70 | 25 681 | 26 694 | 3.94 | 44 751 | 45 775 | 2.29 |
| | France | 13 009 | 12 581 | -3.29 | 22 132 | 22 165 | 0.15 | 37 989 | 38 233 | 0.64 |
| | Germany | 13 489 | 13 629 | 1.04 | 21 935 | 22 320 | 1.75 | 35 535 | 35 824 | 0.82 |
| | Greece | 7 923 | 7 846 | -0.97 | 13 572 | 13 452 | -0.88 | 23 094 | 23 075 | -0.09 |
| | Hungary | 3 770 | 3 703 | -1.79 | 6 290 | 6 248 | -0.66 | 10 165 | 10 114 | -0.50 |
| | Ireland | 15 968 | 15 227 | -4.64 | 29 103 | 28 889 | -0.73 | 48 147 | 48 139 | -0.02 |
| | Latvia | 3 751 | 3 645 | -2.83 | 7 010 | 6 826 | -2.63 | 12 690 | 12 553 | -1.08 |
| | Lithuania | 3 373 | 3 414 | 1.21 | 6 291 | 6 468 | 2.82 | 10 973 | 10 973 | 0.00 |
| | Luxembourg | 22 900 | 23 546 | 2.82 | 39 119 | 39 319 | 0.51 | 66 761 | 67 284 | 0.78 |
| | Malta** | 9 790 | 9 198 | -6.04 | 18 530 | 18 278 | -1.36 | 31 670 | 32 101 | 1.36 |
| | Netherlands | 17 463 | 17 016 | -2.56 | 26 141 | 26 020 | -0.46 | 40 186 | 40 295 | 0.27 |
| | Poland | 4 220 | 4 166 | -1.29 | 7 023 | 6 957 | -0.94 | 11 437 | 11 414 | -0.20 |
| | Portugal | 7 541 | 7 224 | -4.20 | 13 202 | 12 811 | -2.96 | 22 773 | 22 419 | -1.55 |
| | Romania** | 2 411 | 2 393 | -0.73 | 4 014 | 3 996 | -0.45 | 6 312 | 6 323 | 0.17 |
| | Slovak Republic | 6 449 | 6 422 | -0.42 | 9 965 | 9 929 | -0.36 | 14 580 | 14 601 | 0.15 |
| | Slovenia | 8 856 | 8 243 | -6.92 | 15 589 | 15 145 | -2.85 | 24 683 | 25 305 | 2.52 |
| | Spain | 9 984 | 9 892 | -0.92 | 18 090 | 18 186 | 0.53 | 31 025 | 31 124 | 0.32 |
| | Sweden | 14 997 | 14 683 | -2.10 | 24 574 | 24 382 | -0.78 | 40 155 | 39 701 | -1.13 |
| National exercises (national currencies) | Canada | 28 270 | 28 580 | 1.10 | 48 060 | 48 100 | 0.08 | 77 020 | 77 200 | 0.23 |
| | Finland | 14 948 | 15 358 | 2.75 | 26 284 | 26 844 | 2.13 | 45 961 | 45 258 | -1.53 |
| | Korea | 815 177 | 867 063 | 0.06 | 1 774 301 | 1 787 417 | 0.01 | 2 926 392 | 2 927 020 | 0.00 |
| | Switzerland | 46 389 | 46 146 | -0.53 | 71 658 | 71 107 | -0.77 | 105 815 | 104 848 | -0.92 |
| | United Kingdom | 10 688 | 11 307 | 5.47 | 20 773 | 22 239 | 6.59 | 37 479 | 35 470 | -5.66 |
| | United States | 23 652 | 24 073 | 0.02 | 38 277 | 38 064 | -0.01 | 60 845 | 58 180 | -0.04 |

Note: Positive gaps mean that consumption is overestimated in the matched dataset; negative gaps mean that consumption is underestimated in the matched dataset.
Source: Eurostat and national calculations based on the joint ICW micro datasets.

29 In Figure A.1 and Figure A.3 in Annex A, the absolute and relative gaps between the consumption distribution in the matched ICW dataset and in the original HBS data are shown for all countries of Eurostat's centralised exercise.

Synthetic measures of inequality in consumption as derived from the matched dataset can be assessed against official estimates available internationally. Comparisons for consumption inequality focus on EU Member States only, due to the lack of cross-country comparable official statistics of household consumption beyond Europe. As shown in Figure A.5, the mean-to-median ratios are rather close between ICW and official Eurostat statistics.

### Table 3.3. Comparison between the marginal distribution of net wealth in the original and matched datasets

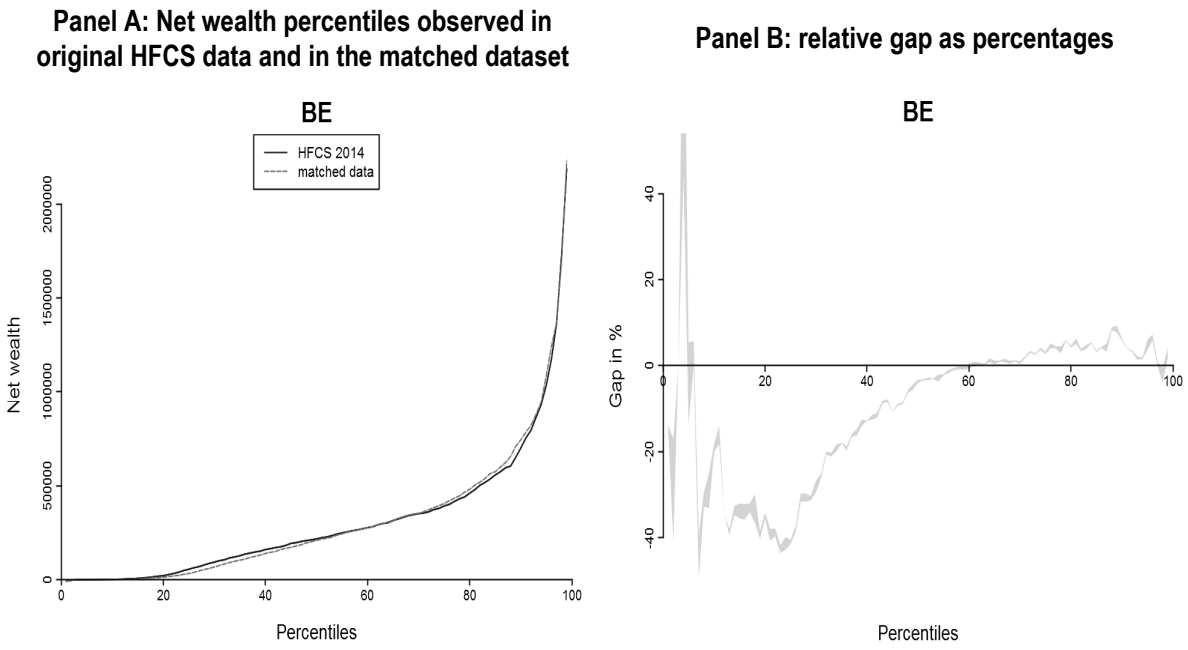
| | | 20th percentile | | | Median | | | 80th percentile | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | original | matched | gap (%) | original | matched | gap (%) | original | matched | gap (%) |
| Eurostat's centralised exercise (euro) | Austria | 6 400 | 6 113 | -4.49 | 85 860 | 101 120 | 17.77 | 364 480 | 377 253 | 3.50 |
| | Belgium | 21 850 | 14 333 | -34.40 | 217 700 | 210 462 | -3.32 | 461 000 | 482 216 | 4.60 |
| | Cyprus** | 27 833 | 17 672 | -36.51 | 173 000 | 172 281 | -0.42 | 436 198 | 454 178 | 4.12 |
| | Estonia | 6 400 | 5 954 | -6.97 | 43 771 | 41 342 | -5.55 | 113 790 | 109 697 | -3.60 |
| | Finland | 3 500 | 3 293 | -5.93 | 108 793 | 108 524 | -0.25 | 291 184 | 308 033 | 5.79 |
| | France | 9 112 | 10 644 | 16.82 | 117 551 | 136 485 | 16.11 | 349 391 | 366 796 | 4.98 |
| | Germany | 2 450 | 3 695 | 50.81 | 61 000 | 72 750 | 19.26 | 274 000 | 311 455 | 13.67 |
| | Greece | 7 000 | 7 488 | 6.98 | 65 000 | 65 020 | 0.03 | 151 425 | 158 921 | 4.95 |
| | Hungary | 6 534 | 8 255 | 26.34 | 26 203 | 27 918 | 6.54 | 68 395 | 72 614 | 6.17 |
| | Ireland | 1 820 | 2 502 | 37.47 | 100 600 | 113 353 | 12.68 | 310 320 | 336 490 | 8.43 |
| | Latvia | 1 330 | 2 690 | 102.26 | 14 207 | 15 590 | 9.73 | 43 188 | 52 478 | 21.51 |
| | Luxembourg | 32 739 | 47 694 | 45.68 | 439 859 | 438 644 | -0.28 | 1 016 000 | 1 046 614 | 3.01 |
| | Malta** | 75 366 | 78 600 | 4.29 | 211 349 | 209 658 | -0.80 | 424 700 | 432 666 | 1.88 |
| | Netherlands | 4 163 | 4 794 | 15.16 | 81 990 | 88 511 | 7.95 | 269 175 | 277 558 | 3.11 |
| | Poland | 11 053 | 16 671 | 50.83 | 57 066 | 64 813 | 13.58 | 142 156 | 154 964 | 9.01 |
| | Portugal | 7 869 | 14 735 | 87.25 | 71 101 | 84 493 | 18.83 | 202 060 | 243 783 | 20.65 |
| | Slovak Republic | 17 039 | 19 209 | 12.74 | 50 287 | 52 760 | 4.92 | 93 438 | 94 386 | 1.01 |
| | Slovenia | 14 330 | 16 263 | 13.49 | 80 295 | 86 713 | 7.99 | 170 009 | 178 997 | 5.29 |
| | Spain | 23 700 | 17 980 | -24.13 | 122 202 | 127 212 | 4.10 | 305 506 | 381 511 | 24.88 |
| National exercises (national currencies) | Canada | : | : | : | : | : | : | : | : | : |
| | Finland | : | : | : | : | : | : | : | : | : |
| | Korea | 3 382 | 3 336 | -1.35 | 17 724 | 17 572 | -0.86 | 42 796 | 42 515 | -0.65 |
| | Switzerland | : | : | : | : | : | : | : | : | : |
| | United Kingdom | : | : | : | : | : | : | : | : | : |
| | United States | : | : | : | : | : | : | : | : | : |

Note: Positive gaps mean that consumption is overestimated in the matched dataset, negative gaps mean that consumption is underestimated in the matched dataset.
Source: Eurostat and national calculations based on the joint ICW micro datasets.

Table 3.3 summarises the comparison between the distribution of net wealth in the original and matched datasets. The distribution of net wealth follows a similar pattern in the matched dataset as compared with original wealth data. However, at specific points of the distribution the relative gaps are very large, in particular at the lower end where small absolute differences result in large relative gaps. Figure 3.2 shows the absolute and relative gaps between the original and the matched distributions of net wealth for Belgium, as an example. The absolute gap below the 20th percentile is small, but it is significant in relative terms due to the very low wealth levels at the bottom of the distribution, which are close to zero below the 15th percentile. A rather large absolute gap is evident for the second quintile in the case of Belgium, but this varies from country to country. The values for other countries of the centralised exercise are in Figure A.3 and Figure A.4 in Annex A. In general, large gaps at the lower end of the

distributions are due to households holding little or no wealth. However, in Eurostat's centralised exercise wealth is often significantly overestimated in the joint distribution as compared with the original HFCS data even for the upper parts of the distribution. For the 80th percentile in Spain, Latvia and Portugal, wealth is overestimated by 20-25% in the matched data and 14% for Germany, whereas the difference between the original HFCS estimates and the matched dataset are below 10% for the other countries.

### Figure 3.2. Net wealth percentiles observed in the original and matched datasets and the corresponding relative gap, Belgium




Note: The grey area designates the 95% confidence interval computed through multiple repetitions of the random hot deck procedure.
Source: Eurostat calculations based on the joint ICW micro dataset. The original dataset is the Household Finance and Consumption Survey.

Figure A.6 compares the mean-to-median ratio of net wealth as derived in the ICW exercise with estimates drawn from the OECD Wealth Distribution Database (OECD WDD). Although the country ranking is broadly preserved and differences in levels are small for a majority of countries, a few notable exceptions stand out. Wealth inequality is lower in the ICW matched dataset than in the original ones for the Netherlands (country-based exercise) and the United States[30]. By contrast, wealth inequality for Spain is higher in the ICW data than in the WDD, even though estimates from both datasets are based on the Spanish HFCS. Differences in the Spanish (and, to a smaller extent, in the French) data require further investigation.

30 Part of the difference between ICW and WDD estimates for the Netherlands (country-based exercise) could be due to the fact that the WDD data rely on a revised series of wealth statistics, released by the CBS in the first quarter of 2021, while the ICW estimates are based on the previous data series. For the United States, the differences could partly be due to the treatment of occupational pensions, which are excluded from the net wealth definition underlying the WDD data while they are included in the ICW definition.

## Table 3.4. Spearman correlation coefficients of ordinal matching variables with consumption in the original and matched datasets

| | | Age class | | Degree of urbanisation | | Income quintiles | | Level of education | | Household size | | Tenure status | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | original | matched | original | matched | original | matched | original | matched | original | matched | original | matched |
| **Eurostat's centralised exercise** | Austria | -0.15 | -0.14 | 0.02 | 0.04 | 0.60 | 0.63 | 0.29 | 0.27 | 0.47 | 0.46 | -0.07 | -0.08 |
| | Belgium | -0.07 | -0.14 | 0.07 | 0.07 | 0.65 | 0.65 | 0.33 | 0.35 | 0.44 | 0.47 | -0.17 | -0.23 |
| | Bulgaria** | -0.46 | -0.36 | -0.25 | -0.22 | 0.86 | 0.87 | 0.33 | 0.37 | 0.60 | 0.55 | 0.11 | 0.03 |
| | Croatia** | -0.43 | -0.45 | -0.13 | -0.07 | 0.79 | 0.75 | 0.47 | 0.37 | 0.64 | 0.69 | -0.01 | 0.01 |
| | Cyprus** | -0.27 | -0.29 | -0.07 | -0.08 | 0.79 | 0.79 | 0.48 | 0.45 | 0.56 | 0.53 | 0.17 | -0.08 |
| | Czech Republic | -0.19 | -0.40 | -0.02 | 0.04 | 0.83 | 0.86 | 0.23 | 0.31 | 0.66 | 0.71 | 0.00 | -0.10 |
| | Denmark | -0.11 | -0.18 | 0.02 | -0.03 | 0.67 | 0.66 | 0.28 | 0.26 | 0.54 | 0.54 | -0.12 | -0.14 |
| | Estonia | -0.44 | -0.39 | -0.04 | -0.02 | 0.71 | 0.66 | 0.28 | 0.21 | 0.59 | 0.61 | 0.07 | 0.00 |
| | Finland | -0.26 | -0.20 | -0.04 | -0.03 | 0.74 | 0.66 | 0.33 | 0.30 | 0.62 | 0.64 | -0.13 | -0.20 |
| | France | -0.12 | -0.21 | 0.01 | -0.01 | 0.67 | 0.64 | 0.45 | 0.38 | 0.37 | 0.50 | 0.03 | -0.08 |
| | Germany | -0.06 | -0.07 | 0.05 | 0.05 | 0.78 | 0.77 | 0.21 | 0.27 | 0.56 | 0.55 | -0.05 | -0.09 |
| | Greece | -0.35 | -0.30 | -0.05 | -0.09 | 0.64 | 0.64 | 0.40 | 0.34 | 0.53 | 0.46 | 0.12 | 0.07 |
| | Hungary | -0.31 | -0.29 | -0.16 | -0.15 | 0.80 | 0.81 | 0.38 | 0.37 | 0.57 | 0.56 | 0.07 | 0.03 |
| | Ireland | -0.29 | -0.30 | -0.05 | -0.12 | 0.71 | 0.77 | 0.42 | 0.36 | 0.58 | 0.59 | 0.01 | -0.05 |
| | Latvia | -0.41 | -0.42 | -0.12 | -0.07 | 0.77 | 0.79 | 0.38 | 0.28 | 0.57 | 0.57 | -0.03 | -0.06 |
| | Lithuania | -0.35 | -0.37 | -0.15 | -0.11 | 0.70 | 0.73 | 0.39 | 0.29 | 0.57 | 0.58 | -0.03 | -0.03 |
| | Luxembourg | 0.05 | -0.03 | 0.13 | 0.07 | 0.63 | 0.67 | 0.27 | 0.23 | 0.33 | 0.40 | -0.14 | -0.19 |
| | Malta** | -0.36 | -0.39 | -0.04 | -0.03 | 0.63 | 0.61 | 0.36 | 0.33 | 0.54 | 0.57 | -0.08 | -0.16 |
| | Netherlands | 0.01 | -0.14 | 0.06 | 0.07 | 0.65 | 0.68 | n/a | 0.24 | 0.51 | 0.57 | -0.21 | -0.34 |
| | Poland | -0.33 | -0.35 | -0.06 | -0.04 | 0.72 | 0.73 | 0.37 | 0.32 | 0.50 | 0.54 | 0.04 | 0.02 |
| | Portugal | -0.32 | -0.28 | -0.17 | -0.12 | 0.66 | 0.67 | 0.44 | 0.40 | 0.44 | 0.44 | -0.01 | -0.01 |
| | Romania** | -0.38 | -0.31 | -0.26 | -0.23 | 0.84 | 0.83 | 0.48 | 0.36 | 0.59 | 0.60 | 0.10 | 0.03 |
| | Slovak Republic | -0.45 | -0.32 | 0.01 | -0.07 | 0.78 | 0.74 | 0.32 | 0.33 | 0.61 | 0.52 | 0.01 | -0.03 |
| | Slovenia | -0.38 | -0.37 | 0.02 | 0.05 | 0.70 | 0.63 | 0.37 | 0.28 | 0.62 | 0.67 | 0.18 | -0.06 |
| | Spain | -0.19 | -0.20 | -0.06 | -0.06 | 0.68 | 0.69 | 0.36 | 0.35 | 0.46 | 0.49 | 0.09 | -0.07 |
| | Sweden | -0.02 | -0.09 | -0.01 | -0.02 | 0.57 | 0.60 | 0.20 | 0.20 | 0.37 | 0.46 | -0.11 | -0.20 |

| | | Age class | | Degree of urbanisation | | Income quintiles | | Level of education | | Household size | | Tenure status | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | original | matched | original | matched | original | matched | original | matched | original | matched | original | matched |
| **National exercises** | Canada | -0.29 | -0.24 | 0.16 | 0.09 | 0.67 | 0.61 | : | : | : | : | : | : |
| | Finland | -0.27 | -0.25 | : | : | 0.73 | 0.76 | : | : | : | : | : | : |
| | Korea | -0.46 | -0.43 | : | : | 0.78 | 0.79 | : | : | : | : | : | : |
| | Switzerland | -0.17 | | : | : | 0.68 | | : | : | : | : | : | : |
| | United Kingdom | -0.21 | -0.16 | : | : | 0.70 | 0.78 | : | : | : | : | : | : |
| | United States | -0.06 | -0.10 | : | : | 0.72 | 0.70 | : | : | : | : | : | : |

Source: Eurostat and national calculations based on the joint ICW micro-datasets.

Another criterion for assessing the quality of the statistical matching in reproducing the 'true' joint distributions is to consider the correlations across variables observed in both datasets. A successful match should lead to similar relationships between common and target variables in the donor and the matched file. Table 3.4 compares the correlation of some ordinal matching variables and the consumption target variable in the original and in the matched dataset. The sign and order of magnitude of the correlations are the same in both datasets. For most countries and variables, the correlation hardly differs at all. The correlations are strongest for income quintiles and household size, whereas for most countries there is no correlation between the degree of urbanisation and consumption expenditure.

As an additional quality check, the Hellinger distance was computed on the target variables of the original and the matched datasets to examine whether their distributions are homogeneous. The Hellinger distance is a measure of the difference or similarity between two probability distributions. It is very convenient in that the distance between the distributions is expressed in a synthetic index, with a maximum value of 1 when there is no overlap in the probability distributions and a low value close to zero when the distributions are very similar[31]. It is generally considered that a Hellinger distance of over 0.05 should raise concerns about the similarities in distributions (Eurostat, 2013[11]). The low values for the Hellinger distance shown in Table 3.5 suggest that the distribution of the target variables in the original and the matched datasets are similar for all countries.

## Table 3.5. Hellinger distance for deciles of consumption and wealth (original versus matched datasets)

| | | Hellinger distance | |
|---|---|---|---|
| | | consumption deciles | wealth deciles |
| Eurostat's centralised exercise | Austria | 0.00039 | 0.00061 |
| | Belgium | 0.00053 | 0.00058 |
| | Bulgaria** | 0.00070 | : |
| | Croatia** | 0.00063 | : |
| | Cyprus** | 0.00093 | 0.00038 |
| | Czech Republic | 0.00049 | : |
| | Denmark | 0.00082 | : |
| | Estonia | 0.00081 | 0.00062 |
| | Finland | 0.00072 | 0.00043 |
| | France | 0.00040 | 0.00057 |
| | Germany | 0.00015 | 0.00024 |
| | Greece | 0.00035 | 0.00012 |
| | Hungary | 0.00033 | 0.00416 |
| | Ireland | 0.00049 | 0.00040 |
| | Latvia | 0.00062 | 0.00116 |
| | Lithuania | 0.00212 | : |
| | Luxembourg | 0.00098 | 0.00076 |
| | Malta** | 0.00065 | 0.00054 |
| | Netherlands | 0.00059 | 0.00046 |
| | Poland | 0.00023 | 0.00027 |
| | Portugal | 0.00028 | 0.00031 |
| | Romania** | 0.00045 | : |
| | Slovak Republic | 0.00099 | 0.00092 |

[31] Given two probability distributions, P and Q, the Hellinger distance is defined as $h(P,Q) = \left(1/\sqrt{2}\right) ||\sqrt{P} - \sqrt{Q}\,||_2$ and quantifies the difference between P and Q. The Hellinger distance is equal to 0 when the two probability distributions are identical and equal to 1 when there is no overlap between them.

| | | Hellinger distance | |
|---|---|---|---|
| | | consumption deciles | wealth deciles |
| | Slovenia | 0.00088 | 0.00103 |
| | Spain | 0.00025 | 0.00026 |
| | Sweden | 0.00105 | : |
| National exercises | Canada | 0.00579 | : |
| | Finland | 0.0011 | : |
| | Korea | 0.0242 | 0.0023 |
| | Switzerland | 0.00041 | : |
| | United Kingdom | 0.0083 | : |
| | United States | 0.0836 | : |

Source: Eurostat and national calculations based on the joint ICW micro-datasets.

A next validation level is to compare the correlation coefficients between income and consumption in the original donor dataset and in the fused file (see Table 3.6). This is possible due to the auxiliary information provided by the income variable in the donor dataset containing the consumption target. The comparison shows that the correlation between income and consumption is almost fully maintained in all countries. Only in a very few countries does the correlation decrease more than what would be expected from two random samples drawn from the same population.

### Table 3.6: Correlation between ordinal income and consumption ventiles in the original and matched consumption datasets

| | | Spearman correlation coefficient | |
|---|---|---|---|
| | | original | matched |
| Eurostat's centralised exercise | Austria | 0.60 | 0.63 |
| | Belgium | 0.65 | 0.65 |
| | Bulgaria** | 0.86 | 0.87 |
| | Croatia** | 0.79 | 0.75 |
| | Cyprus** | 0.79 | 0.79 |
| | Czech Republic | 0.83 | 0.86 |
| | Denmark | 0.67 | 0.66 |
| | Estonia | 0.71 | 0.66 |
| | Finland | 0.74 | 0.66 |
| | France | 0.67 | 0.64 |
| | Germany | 0.77 | 0.76 |
| | Greece | 0.64 | 0.63 |
| | Hungary | 0.79 | 0.80 |
| | Ireland | 0.70 | 0.77 |
| | Latvia | 0.77 | 0.79 |
| | Lithuania | 0.69 | 0.73 |
| | Luxembourg | 0.63 | 0.68 |
| | Malta** | 0.63 | 0.61 |
| | Netherlands | 0.65 | 0.68 |
| | Poland | 0.72 | 0.73 |
| | Portugal | 0.66 | 0.67 |
| | Romania** | 0.84 | 0.83 |
| | Slovak Republic | 0.77 | 0.74 |
| | Slovenia | 0.70 | 0.63 |
| | Spain | 0.68 | 0.69 |
| | Sweden | 0.56 | 0.60 |

| | | Spearman correlation coefficient | |
|---|---|---|---|
| | | original | matched |
| National exercises | Canada | 0.67 | 0.62 |
| | Finland | 0.73 | 0.76 |
| | Korea | 0.81 | 0.59 |
| | Switzerland | 0.69 | 0.72 |
| | United Kingdom | 0.69 | 0.78 |
| | United States | 0.72 | 0.74 |

Source: Eurostat and national calculations based on the joint ICW micro-datasets.

Additional validation levels set out in the literature (see Rässler (2004[23])), in other words, the preservation of the true joint distribution in the fused file and the preservation of the true individual values, cannot be tested for, and at least the latter is almost impossible to be achieved. One should not forget that the aim of any statistical matching exercise is not to reproduce individual values but to make valid statistical inference on the two or three economic concepts and the adherent (socio-demographic) variables in the joint dataset. It is thus important to keep in mind that the joint dataset makes it possible to produce estimates that are valid for the underlying population and for sufficiently large sub-populations, but not for small groups or individual households.

### *3.3.2. Validity of the conditional independence assumption*

A key issue for statistical matching and modelling is that the relationship between the target variables (income, consumption and wealth) can be estimated only indirectly, relying on assumptions that are difficult to test. The most common one is the conditional independence assumption (CIA), which states that the relationship between the target variables is entirely explained by the values of the common variables whereas the target variables are independent of each other. Since the CIA is a very strong assumption, which might not hold in practice, the use of auxiliary information greatly helps to make the assumption more plausible and thereby improve the quality of the matching. Such auxiliary information can be either a proxy variable of Z in the recipient dataset or a proxy of Y in the donor dataset (see Figure 3.1). Eurostat and all EG members relying on statistical matching used an income proxy available in the consumption dataset as matching variable to relax the CIA. Moreover, Eurostat used gross income, available in both EU-SILC and HFCS, as a proxy for total disposal income and food consumption, available in both HBS and HFCS, as a proxy for consumption to join wealth data to the income-consumption dataset.

In 2017, some EU Member States collected data on ‘over-indebtedness, consumption and wealth’ (OCW) through an ad hoc module of the EU-SILC survey. This OCW module, containing limited information on consumption and wealth together with details on income, enabled Eurostat to test the effectiveness of using proxy variables of the targets as matching variables to justify the CIA. Using partial correlations between the target variables, for given values of the proxy variables, Eurostat could prove that the CIA hypothesis holds, meaning that independence between the target variables can be assumed when controlling for the matching variables (Lamarche, Oehler and Rioboo, 2020[4]). The results indicate that the partial correlation between the target variables, given the proxy variables, are very low, being in most cases not statistically significant. This implies that the existing relationship among these target variables is mainly explained by the proxy variables.

### *3.3.3. Uncertainty of the estimates*

Subsection 3.3.1 showed that the statistical matching procedures applied by Eurostat and the EG members relying on statistical matching resulted in joint distributions that are comparable to those in the original data. In Subsection 3.3.2, the conditional independence assumption underlying the procedure was discussed and (as far as possible) validated. This last section of the quality assessment reviews the general uncertainty that characterises the ICW estimates as a result of imputing the missing target variable through statistical matching, a procedure that is unable to discriminate among a set of plausible (joint) distributions for (Y, Z) given X. This means that the statistical matching picks a specific joint distribution (y,z | x) and uses it as if it was the true joint distribution, although other options within the boundaries set by the matching variables would have been possible (see also Conti, Marella and Neri (2017[24])). Two tools were used to quantify the uncertainty range: multiple imputations and Fréchet bounds[32].

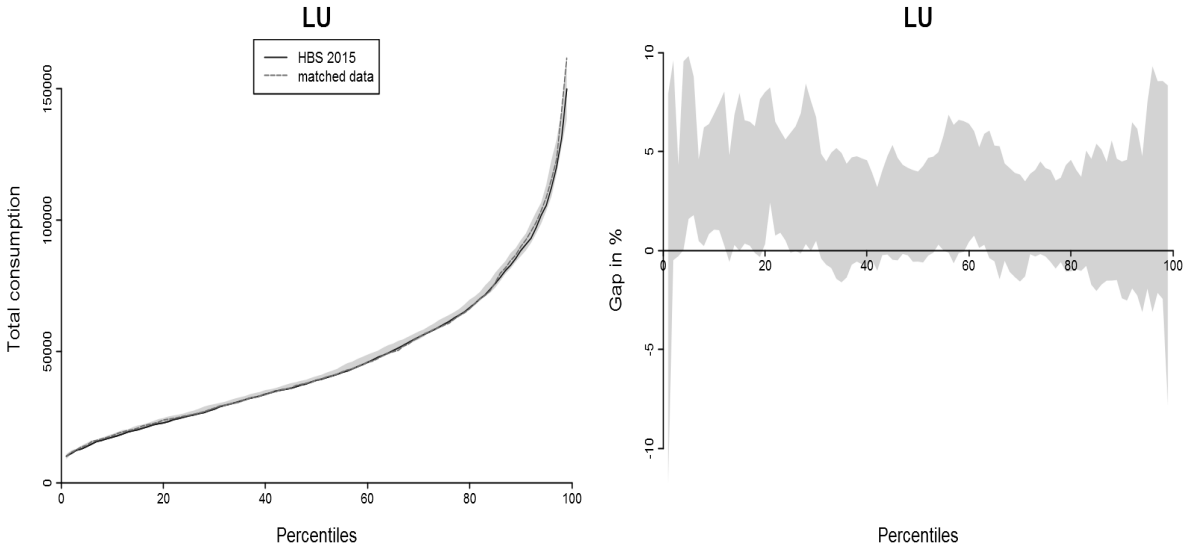
The EG operational guidelines recommended experts to perform 1 000 imputations in the case of a single matching, and 100*100 in the case of a double matching, in order to evaluate the dispersion of the values estimated through the matching. The **multiple imputation** results make possible the derivation of confidence intervals containing – in this case – 95% of the estimates that may result from the matching. Using 100*100 imputations, however, significantly increases the size of intermediary datasets and the processing time. Eurostat thus reduced the 100*100 imputations to 30*30 after the comparison of results for two test datasets had shown that the confidence intervals did not significantly vary using fewer imputations. The confidence intervals obtained through the multiple imputations shown by two examples in Figure 3.2 and Figure 3.3 are rather narrow, indicating that there is limited variability induced by the statistical matching once data have been stratified by the set of matching variables. Instead of selecting one specific joint distribution from the set of plausible ones which are all equally likely to be true, Eurostat used the mean value of the multiple imputations for all indicators shown in Section 4.

[32] Fréchet bounds are interval estimates which can be applied in a non-parametric setting to measure the uncertainty range. In our case they provide lower and upper bounds for the contingency table that crosses income and consumption or wealth quintiles (see also D'Orazio, Di Zio and Scanu (2006[21])). The intervals correspond to the probability of a household falling into the respective cell, for example the first quintile of income and consumption.

## Figure 3.3. Consumption percentiles observed in the original and matched datasets and the corresponding relative gap, Luxembourg

**Panel A: Total consumption percentiles observed in original HFCS data and in the matched dataset**

**Panel B: relative gap as percentages**




Note: The grey area in Panel B indicates the 95% confidence interval computed through multiple repetitions of the random hot-deck procedure.
Source: Eurostat calculations based on the joint ICW micro dataset. The original dataset is the Household Budget Survey.

**Fréchet bounds** were computed from the observed marginal distributions of two target variables (income, consumption and/or wealth). This made it possible to measure the degree of uncertainty of estimates belonging to a certain income and consumption or wealth quantile, by providing upper and lower bounds for the contingency table. Each bound is an interval for each pair of variables' categories, and it provides a lower and a higher value for their probability of occurrence, covering all the possible values that are compatible with the data available and thus providing an uncertainty space. The inclusion of the set of matching variables in the estimation of the cells probabilities improves the bounds by narrowing them. Table 3.7 shows the Fréchet bounds for a selection of categories for pairs of variables. Generally, the uncertainty of the estimates is lower for income-consumption than for income-wealth. Indeed, for the 1st and 10th income and consumption deciles, the bounds are very narrow, at least when conditioned on the matching variables. Larger, but still acceptable, bounds are found for the 1st and 10th income and wealth deciles. For most countries, relatively narrow Fréchet bounds delimit the probability range for a household to fall into the following categories: 'at risk of income and consumption poverty', defined as being below the twofold risk-of-poverty threshold of 60% of median income and consumption; and 'at risk of income and asset poverty', defined as being below the risk-of-poverty threshold of 60% of median income and lacking the liquid financial assets needed to support someone at the level of the income risk-of-poverty threshold. Note that, when not conditioned on the matching variables, the difference between the upper and lower bounds becomes larger.

## Table 3.7. Estimated Fréchet bounds (conditioned on the list of matching variables) for a selection of categories of pairs of variables

Pairs of variables refer to the binary 'at risk of income and consumption poverty' (IC poverty) and 'at risk of income and asset poverty' (IW poverty) variables, as well as to the bottom and top deciles of the joint distributions of income and consumption (IC) and of income and wealth (IW)

| | | IC poverty | | IW poverty | | IC 1st decile | | IW 1st decile | | IC 10th decile | | IW 10th decile | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | lower | upper | lower | upper | lower | upper | lower | upper | lower | upper | lower | upper |
| Eurostat's centralised exercise | Austria | 0.0378 | 0.0433 | 0.0838 | 0.1034 | 0.0330 | 0.0333 | 0.0288 | 0.0407 | 0.0366 | 0.0371 | 0.0235 | 0.0403 |
| | Belgium | 0.0325 | 0.0529 | 0.0916 | 0.1227 | 0.0375 | 0.0378 | 0.0368 | 0.0525 | 0.0374 | 0.0376 | 0.0119 | 0.0358 |
| | Bulgaria** | 0.0739 | 0.0807 | : | : | 0.0498 | 0.0499 | : | : | 0.0496 | 0.0500 | : | : |
| | Croatia** | 0.0882 | 0.1032 | : | : | 0.0550 | 0.0551 | : | : | 0.0475 | 0.0476 | : | : |
| | Cyprus** | 0.0672 | 0.0849 | 0.0125 | 0.1714 | 0.0477 | 0.0481 | 0.0002 | 0.0543 | 0.0376 | 0.0388 | 0.0000 | 0.0678 |
| | Czech Republic | 0.0319 | 0.0383 | : | : | 0.0580 | 0.0580 | : | : | 0.0576 | 0.0576 | : | : |
| | Denmark | 0.0272 | 0.0298 | : | : | 0.0356 | 0.0364 | : | : | 0.0315 | 0.0321 | : | : |
| | Estonia | 0.1157 | 0.1418 | 0.1419 | 0.1904 | 0.0395 | 0.0397 | 0.0264 | 0.0391 | 0.0387 | 0.0389 | 0.0275 | 0.0372 |
| | Finland | 0.0484 | 0.0598 | 0.0576 | 0.1179 | 0.0357 | 0.0362 | 0.0000 | 0.0442 | 0.0441 | 0.0450 | 0.0000 | 0.0764 |
| | France | 0.0402 | 0.0576 | 0.0969 | 0.1227 | 0.0372 | 0.0379 | 0.0195 | 0.0348 | 0.0394 | 0.0399 | 0.0339 | 0.0545 |
| | Germany | 0.0667 | 0.0745 | 0.1303 | 0.1468 | 0.0622 | 0.0625 | 0.0099 | 0.0259 | 0.0478 | 0.0492 | 0.0201 | 0.0469 |
| | Greece | 0.0490 | 0.0524 | 0.0745 | 0.1160 | 0.0262 | 0.0263 | 0.0118 | 0.0334 | 0.0384 | 0.0388 | 0.0148 | 0.0327 |
| | Hungary | 0.0548 | 0.0726 | 0.0520 | 0.1104 | 0.0520 | 0.0520 | 0.0102 | 0.0329 | 0.0527 | 0.0528 | 0.0321 | 0.0570 |
| | Ireland | 0.0668 | 0.0791 | 0.1062 | 0.1282 | 0.0456 | 0.0461 | 0.0029 | 0.0055 | 0.0407 | 0.0414 | 0.0166 | 0.0393 |
| | Latvia | 0.1199 | 0.1372 | 0.1870 | 0.2187 | 0.0489 | 0.0490 | 0.0225 | 0.0304 | 0.0542 | 0.0543 | 0.0354 | 0.0495 |
| | Lithuania | 0.0858 | 0.1020 | : | : | 0.0350 | 0.0350 | : | : | 0.0446 | 0.0447 | : | : |
| | Luxembourg | 0.0523 | 0.0730 | 0.0825 | 0.1167 | 0.0475 | 0.0477 | 0.0210 | 0.0402 | 0.0402 | 0.0404 | 0.0268 | 0.0449 |
| | Malta** | 0.0511 | 0.0731 | 0.0711 | 0.1027 | 0.0334 | 0.0342 | 0.0265 | 0.0338 | 0.0337 | 0.0344 | 0.0264 | 0.0328 |
| | Netherlands | 0.0127 | 0.0242 | 0.0964 | 0.1189 | 0.0345 | 0.0345 | 0.0053 | 0.0213 | 0.0472 | 0.0473 | 0.0081 | 0.0276 |
| | Poland | 0.0513 | 0.0744 | 0.0697 | 0.1218 | 0.0535 | 0.0537 | 0.0241 | 0.0421 | 0.0490 | 0.0494 | 0.0158 | 0.0325 |
| | Portugal | 0.0653 | 0.0807 | 0.1337 | 0.1677 | 0.0392 | 0.0394 | 0.0165 | 0.0318 | 0.0426 | 0.0431 | 0.0251 | 0.0443 |
| | Romania** | 0.0895 | 0.1015 | : | : | 0.0620 | 0.0626 | : | : | 0.0565 | 0.0571 | : | : |
| | Slovak Republic | 0.0181 | 0.0476 | 0.0392 | 0.0904 | 0.0597 | 0.0598 | 0.0144 | 0.0269 | 0.0381 | 0.0381 | 0.0263 | 0.0427 |
| | Slovenia | 0.0463 | 0.0566 | 0.0350 | 0.1657 | 0.0382 | 0.0385 | 0.0007 | 0.0588 | 0.0406 | 0.0411 | 0.0000 | 0.0708 |
| | Spain | 0.0935 | 0.1103 | 0.1364 | 0.1701 | 0.0435 | 0.0436 | 0.0192 | 0.0335 | 0.0452 | 0.0454 | 0.0345 | 0.0472 |
| | Sweden | 0.0487 | 0.0712 | : | : | 0.0417 | 0.0417 | : | : | 0.0347 | 0.0348 | : | : |

| | | IC poverty | | IW poverty | | IC 1st decile | | IW 1st decile | | IC 10th decile | | IW 10th decile | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | lower | upper | lower | upper | lower | upper | lower | upper | lower | upper | lower | upper |
| National exercises | Canada | 0.0196 | 0.0348 | 0.0332 | 0.1204 | 0.0339 | 0.0542 | 0.0272 | 0.0522 | 0.0223 | 0.0527 | 0.0165 | 0.0442 |
| | Finland | : | : | : | : | : | : | : | : | : | : | : | : |
| | Korea | 0.1266 | 0.1500 | : | : | 0.0390 | 0.0597 | : | : | 0.0135 | 0.0560 | : | : |
| | Switzerland | 0.0314 | 0.0357 | : | : | 0.0354 | 0.0367 | : | : | 0.0356 | 0.0370 | : | : |
| | United Kingdom | : | : | : | : | : | : | : | : | : | : | : | : |
| | United States | 0.0893 | 0.1441 | 0.1803 | 0.2209 | 0.0041 | 0.0736 | 0.0016 | 0.0431 | 0.0299 | 0.1068 | 0.0091 | 0.0776 |

Note: Fréchet bounds conditioned on the list of matching variables in Table A.4.
Source: Eurostat and national calculations based on the joint ICW micro-datasets.

# 4 Key empirical results

The objective of this section is to highlight how an integrated framework can enhance understanding both of how economic resources are distributed in society and of the heterogeneity in household behaviour, for instance with respect to saving rates and household spending. As mentioned already, caution should be taken when reviewing the ICW evidence below, as it relies on estimates that are experimental in nature and based on assumptions and methods that are being tested and still subject to modification. Moreover, cross-country comparability remains somewhat limited (for example, due to different I, C, and/or W concepts used or the statistical methods applied to derive the joint distribution estimates). It should also be borne in mind that the evidence on multi-dimensional inequality reviewed below partly reflects the nature of the underlying data sources on income, consumption and wealth. For instance, the finding that the United States systematically stands out as the most unequal country in a multi-dimensional setting could be partly driven by the fact that the US Survey of Consumer Finances – from which information on both household income and wealth is taken – oversamples the rich, which is not always the case with other country surveys.

In general, national estimates produced by some EU Member States are shown in the following analysis instead of those produced by Eurostat's centralised exercise[33]. This is the case for Finland, France and the Netherlands. It is worth noting that the two sets of estimates may differ from each other due to differences in data sources and methods used to produce the matched dataset underlying the estimates. In agreement with the Austrian experts, Eurostat's estimates are used instead of national estimates in the following analysis.

The analysis in this section is based on semi-aggregated tabulations. This approach makes it possible to cover countries for which sharing anonymised microdata was not a viable option but comes at the expense of limiting the depth of the analysis. For instance, while semi-aggregated tabulations can depict differences across population groups, they clearly cannot account for disparities within groups and they cannot be used to analyse how the interaction of different household and individual characteristics may impact the joint distribution of ICW.

## 4.1. How correlated are household income, consumption and wealth?

A dimension-by-dimension approach to inequality may lead to a mixed picture. For example, Nordic countries typically exhibit low levels of income inequality but highly skewed wealth distributions. Looking at inequalities in an integrated framework may either help solve this and other puzzles or reinforce the evidence based on uni-dimensional metrics. At household level, the association between income, consumption and wealth is difficult to predict. On the one hand, the three economic dimensions may be expected to be highly related at the household level, since people with higher incomes are more likely to accumulate wealth, and both resources can be used to support high spending. On the other hand, life-cycle effects may weaken the association between the three variables: for instance, people early in

33 For Eurostat's centralised results see also the experimental ICW statistics published at: https://ec.europa.eu/eurostat/web/experimental-statistics/income-consumption-and-wealth.

their careers may have high incomes but had little time to build wealth, while retired people tend to have lower incomes but often hold substantial wealth.

The analysis of inequality in a multi-dimensional setting first focuses on the three different pairs – IW, IC and CW – and then moves to the joint analysis of the three aspects, ICW. Two different indicators are considered: *i)* the share of *households* at different points of the joint distribution (for example, share of households in the top 20% (or top 10%) and in the bottom 20% (or bottom 10%) of the joint distribution); and *ii)* the share of *resources* held by those same households.

### *4.1.1. Inequality in two dimensions*

Figure 4.1 shows the share of households that belong to the highest and lowest quintile of the joint distribution of **income and wealth** (Panel A), as well as the share of those that are in the top 10% of the joint distribution (Panel B). If the association were perfect, then all the households in a given quintile (decile) of the income distribution would belong to the same quintile (decile) of the wealth distribution. On the contrary, if the association were nihil, then a household in a given quintile (decile) of the income distribution would be equally likely to belong to any of the quintiles (deciles) of the wealth distribution. The red horizontal line in Figure 4.1 Panel A, set at 4%, represents the case of no correlation between households' positions in the distribution of income and of wealth (when analysed by quintiles). By contrast, the closer the bars are to 20% (the perfect correlation scenario when analysed by quintiles) the stronger is the association between the ranking of households in the distribution of income and their ranking in the distributions of wealth[34]. Likewise, the black horizontal line at 1% in Figure 4.1 Panel B represents the case where the top 10% of households is randomly distributed along the joint distribution, whereas a perfect association would mean that the share of households in the top 10% of both income and wealth is 10%.

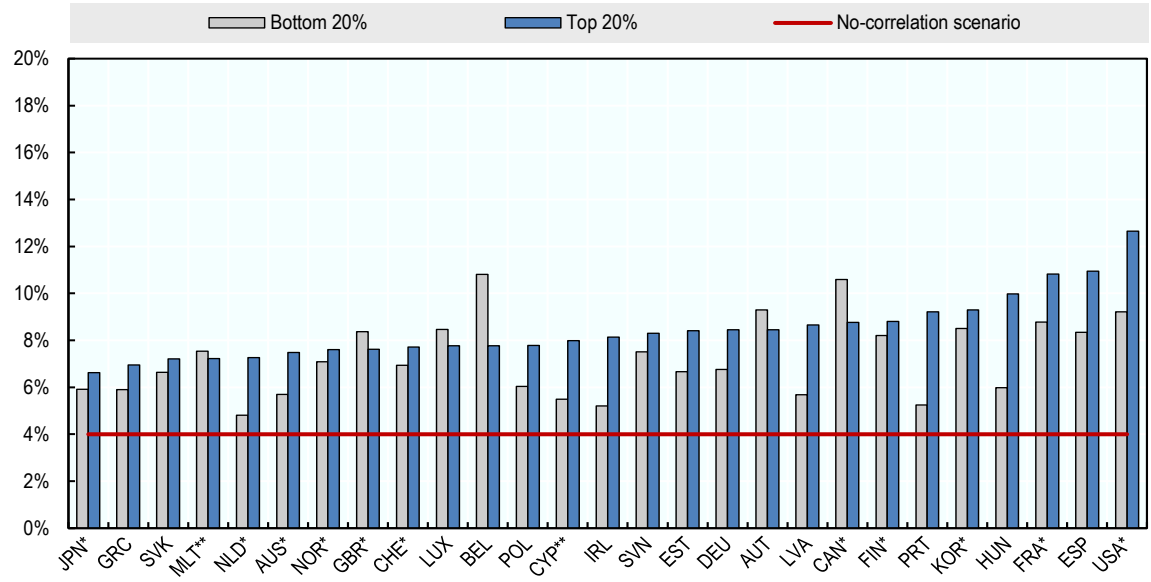
*Income and wealth tend to be more correlated at the top than at the bottom of the joint distribution*

Not surprisingly, across countries, there is a positive (although far from perfect) correlation between a household's position on the income ladder and its position in the wealth distribution. The share of households that are in the top 20% of the joint income and wealth distribution ranges from a low of 7% in Japan to a high of 13% in the United States. In a number of countries, such as Malta, Finland and Korea, the association is equally strong at both ends; in some other countries, including Belgium and Canada, the association is stronger at the bottom of the distribution, although in the majority of countries the reverse holds true (Figure 4.1 Panel A). The association at the top end of the joint distribution is stronger when the highest decile is considered. As shown in Figure 4.1 Panel B, in all countries, households in the top 10% of the income distribution are also likely to be among those belonging to the top 10% of the wealth distribution, with this probability being strongest in France, Spain and the United States.

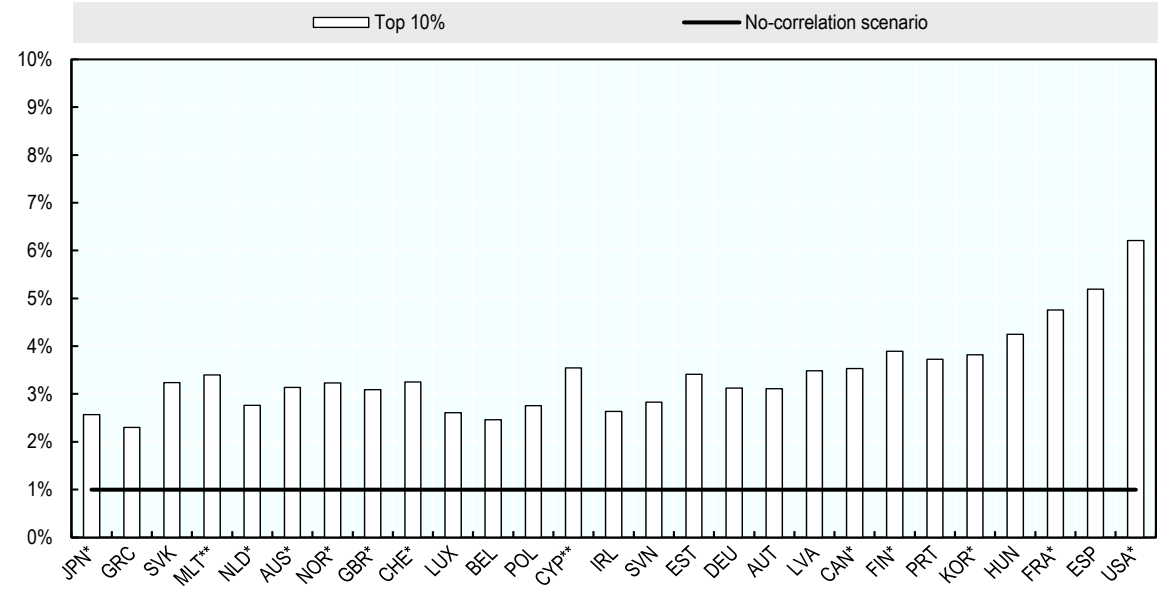
[34] There are 25 possible combinations of quintiles characterising the joint distribution of income and wealth, in other words, 5 (quintiles) to the power of 2 (distributions). This implies that, if there were zero correlation between income and wealth, a household in a given quintile of the income distribution would be equally likely to belong to any of the 25 possible combinations, which would then hold 4% of households each (in other words a 1/25 share).

## Figure 4.1. Share of households at different points of the joint distribution of income and wealth, 2015 or closest available year

Panel A: Share of households in the bottom and top 20% of the joint distribution




Panel B: Share of households in the top 10% of the joint distribution




Note: In both panels, countries are ranked in ascending order of the share of households belonging to the top 20% of the joint distribution of income and wealth. * denotes national or country-based estimates, ** denotes non-OECD countries. Panel A: the horizontal line set at 4% represents the case of no correlation between a household's quintile in the distribution of income and its quintile in the distribution of wealth; while the closer the bars are to 20% (the perfect correlation scenario) the stronger is the association between a household's quintile in the distribution of income and its quintile in the distribution of wealth. Panel B: the horizontal line set at 1% represents the case of no correlation between a household's decile in the distribution of income and its decile in the distribution of wealth. The closer the bars are to 10% (the perfect correlation scenario) the stronger is the association at the top of the two distributions. A full list of countries and their three-letter country codes is provided on page 5(.

Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: Panel A[icw_res_01] and Panel B [icw_res_01]).

Focusing only on the top and bottom of the distribution presents an incomplete picture. Figure B.1 in Annex B shows two-dimensional inequality across the entire distribution using quintile on quintile tabulations, which show the share of households in the various quintiles of the joint distribution. In case of a perfect association between households' income and wealth positions, the observations in Figure B.1 would be entirely concentrated in the diagonal cells from top left to bottom right (from quintile pairs 1,1 to 5,5), each with a value of 20, while the off-diagonal cells would have no observations and each have a value of 0. By contrast, if income and wealth were not correlated at all, each cell in the matrix would contain 4 % of households. In practice, the association between income and wealth is stronger at the tails (around quintile pairs 1,1 and 5,5) than around the middle and a sizeable share of households is found throughout the distribution. For example, in Australia, Cyprus, Finland and Switzerland, about 10% of households in the bottom wealth quintile are located in the second or third income quintiles (see Figure B.1).

This type of distribution indicates that, on the one hand, some low-income households are able to rely on wealth to support consumption and income fluctuations, while on the other hand, households with little wealth and high income are vulnerable to sudden income or labour-market shocks[35]. Life-cycle considerations likely play a role in explaining this pattern: young households with high income (prospects) still have to accumulate wealth or may be highly leveraged. Figure B.2 shows the joint distribution of income and wealth among the working-age population and confirms that households with a working-age head tend to be overrepresented among those with high-income and low-wealth levels (see for example Canada and Australia).

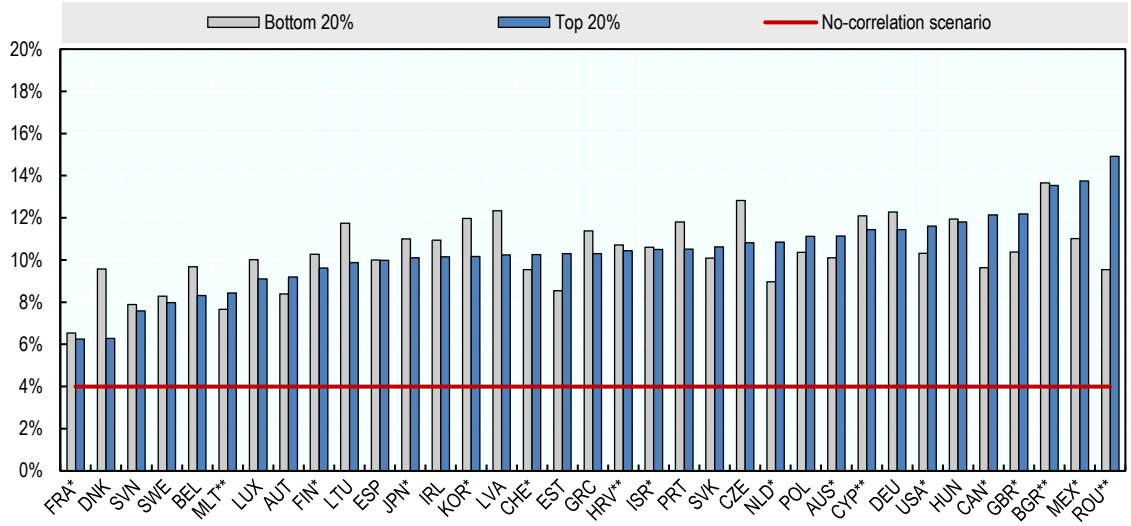
### *In a majority of countries, the association between income and consumption is equally strong at both tails of the joint distribution*

The joint distribution of **income and consumption** shows a number of distinctive findings. First, in most countries the probability for a household in the lowest quintile of the income distribution to be in the lowest quintile of the consumption distribution is about the same as that of a high-income household to be among those with the highest consumption levels (Figure 4.2, Panel A). Moreover, for most countries the association between income and consumption seems stronger than that between income and wealth, regardless of where along the joint distribution it is computed (compare Figure B.1 and Figure B.3). In addition, the association tends to be stronger for working-age households than for the total population (compare Figure B.3 and Figure B.4). Finally, as with income and wealth, the association between income and consumption is even stronger in the top decile than in the top quintile of the joint distribution (Figure 4.2, Panel B).

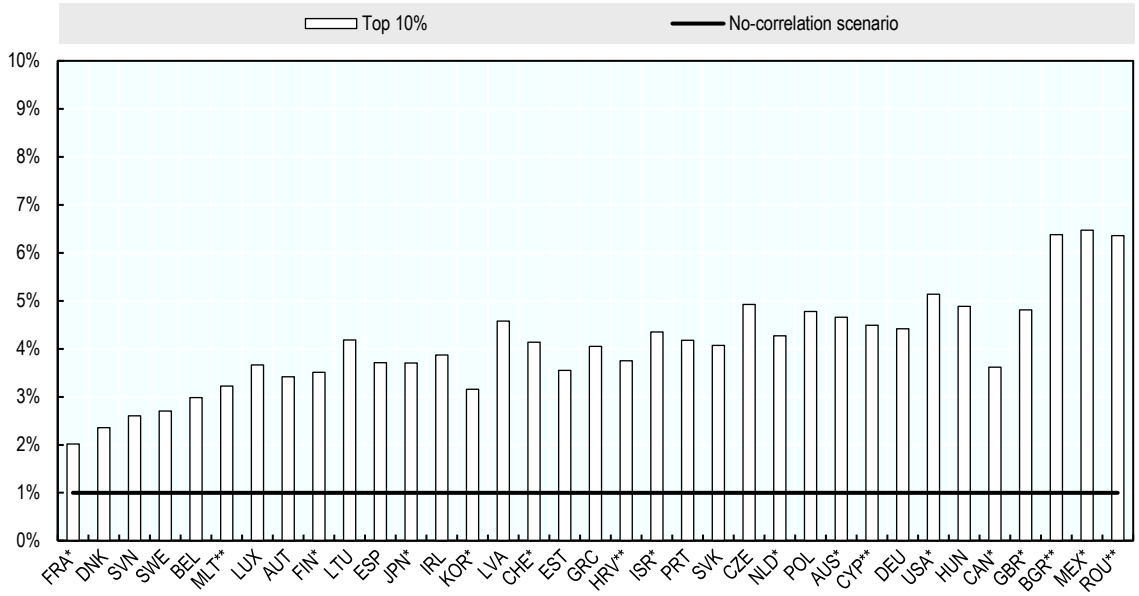
35 This finding is in line with evidence shown in Balestra et al. (forthcoming[33]), based on HFCS wealth data and disposable income imputed via a machine learning model.

## Figure 4.2. Share of households at different points of the joint distribution of income and consumption, 2015 or closest available year

Panel A: Share of households in the bottom and top 20% of the joint distribution




Panel B: Share of households in the top 10% of the joint distribution



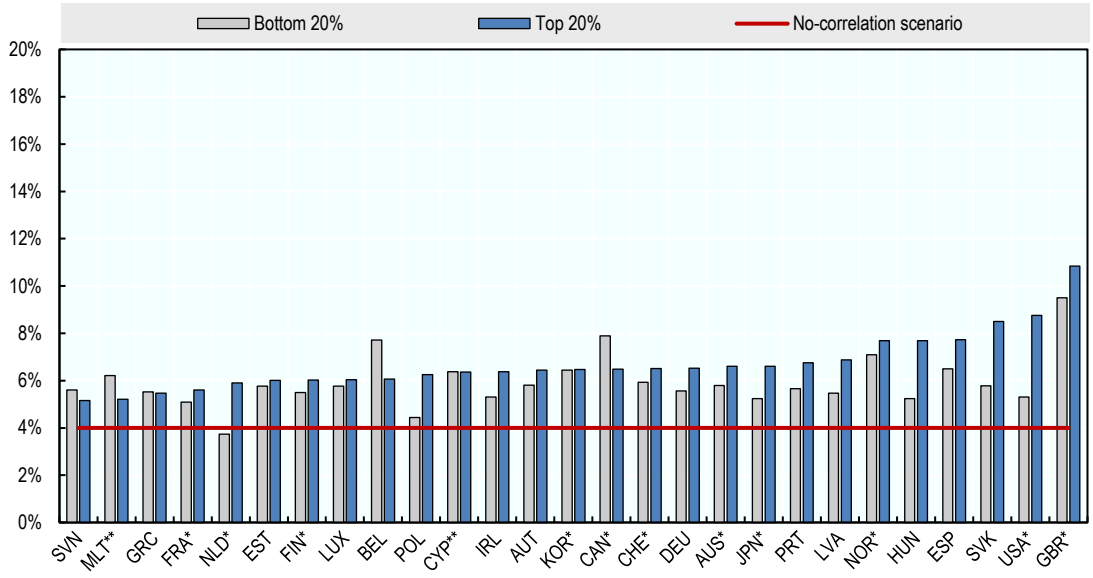

Note: In both panels, countries are ranked in ascending order of the share of households belonging to the top 20 % of the joint distribution of income and consumption. Panel A: the horizontal line set at 4% represents the case of no correlation between a household's quintile in the distribution of income and its quintile in the distribution of consumption; while the closer the bars are to 20% (the perfect correlation scenario) the stronger is the association between a household's quintile in the distribution of income and its quintile in the distribution of consumption. Panel B: the horizontal line set at 1% represents the case of no correlation between a household's decile in the distribution of income and its decile in the distribution of consumption. The closer the bars are to 10% (the perfect correlation scenario) the stronger is the association at the top of the two distributions.

Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: Panel A [icw_res_01] and Panel B [icw_res_01]).

*Wealth and consumption tend to be more correlated at the top than at the bottom of the joint distribution*

In most countries, the association between **wealth and consumption** tends to be stronger at the top than at the bottom of the joint distribution (Figure 4.3). While the shares of households belonging to the top 20% of the joint distribution of wealth and consumption are similar to those observed for the joint distribution of income and wealth, country rankings are not always preserved. However, the United Kingdom and the United States hold on to their position among the countries with the highest shares of households in the top 20% of joint distributions. In the Netherlands and Poland, households in the lowest quintile of the wealth distribution are equally likely than their wealthier peers to belong to the lowest quintile of the consumption distribution (shares around 4%, the no-correlation scenario)[36].

## Figure 4.3. Share of households at different points of the joint distribution of wealth and consumption, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of households belonging to the top 20% of the joint distribution of wealth and consumption. The horizontal line set at 4% represents the case of no correlation between a household's quintile in the distribution of wealth and its quintile in the distribution of consumption; while the closer the bars are to 20% (the perfect correlation scenario) the stronger is the association between a household's quintile in the distribution of wealth and its quintile in the distribution of consumption.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

When examining the entire joint distributions, the higher values on the diagonals (from quintile pair 1,1 to 5,5) for the joint income and consumption distributions compared with the wealth and consumption distributions indicate that consumption and income are more closely linked at the household level than consumption and wealth. This result holds for both the total population and for working-age households (see Figure B.5 and Figure B.6, respectively), and it is in line with the literature on marginal propensities

36 Evidence on the association between wealth and consumption in the top decile of the joint distribution is not shown in Figure 4.3, since wealth and consumption are not directly matched, and the uncertainty related to the statistical matching prevents focussing on the top tail of the joint distribution.

to consume, which postulates a higher marginal propensity to consume out of wealth than out of income (see, for example, Carroll et al. (2017[25])).

*Wealth is more concentrated than income and consumption at the top of the joint distributions*

When the *share of income, consumption and wealth* held by households in the top quintile of the different joint distributions reviewed above is considered, the usual finding of household **wealth** being more concentrated at the top than **income** is confirmed in the joint dataset. In general, 8% of households hold at the same time 19% of total income and almost 35% of total wealth. The largest concentration of resources at the top of the joint distribution is in the United States, where 13% of households hold almost half of total income and more than three quarters of total wealth (Table 4.1). The concentration of resources among the top 20% of the joint distribution of income and wealth is also high in the United Kingdom, Switzerland and the Netherlands[37]. When the top 10% of the joint distribution of income and wealth is considered, the picture becomes even more unequal (Figure B.7), with a mere 3% of households (right vertical axis) simultaneously holding almost 10% of the average country's income and more than a fifth of wealth (left vertical axis). Again, the United States, the United Kingdom and Switzerland stand out as the countries where economic resources are most concentrated at the top of the joint distribution.

**Table 4.1. Share of income, consumption and wealth held by households in the top quintile of the joint distributions, total population, 2015 or closest available year**

| | IW joint distribution | | | IC joint distribution | | | CW joint distribution | | |
|---|---|---|---|---|---|---|---|---|---|
| | Share of households | Share of income | Share of wealth | Share of households | Share of income | Share of consumption | Share of households | Share of consumption | Share of wealth |
| **Australia*** | 7.5% | 18.1% | 31.6% | 11.1% | 24.6% | 23.7% | 6.6% | 14.8% | 28.1% |
| **Austria** | 8.5% | 16.8% | 35.3% | 9.2% | 18.3% | 18.2% | 6.4% | 12.6% | 23.2% |
| **Belgium** | 7.8% | 14.7% | 28.4% | 8.3% | 15.4% | 17.5% | 6.1% | 11.6% | 18.3% |
| **Bulgaria**** | : | : | : | 13.5% | 30.6% | 27.6% | : | : | : |
| **Canada*** | 8.8% | 21.2% | 37.0% | 12.1% | 24.2% | 19.2% | 6.5% | 9.9% | 24.1% |
| **Czech Republic** | : | : | : | 10.8% | 21.5% | 18.7% | : | : | : |
| **Croatia**** | : | : | : | 10.4% | 20.6% | 19.7% | : | : | : |
| **Cyprus**** | 8.0% | 18.0% | 32.6% | 11.4% | 23.0% | 23.1% | 6.4% | 12.5% | 24.2% |
| **Denmark** | : | : | : | 6.3% | 12.3% | 13.5% | : | : | : |
| **Estonia** | 8.4% | 18.5% | 35.3% | 10.3% | 20.7% | 23.3% | 6.0% | 13.7% | 28.0% |
| **Finland*** | 8.8% | 19.0% | 38.0% | 9.6% | 19.4% | 19.1% | 6.0% | 11.4% | 25.5% |
| **France*** | 10.8% | 22.5% | 42.2% | 6.3% | 12.2% | 12.3% | 5.6% | 10.9% | 19.4% |
| **Germany** | 8.5% | 17.2% | 39.2% | 11.4% | 23.8% | 22.9% | 6.5% | 12.3% | 26.9% |
| **Greece** | 7.0% | 13.9% | 21.6% | 10.3% | 21.0% | 23.6% | 5.5% | 11.2% | 16.8% |
| **Hungary** | 10.0% | 20.7% | 36.8% | 11.8% | 22.2% | 22.4% | 7.7% | 14.7% | 26.4% |
| **Ireland** | 8.1% | 16.9% | 31.7% | 10.2% | 20.4% | 19.7% | 6.4% | 11.5% | 24.1% |
| **Israel*** | : | : | : | 10.5% | 22.2% | 21.1% | : | : | : |
| **Japan*** | 6.6% | 15.1% | 23.3% | 10.1% | 21.5% | 20.5% | 6.6% | 13.9% | 21.8% |
| **Korea*** | 9.3% | 23.1% | 33.5% | 10.2% | 22.8% | 20.5% | 6.5% | 12.9% | 21.5% |
| **Latvia** | 8.7% | 22.1% | 35.2% | 10.2% | 24.2% | 26.3% | 6.9% | 17.2% | 26.6% |
| **Lithuania** | : | : | : | 9.9% | 22.2% | 22.9% | : | : | : |

37 The concentration of economic resources is computed as the share of resources held by the average household at the top of the joint distribution.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Luxembourg** | 7.8% | 14.7% | 28.7% | 9.1% | 18.2% | 18.7% | 6.0% | 11.0% | 20.4% |
| **Malta**** | 7.2% | 14.6% | 31.7% | 8.4% | 15.6% | 17.7% | 5.2% | 9.9% | 17.7% |
| **Mexico*** | : | : | : | 13.7% | 37.6% | 34.8% | : | : | : |
| **Netherlands*** | 7.3% | 14.4% | 38.1% | 10.9% | 21.4% | 21.0% | 5.9% | 11.4% | 31.1% |
| **Norway*** | 7.6% | 17.1% | 34.5% | : | : | : | 7.7% | 17.7% | 35.1% |
| **Poland** | 7.8% | 15.3% | 24.7% | 11.1% | 20.8% | 22.5% | 6.2% | 12.0% | 19.2% |
| **Portugal** | 9.2% | 19.3% | 31.6% | 10.5% | 21.2% | 24.9% | 6.7% | 15.2% | 22.1% |
| **Romania**** | : | : | : | 14.9% | 26.5% | 26.8% | : | : | : |
| **Slovak Republic** | 7.2% | 12.7% | 24.1% | 10.6% | 17.2% | 18.3% | 8.5% | 12.1% | 22.6% |
| **Slovenia** | 8.3% | 14.2% | 30.8% | 7.6% | 14.0% | 15.3% | 5.2% | 9.4% | 15.1% |
| **Spain** | 10.9% | 23.1% | 55.1% | 10.0% | 20.9% | 20.8% | 7.7% | 15.4% | 35.1% |
| **Sweden** | : | : | : | 8.0% | 14.4% | 16.7% | : | : | : |
| **Switzerland*** | 7.7% | 17.0% | 38.2% | 10.3% | 19.8% | 17.8% | 6.5% | 11.6% | 27.0% |
| **United Kingdom*** | 7.6% | 19.6% | 36.8% | 12.2% | 29.6% | 23.6% | 10.8% | 21.8% | 46.4% |
| **United States*** | 12.7% | 46.4% | 76.2% | 11.6% | 41.3% | 25.7% | 8.8% | 19.8% | 57.4% |

Note: Data on the joint distribution of income and consumption for Norway are not presented in the table due to possible data issues. * denotes national or country-based estimates, ** denotes non-OECD countries.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

In most countries, households in the top 20% of the joint distribution of **income and consumption** own about the same share of the total of each of the two resources (Table 4.1). In general, 10% of households hold almost 22% of total income and 21% of total consumption. The smallest concentration of resources at the top of the joint distribution are in the Slovak Republic, where 11% of households hold 17% of total income and about 18% of total consumption. At the other end of the spectrum, 12% of households in the United States hold 26% of total income and 41% of total consumption.

The share of total wealth held by households belonging to the top 20 % of the joint distribution of **wealth and consumption** is twice as large as the share of total consumption (Table 4.1). In general, 7 % of households hold 13 % of total consumption and 26 % of total wealth. The smallest concentrations at the top of the joint distribution are observed in the Slovak Republic, Belgium and Poland, while (at the other extreme) 9 % of households in the United States hold one quarter of total consumption and more than half of total wealth.

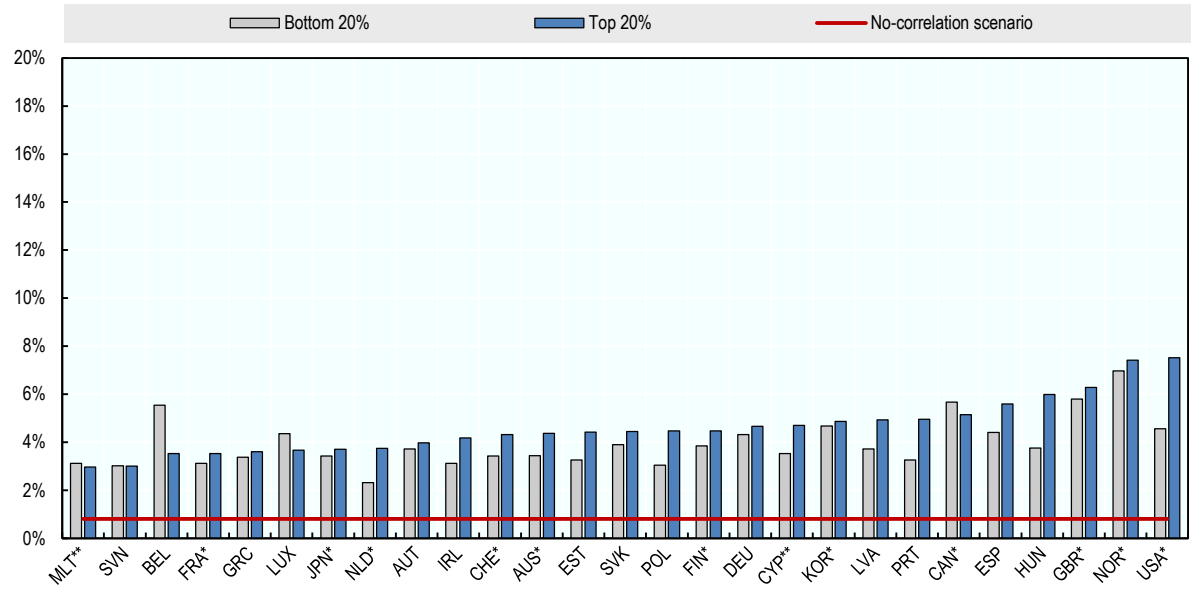
### *4.1.2. Inequality in three dimensions*

*On average, 5 % of households belong to the top quintile of all three dimensions*

Figure 4.4 shows the share of households that belong to the bottom and top quintiles of all three (in other words, **income, consumption and wealth**) distributions. If the association were perfect, then all the households in a given quintile of the income distribution would belong to the same quintile of the wealth and consumption distributions. On the contrary, if the association were nihil, then a household in a given quintile of the income distribution would be equally likely to belong to any of the quintiles of the wealth and consumption distributions. The horizontal line set at 0.8% in Figure 4.4 represents the case of no correlation between households' positions in the joint distribution of income, consumption and wealth, while the closer bars are to 20% (the perfect correlation scenario) the stronger is the association between the ranking of households in the distribution of income and their ranking in the

distributions of wealth and consumption[38]. On average, 5% of households belong to the top 20% of all three dimensions, in other words, a quarter of what would be observed if the three distributions overlapped perfectly, but over six times higher than the outcome in a random scenario. The three distributions tend to overlap more strongly at the top than at the bottom. On average (across all countries), only 4% of all households belong to the bottom 20% of each distribution. However, this is not the case everywhere: in Belgium, and to a lesser extent in Luxembourg and Canada, the overlap is stronger for the first quintile (Figure 4.4).

### Figure 4.4. Share of households at different points of the joint distribution of income, consumption, and wealth, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of households belonging to the top 20% of the each of the three distributions of income, consumption and wealth. For instance, 8% of United States' households belong simultaneously to the top quintile of the distribution of income, consumption and wealth. The horizontal line set at 0.8% represents the case of no correlation between a household's quintile in the distribution of income and its quintile in the distributions of consumption and wealth; while the closer the bars are to 20% (the perfect correlation scenario) the stronger is the association between a household's quintile in the distribution of income and its quintile in the distributions of consumption and wealth.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data codes: icw_res_01 and icw_res_01).

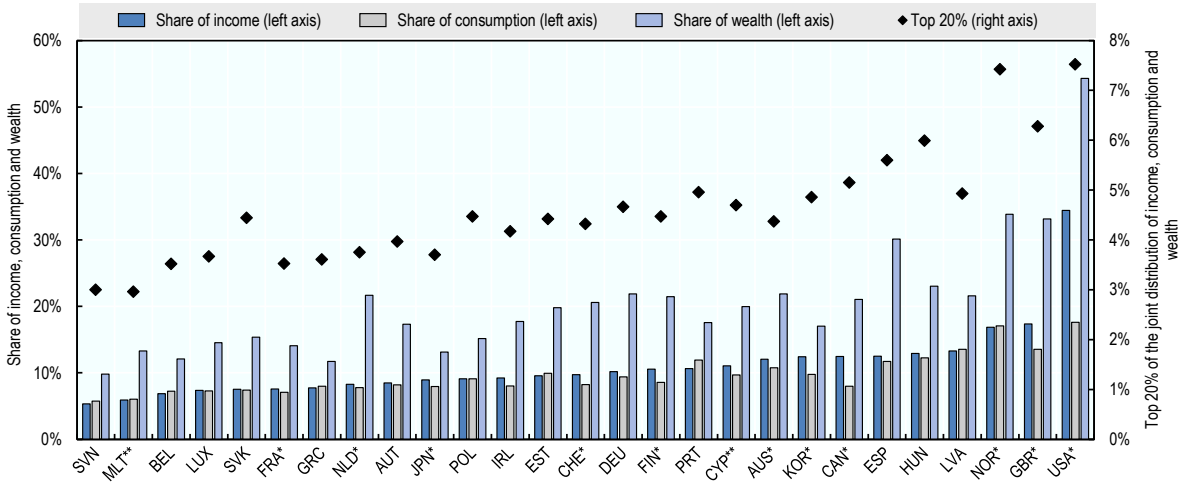
### *Wealth tends to be about twice as concentrated as income and consumption in the top quintile of the joint distribution of the three dimensions*

On average, the 5% of households in the top 20% of income, consumption, and wealth (Figure 4.5, right vertical axis) contribute to 10% of total consumption and hold 11% of income and 20% of wealth (left vertical axis). The share of wealth held by households in the top 20% of the joint distribution of

[38] There are 125 possible combinations of quintiles characterising the joint distribution of income, consumption and wealth, in other words, 5 (quintiles) to the power of 3 (distributions). This implies that, if there were zero correlation between ICW, a household in a given quintile of the income distribution would be equally likely to belong to any of the 125 possible combinations, which would then hold 0.8% of households each (in other words, a 1/125 share).

income, consumption and wealth is by far the largest in the United States (over 50%), followed by Norway and the United Kingdom. The share of income and consumption held by the same households varies less among countries, and in a majority of countries it remains below or around 10%. However, the share of income held by households at the top of the ICW distribution in the United States (34%) stands out; the same share is only 17% in both Norway and the United Kingdom, the second highest countries by this measure.

### Figure 4.5. Share of income, consumption and wealth held by households in the top 20% of the joint distribution of income, consumption and wealth, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of households belonging to the top 20% of the joint distribution of income, consumption and wealth. For instance, in Slovenia, 3% of households are in the top 20% of income, consumption and wealth holding 5% of the total income, 6% of the total consumption and 10% of the total wealth of all households in the country.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

## 4.2. Who is most at risk of poverty in more than one dimension?

The analysis in the previous section confirmed the intuition that income, consumption and wealth are correlated. This means that higher-income households tend to accumulate more wealth, and this may be used as a buffer to smooth their typically higher living standards in case of income shocks. Conversely, the correlation of income, wealth and consumption in the bottom of the distribution entails that the most disadvantaged households have little financial means to draw on in case of need.

This subsection discusses the prevalence of two-dimensional risk of poverty – in other words, the share of individuals (rather than households, as was the case in the previous section) who are poor in two dimensions (for example, in terms of *both* income *and* consumption) – and its relation with standard, uni-dimensional risk-of-poverty measures. The interplay between different risk of poverty measures may provide valuable information to policy makers interested in identifying a target group of individuals with more urgent economic needs. The section also examines how the demography of two-dimensional risk of poverty (defined along the lines of age and education level (of the individual, rather than of the household head), family type, activity status and housing tenure) depends on the resource measure used, in other words, income, consumption or wealth.

As described in more detail below, for some population groups all measures show the same picture: unemployed people, renters and people with at most a lower secondary level of education systematically face a higher risk of poverty, regardless of the dimension considered. By contrast, while

people below 25 years and people aged 65 years and over are both worse off than adults between 25 and 64 years in income terms, only young people disproportionately belong to the lower reaches of the wealth and consumption distributions.

As mentioned above, in contrast to the evidence on multi-dimensional inequality portrayed in the previous subsection, the focus here is on individuals rather than households. To account for economies of scale in the use of economic resources within the household, as explained in Box 2.2, two sets of ICW experimental estimates were derived, based on two different hypotheses on the size of households' 'economies of scale': either the square root of household size used in OECD statistics on income inequality and poverty, or the 'OECD modified scale' used by Eurostat[39]. For the sake of brevity, this section discusses only results derived using the OECD modified scale[40]. In the future, sensitivity analysis could be conducted to highlight possible differences in key patterns due to the use of alternative scales[41].

An analysis of the relative risk of poverty also requires the choice of a risk-of-poverty threshold, which in the case of Eurostat official estimates is set at 60% of median income, while it is set at 50% of median income for OECD reporting. In the following, the analysis refers to the threshold set at 60% of median levels. If the 50% of median threshold were chosen, lower risk-of-poverty rates would emerge, although similar patterns as those reviewed below are likely be found. Box 4.1 defines the different concepts used in the analysis.

### Box 4.1. Defining risk of poverty in a multi-dimensional setting

While different concepts can be used to define risk of poverty (in terms of income, consumption and wealth), the analysis in this section is based on the following:

**Income risk-of-poverty threshold** – set at 60% of the median household equivalised income of the total population. The **income poor** are people whose equivalised household disposable income falls below this risk-of-poverty threshold.

**Consumption risk-of-poverty threshold** – set at 60% of the median household equivalised consumption of the total population. The **consumption poor** are people (in different demographic groups) whose equivalised household consumption falls below this risk-of-poverty threshold.

**Asset risk-of-poverty threshold** – defined relative to either income or consumption. The asset risk-of-poverty threshold is set at 25% of the annual income (consumption) risk-of-poverty threshold. As such, the **asset poor** are people who lack sufficient *liquid financial assets* to maintain a risk-of-poverty-level living standard for at least three months. Illiquid assets are excluded from the definition of asset

39 Using household size as the determinant, equivalence scales can be expressed through an 'equivalence elasticity', in other words the power by which economic needs change with household size. The equivalence elasticity can range from 0 (when unadjusted household measures are considered) to 1 (when a per person approach is used). The smaller the value for this elasticity, the higher the economies of scale in the use of economic resources. The 'modified OECD scale', with an equivalence elasticity of 0.53, assumes slightly higher economies of scale relative to the square root scale (0.5).

40 Note that for Korea and Japan, due to the lack of information on household composition, estimates in this section are adjusted by the square root of the household size.

41 Different equivalence scales can have a strong impact on the measured level of risk-of-poverty, particularly when decomposed by different family types. Single adult households with children and households with a larger number of adults are more sensitive to these changes. However, previous analysis suggests that the change in the equivalence scale does not affect country ranking as such Buhmann et al. (1988[36]).

risk of poverty used in this paper because they cannot easily be converted into cash and are less useful than savings and other liquid assets in times of sudden financial stress.

Based on the concepts above, it is possible to define the following groups of people.

**Income and asset poor** – defined as the share of people who are both income and asset poor.

**Income and consumption poor** – defined as the share of people who are both income and consumption poor.

**Asset and consumption poor** – defined as the share of people who are both asset and consumption poor.

### *4.2.1. Asset risk of poverty is much more widespread than income risk of poverty*

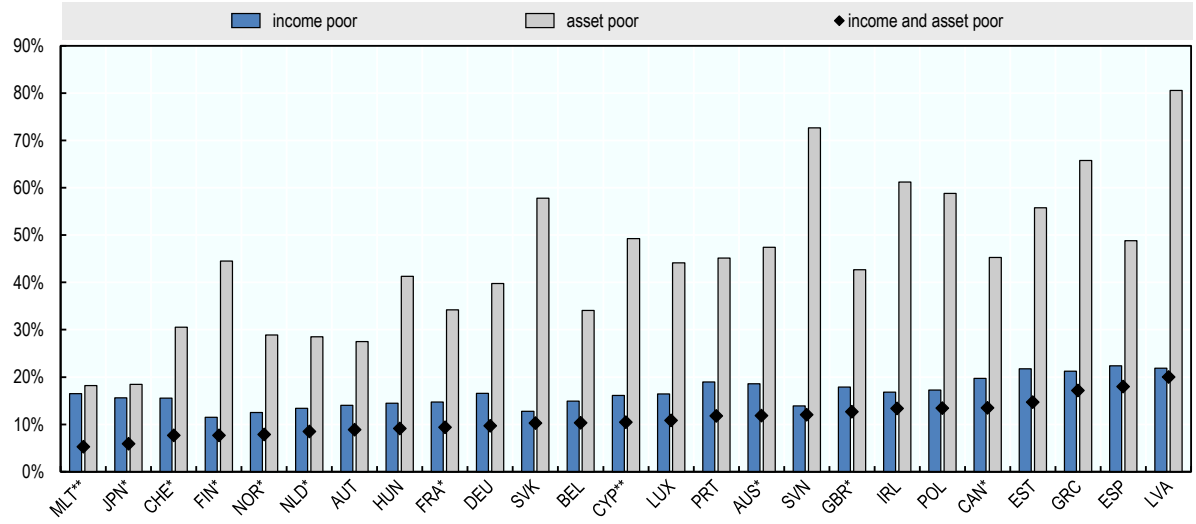
The proportion of people who are **asset poor (defined relative to income levels)** is substantially higher than the income risk-of-poverty rate in almost all countries (Figure 4.6), reflecting that in most EU Member States and OECD countries, a significant share of people live in households with little or no positive net wealth[42]. On average, 17% of people are income poor, while 45% are asset poor, in other words, they lack sufficient liquid financial assets to keep them above the income risk-of-poverty threshold for at least three months[43]. The scope of the problem varies widely across countries. In eastern and Baltic EU Member States as well as in Greece and Ireland, asset risk of poverty is widespread. In Latvia, 8 in 10 individuals lack the liquid financial wealth needed to support them at the level of the income risk-of-poverty threshold for at least three months. Over two-thirds of the population are also asset poor in Slovenia and Greece. By contrast, only 18% of individuals are asset poor in Japan and Malta[44].

42 Due to a number of methodological differences, previous OECD estimates of asset risk-of-poverty, based on the OECD Wealth Distribution Database (WDD, oe.cd/wealth), are only partially comparable with those shown here. In particular, previous OECD estimates referred to the 50% threshold and squared-root elasticity that are typically used in OECD analysis. Moreover, for a number of countries covered in the WDD, lack of information on disposable income on household wealth surveys implied that the risk-of-poverty line underlying the concept of asset risk of poverty was based on gross income. See Balestra et al. (forthcoming[33]) for a comparable analysis of income and asset risk of poverty in a sub-set of OECD countries.

43 Asset-based risk of poverty is often construed as a proxy for financial insecurity (see, for instance, Balestra and Tonkin (2018[18])), as it signals the lack of financial assets that could support living standards in the event of a sudden loss of income. However, it should be kept in mind that this measure provides only a partial view of financial insecurity, as it does not take into account social transfers (for example, unemployment benefits) that people may receive in the event of some types of shocks – depending on their individual circumstances. A broader approach to economic insecurity is provided by Hacker (2018[34]).

44 Country rankings are likely to reflect, at least partly, cross-country differences in outright homeownership rates. In countries where outright homeownership is widespread, such as Canada, Baltic and Eastern-European countries, liquid financial assets make up a smaller share of net wealth for those of the bottom of the distribution. The opposite applies in Austria and Norway, where homeownership rates are comparatively lower.

## Figure 4.6. Share of individuals who are income poor, asset poor, and both income and asset poor, total population, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of individuals who are both income and asset poor. Data for Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_10).

The primary value of the joint income and wealth distribution rests on the analysis of the extent to which people are both **income and asset poor**. While not all income poor are also asset poor, people belonging to both these groups are obviously worse off than those who are only income poor, as they have limited liquid assets that might be used to alleviate the impact of their low income[45]. In Latvia, one in five individuals are both income and asset poor, while around 6% of individuals in Japan and Malta are in this group. This partly reflects the large proportion of people at relative income risk of poverty also having few liquid financial assets in countries with a higher incidence of income and asset risk of poverty: in Latvia, 92% of people who are income poor are also asset poor. By contrast, most income-poor people in Malta and Japan are not asset poor, meaning that they have sufficient liquid financial assets to supplement their income and support their material living standards, at least for a short spell.

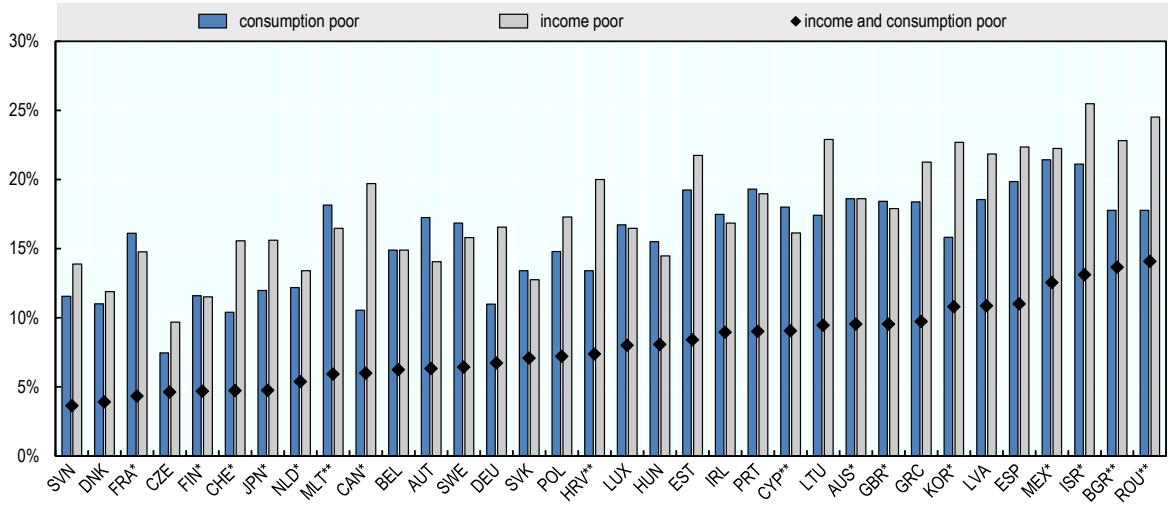
### *4.2.2. Income risk of poverty tends to be higher than consumption risk of poverty*

In general, **consumption risk of poverty** is lower than (or about as high as) income risk of poverty. In a small number of countries, however, the consumption risk-of-poverty rate is higher (Figure 4.7) than the income risk-of-poverty rate based on a similar threshold. The degree of overlap between **income and consumption risk of poverty** is far from perfect and varies substantially across the countries examined. On the one hand, this suggests that many income poor engage in consumption smoothing and are able to maintain their standard of living above the consumption risk-of-poverty threshold. In some cases, such behaviour may be driven by knowledge or expectation that household income will increase in the near future, for example, for people starting a new job soon or for students. However, many households of this type will remain vulnerable to the risk of poverty as the resources they are

45 Social insurance provides a first line of defence against negative income shocks due to job loss, sickness, childbirth, etc. Hence, wealth as a buffer against income shocks is needed primarily in countries where social insurance is less generous or for groups who are not well covered.

relying on are finite and the situation cannot continue indefinitely. On the other hand, consumption risk of poverty in the absence of income risk of poverty could be an indication of perceived uncertainty over future income levels and a lack of adequate assets that could be used to maintain living standards if income were to drop (Serafino and Tonkin, 2017[26]).

### Figure 4.7. Share of individuals who are income poor, consumption poor, and both income and consumption poor, total population, 2015 or closest available year

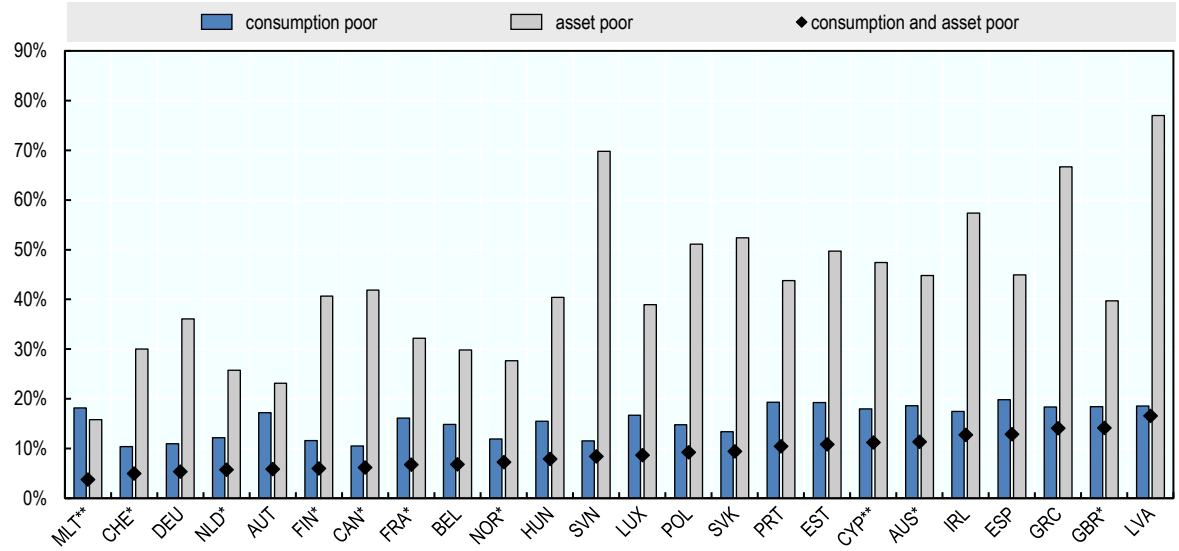



Note: Countries are ranked in ascending order the share of individuals who are both income and consumption poor. Data for the United States are not available. Data for Norway are not shown due to possible data issues.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_10).

### *4.2.3. On average, 1 in 10 individuals are both consumption and asset poor*

The share of people who are consumption poor is substantially lower than the **asset poor (defined relative to consumption levels)**, in other words, people who lack sufficient liquid financial assets to maintain their consumption level above the risk-of-poverty threshold for at least three months. Moreover, asset risk-of-poverty rates are slightly lower when defined in terms of consumption rather than income (43% versus 45% on average), as median consumption is lower than median income. However, the country ranking is broadly preserved under either measure. Asset risk of poverty is highest in Latvia and lowest in Malta, regardless of the definition applied. On average, 1 in 10 individuals are both **consumption and asset poor**, even though the degree of overlap between consumption and asset risk of poverty varies greatly across the countries with available information (Figure 4.8).

**Figure 4.8. Share of individuals who are asset poor, consumption poor, and both consumption and asset poor, total population, 2015 or closest available year**



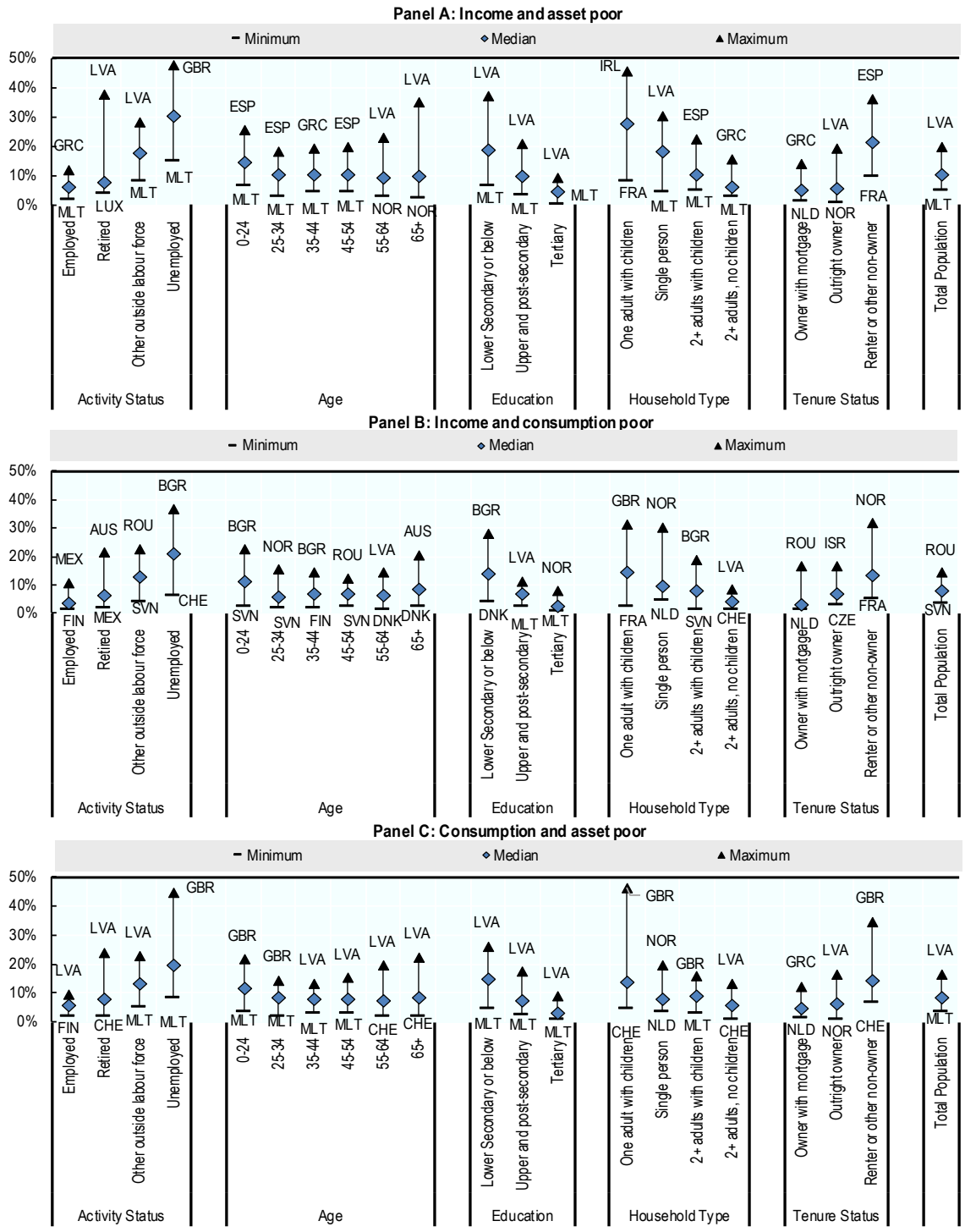

Note: Countries are ranked in ascending order of the share of individuals who are both consumption and asset poor. Data for Japan, Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code:icw_pov_10).

### *4.2.4. Unemployed people, renters and people with at most a lower secondary level of educational attainment face a higher risk of poverty in more than one dimension*

While the analysis above suggests that the prevalence of two-dimensional risk of poverty is relatively limited, it affects people who would most benefit from public support to meet their needs. Thus, the next question to ask is 'who exactly are the worst off?' In the following, this paper will explore the demographic composition of people most at risk of poverty in more than one dimension.

Figure 4.9 provides a high-level overview of how likely different population groups are to be poor in more than one dimension across the countries with available information. In each panel and for each population group, end points represent countries with the smallest and largest share of people who are poor in any two given dimensions, while squares represent median levels across all countries with available information. For instance, the share of income and asset poor (Panel A) among employed people is smallest in Malta and largest in Greece. Country-level information is provided in Annex B.

## Figure 4.9. Share of individuals who are income and asset poor (Panel A), income and consumption poor (Panel B), and consumption and asset poor (Panel C) by socio-demographic characteristics, selected countries, 2015 or closest available year




Note: In each panel and for each population group, end points represent countries with the smallest and largest share of people who are poor in any two given dimensions, while squares represent median levels across all countries with available information.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data codes: icw_pov_10, icw_pov_11and icw_pov_12).

A number of general patterns emerge from this analysis. First, the share of people who are both **income and asset poor** tends to decrease with *age* (Figure 4.9, Panel A). There are, however, some notable exceptions to this general pattern: for instance, in Cyprus and Malta the proportion of people who are income and asset poor remains somewhat steady across age groups. Similarly, while in Australia the age profile of income and asset risk of poverty is U-shaped, in both Slovenia and Latvia it is highest among people aged 65 years and over (Figure B.8). Income and asset risk of poverty generally follow different age patterns. While in most countries the age profile of income risk of poverty is U-shaped, asset risk of poverty is typically highest for the youngest age group and falls steadily in older ages as people usually accumulate wealth over the course of their life, implying that older people are more likely to have accumulated significant assets than their younger counterparts. This result illustrates how the picture of who is poor by age changes dramatically when considering assets in addition to income[46].

Second, the share of people who are both **income and consumption poor** tends to mimic a U-shaped (or upward) age pattern (Figure 4.9, Panel B). This finding has been confirmed in other studies, which show a reduction in consumption expenditure following the income drop associated with retirement (see, for instance, Fisher et al. (2009[27])). In a significant minority of countries, however, the risk of being both income and consumption poor decreases steadily with age: this is the case, for instance, in many European countries (Figure B.9). Conversely, the share of people who are both **consumption and asset poor** tends to decrease with age (Figure 4.9, Panel C), although again with some notable exceptions: in particular, in Slovenia, Estonia and Latvia the prevalence of asset and consumption risk of poverty is highest among people aged 65 years and over (see Figure B.10).

Third, in terms of *household structure*, in nearly all countries the share of people who are both **income and asset poor** is highest among single adult households with children and lowest among households composed of two or more adults without children (Figure 4.9, Panel A). The second highest risk of income and asset risk of poverty is usually observed among single-person households (Figure B.11). On average, **income and consumption risk of poverty** is highest for people in households of single adults with children and lowest for people in households of two or more adults without children (Figure 4.9, Panel B). However, in a number of European countries, single-person households have the highest income and consumption risk-of-poverty rate among all household types (Figure B.12). Finally, in virtually all countries the share of people who are both **asset and consumption poor** tends to be highest among people in households composed of single adults with children and lowest among people in households composed of two or more adults without children (Figure 4.9, Panel C). The risk of facing consumption and asset poverty is more equally distributed across household types in Poland and Latvia (Figure B.13).

Fourth, people having completed at most a primary or lower secondary level of *education* are significantly more likely to be **income and asset poor** compared with people with higher levels of educational attainment (Figure 4.9, Panel A and Figure B.14). A similar pattern emerges when **income and consumption risk of poverty** (Figure 4.9, Panel B), and **consumption and asset risk of poverty** (Figure 4.9, Panel C) are considered[47].

---

46 The case of Switzerland best exemplifies this pattern: in this country, income risk-of-poverty rates tend to be higher in older age due to exclusion from income of lump-sum transfers that are common in Swiss retirement schemes. These transfers, however, increase the liquid financial assets of older people. Hence, extending the analysis to asset risk of poverty shows high income risk-of-poverty rates coupled with low asset risk-of-poverty rates, which means that older Swiss people are able to sustain their standard of living despite high income risk-of-poverty rates.

47 Cross-country comparisons by educational level are presented in Figure B.14-Figure B.16 but are not discussed in detail here because they do not highlight any particular divergence from the general pattern that is worth mentioning.

Fifth, as for *activity status*, unemployed people are systematically more exposed to the risk of multi-dimensional poverty compared with the rest of the population (Figure 4.9). The picture is more mixed among retired people. In many countries they face a lower risk of **income and asset poverty** than unemployed people, but higher than employed people. However, in a few cases the retired are even better off than employed people. In Baltic Member States, on the contrary, retired people are very likely to experience low income and assets (Figure B.17). Relatively high levels of **income and consumption risk of poverty** (over 10%) among retired people are observed in Australia, Baltic Member States, Israel and the United Kingdom (Figure B.18); similarly, in Slovenia and Estonia the assets accumulated over a lifetime do not seem to adequately protect retired people from a higher risk of **consumption and asset poverty** (Figure B.19).

Finally, concerning *tenure status* **income and asset risk of poverty** tends to be highest among renters (Figure 4.9, Panel A). Since the definition of asset risk of poverty used in this section relies only on liquid financial wealth and excludes real estate liabilities, owners with a mortgage are not necessarily more likely than the average for the total population to be both income and asset poor: an explanation could be that they need to save and accumulate liquid assets in order to repay their debt. Renters, on the other hand, are more likely to be income and asset poor (Figure B.20). Similar patterns emerge in the case of **income and consumption risk of poverty** (Figure 4.9, Panel B), and **consumption and asset risk of poverty** (Figure 4.9, Panel C)[48]. This evidence suggests that living in an owner-occupied home can provide resilience from day to day, because of lower housing costs (see Subsection 4.3).

## 4.3. Saving rates, essential expenditures and discretionary income

Saving rates, essential expenditures and discretionary income are other experimental metrics that may lead to a better understanding of how households sustain their standard of living. Figure 4.10 shows the proportion of **dissaving households**, in other words those households whose consumption expenditure exceeds their current income[49]. In general, almost one in three households spends more than it earns.

### *4.3.1. A significant share of households spend more than they earn*

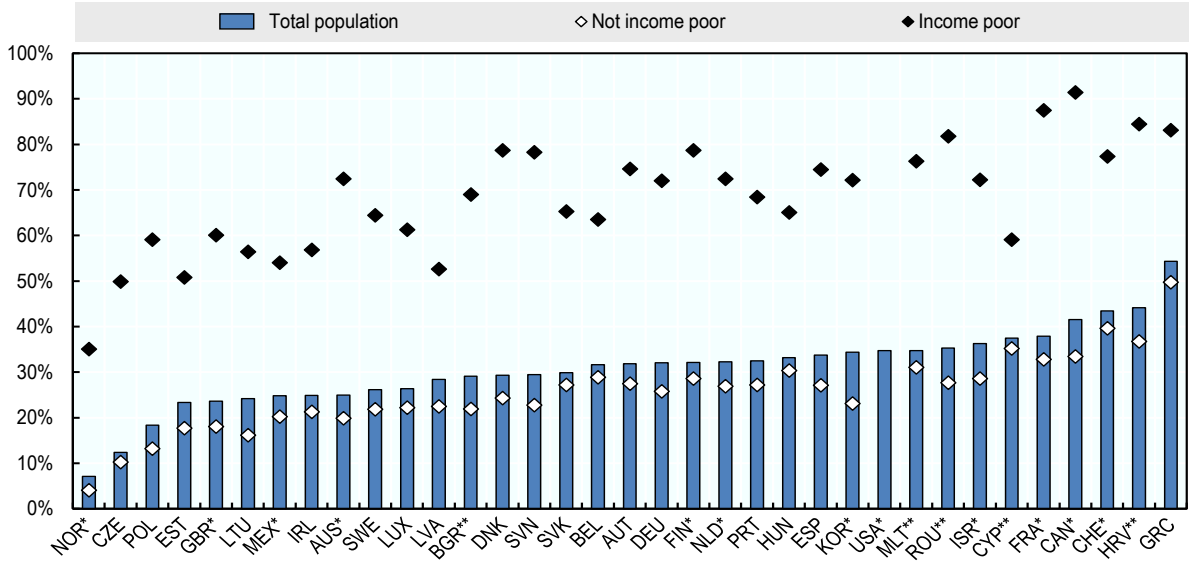
Several factors may explain why households report consumption expenditure exceeding their income. A possibility is that these households may experience a temporary spell of low income that is managed by supporting their expenditure through other means, such as benefiting from one-off government transfers, relying on family or friends through non-recurrent (in other words, capital) transfers, drawing down on their savings or incurring debt. However, the evidence of two-dimensional risk of poverty reviewed in the previous subsection suggests that a large proportion of people have little liquid assets to draw on in case of need, which leaves them more exposed to economic shocks. Another factor relates to measurement error; reporting of consumption expenditures requires the use of diaries, which imposes a higher reporting requirement on respondents than in the case of income; in addition, income is more likely than expenditure to be underreported at the bottom of the distribution (Brewer, Etheridge and O'Dea, 2017[28]).

---

48 Cross-country comparisons by tenure status are presented in Figure B.20-Figure B.22 but are not discussed in detail here because they do not highlight any particular divergence from the general pattern that is worth mentioning.

49 According to the definition of savings outlined in Box 2.2, a household is overspending or dissaving if its total current expenditure exceeds its gross income by more than 1%. However, Eurostat and the United States diverged from this concept by defining savings as the difference between annual net disposable income and total consumption expenditure (net of imputed rents).

## Figure 4.10. Share of dissaving households by income risk-of-poverty status, total population, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of dissaving households in the total population. Data are not available for Japan. Information by risk-of-poverty status is not available for the United States. Possible data issues with Norwegian data.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_06).

Not surprisingly, dissaving households are concentrated at the bottom of the *income* distribution. On average, almost 70% of income-poor households have consumption levels that exceed their income, a finding consistent with the limited degree of overlap between income and consumption risk of poverty reviewed in the previous subsection (Figure 4.10). Households whose consumption expenditure is lower than their income will be more able to withstand income shocks without drastic changes in their living standard, simply by saving less and/or using their accumulated savings, although the leeway of the income poor is limited as a sizeable share of their income goes to meeting basic needs (see discussion on essential expenditures below).

### *4.3.2. There are large differences in median saving rates of the bottom and top income quintiles*

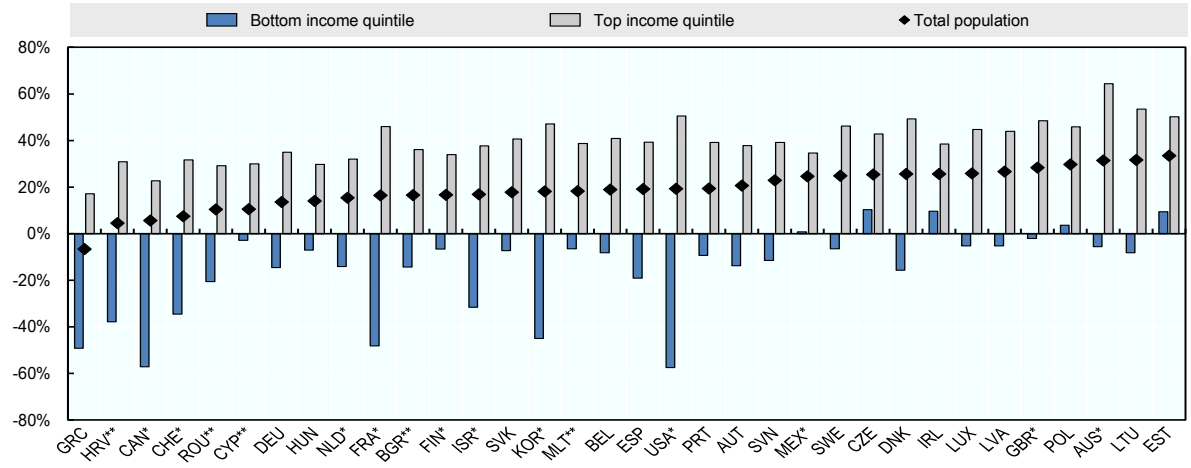
As with shares of dissaving households, there are notable differences in **saving rates** across countries. Figure 4.11 shows the median saving rates for the total population as well as for households in the bottom and top *income quintiles*[50]. The median saving rate for the total population ranges from -7% in Greece to 30% and above in Poland, Australia, Lithuania and Estonia. Country rankings partly reflect institutional factors, such as differences in social security systems and home-ownership rates. For instance, Denmark, where renting is the norm and the welfare state is generous, features lower saving

50 Data are presented for medians, rather than averages, since the median is robust to extreme values that might be observed in saving rates.

rates than Poland, which is characterised by a higher share of households owning their home outright and comparatively lower levels of social spending[51].

In the majority of countries, the median saving rate of households in the first income quintile is negative, with some saving rates being close to or even below -50%. In Greece even the median saving rate of the total population is negative. Conversely, the median saving rate in the top quintile of income is usually quite high, in particular in Australia (64%) and Lithuania (54%). The income gradient is strongest in the United States, and weakest in Ireland[52].

## Figure 4.11. Median saving rates by income quintile, 2015 or closest available year




Note: Countries are ranked in ascending order of median saving rates for the total population. Data for Japan and Norway are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_sr_03).

The conclusion that households with high *current* income save more than those with low *current* income is well established in the literature, although the steepness of the income gradient varies across studies. However, there is more disagreement about whether households with high *permanent or lifetime* income do indeed save more (see for instance Bozio et al. (2017[29]) and Garbinti and Lamarche (2014[30]) for reviews on the topic). Indeed, there are limited comparative data on how savings rates vary along the permanent income ladder. The results shown in Figure 4.11 are broadly in line with national studies[53].

The estimates for how much households save should not be taken as overly precise, given the experimental nature of these statistics. Furthermore, differences in data quality underlying these

51 The definition of savings used here does not include contributions to compulsory pension schemes, which might play a role for some countries, such as the Netherlands.

52 If households were ranked in terms of consumption rather than income, savings rates would not be overly negative, reflecting the fact that income is more likely than consumption to be under-reported at the bottom of the distribution.

53 For instance, Garbinti and Lamarche (2014[20]) show similar patterns for France based on data from the *Enquête Patrimoine* for 2009. The lower median rates reported by the authors may reflect both the different wave of the survey used and the bottom-coding approach, which led them to exclude from the analysis negative or very low incomes.

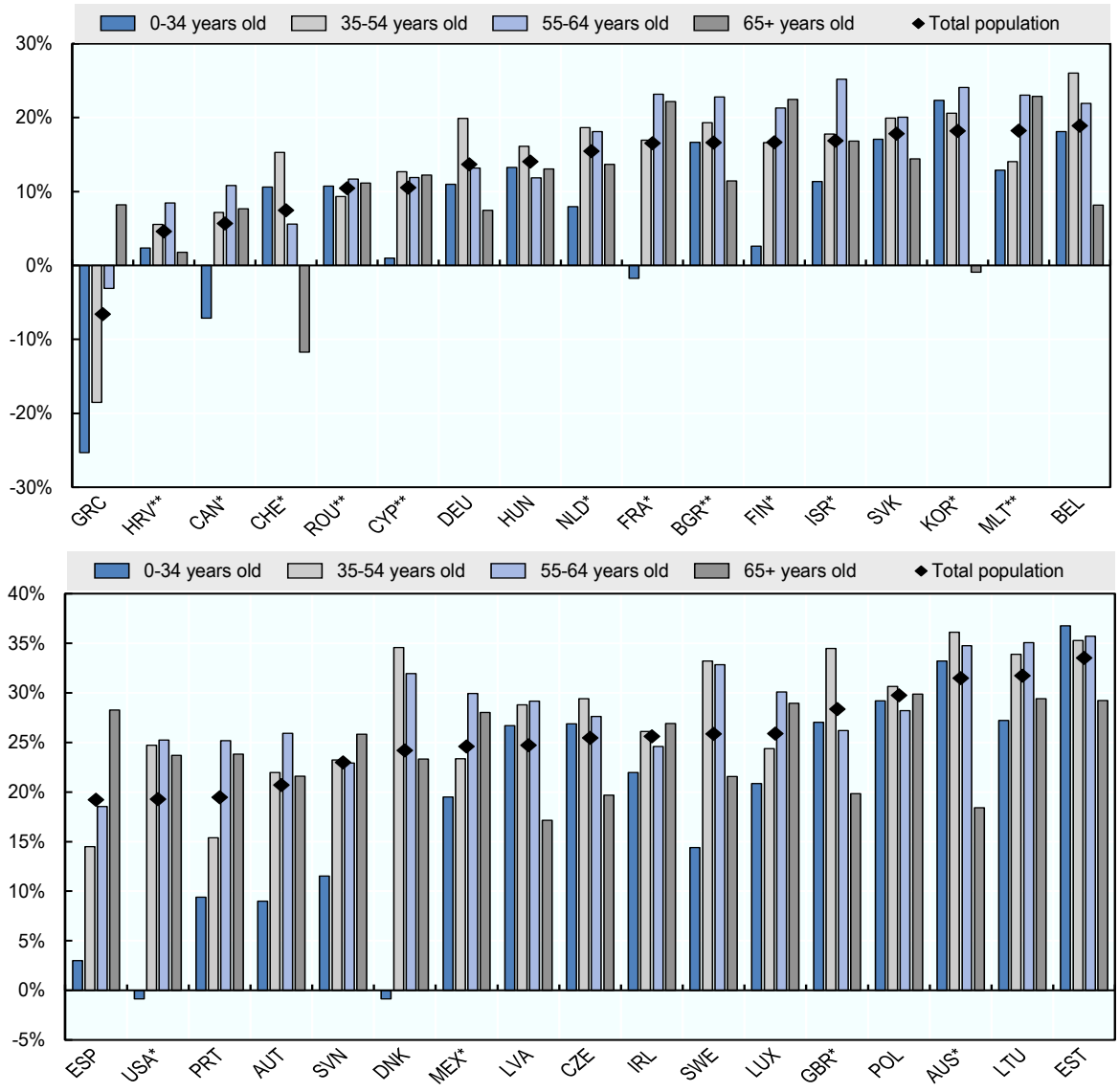
estimates may result in imprecise estimates. For instance, estimates of saving rates for low-income households (in other words, the first quintile) should be taken with caution, as income is more likely than expenditure to be underreported at the bottom of the distribution. Since household saving is the difference of two large aggregates (income and total consumption), in general – and especially at the bottom of the distribution – even a small inaccuracy in one of those components leads to a substantial error in the balance item. The definitions of income and consumption underlying the estimates are also likely to affect their reliability and cross-county comparability. Despite the need for caution, it is still a cause for concern to see that budgets are so tight for many low-income households. The estimates show that in a number of countries, including Canada, Switzerland and the United States, even the median household in the second income quintile records a negative saving rate. Dissaving not only threatens a household's financial resilience by limiting flexibility in response to a shock, but also prevents asset accumulation.

### *4.3.3. Saving decisions change over the course of a lifetime*

Individuals may change their saving behaviour as they *age*, since they tend to adjust their consumption in order to smooth out income variations. In particular, younger and older people would be expected to have lower saving rates, since they are generally earning less than the average over their lifetime. However, the evidence at hand is not conclusive in this regard since there is considerable variation from country to country. Households with younger heads (younger than 35 years) most often save less than households whose main earner is older. However, the expected decrease of saving rates after retirement (for example, age group 65 years and over) is observed in only a small majority of the countries. In a number of European countries (for example, Greece, Slovenia and Spain), older people have higher saving rates than their younger counterparts (Figure 4.12). This finding suggests that people may reduce their consumption as a response to lower income after retirement in order to pass their assets on to their descendants or as a precaution against potential future higher expenditures (for example, due to the costs of long-term care)[54].

[54] In Australia and Switzerland, the negative or comparatively low median saving rates for those aged 65 years and over are likely to reflect the exclusion from the income definition of lump-sum transfers that are frequent in national retirement schemes.

## Figure 4.12. Median saving rates by age group, 2015 or closest available year



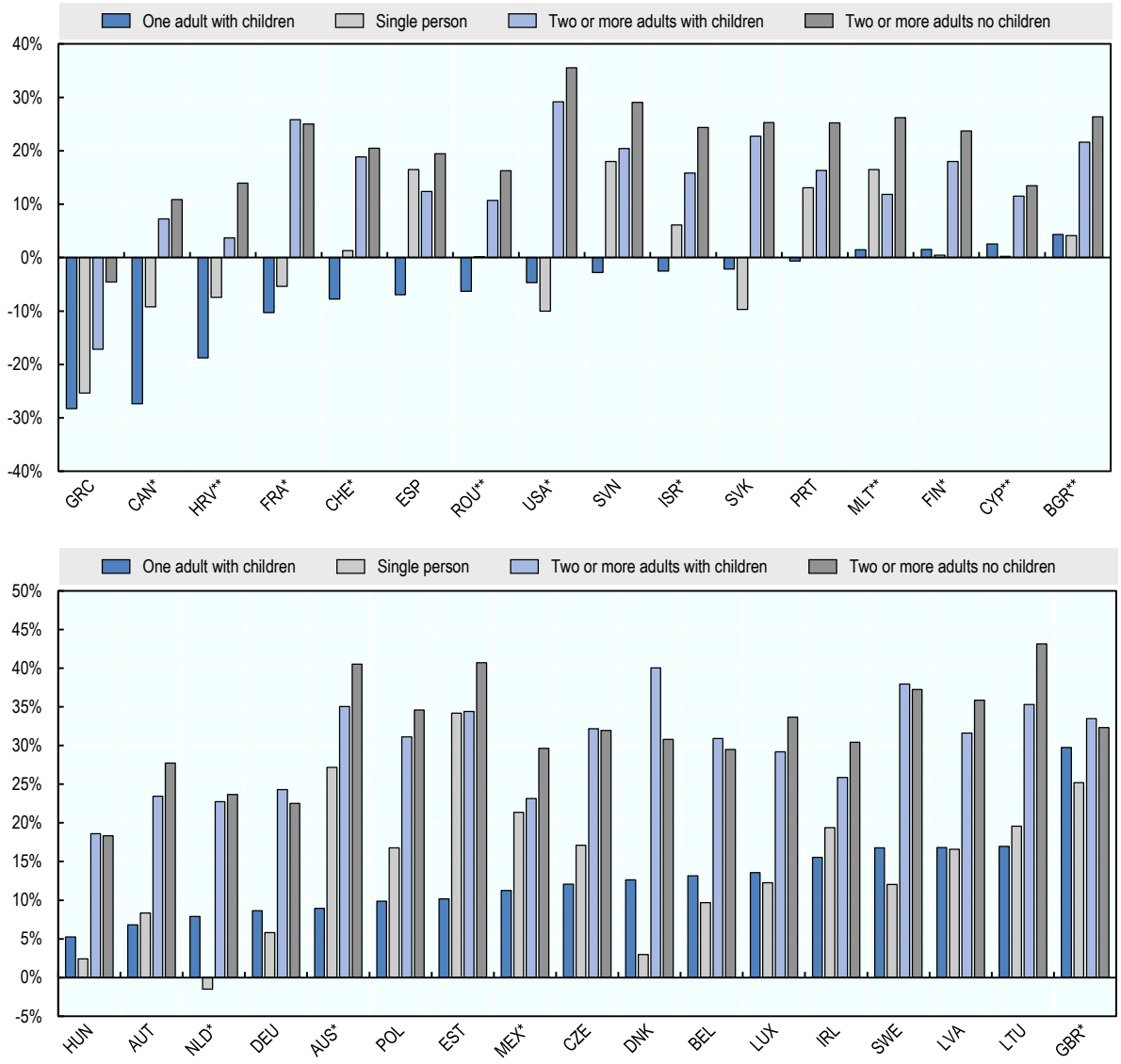

Note: Countries are ranked in ascending order of median saving rates for the total population. Age groups 35–44 and 45–54 years have been combined and the weighted average (35–54 years) is shown. Data for Japan and Norway are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code:icw_sr_01).

### *4.3.4. Households with two or more adults tend to save more than single-adult households (with or without children)*

Saving decisions are also closely linked to *household structure* and *tenure status*. In all countries, households with a smaller number of adults feature lower saving rates, with households composed of a single adult with children often being the only household type displaying negative median saving rates – a finding that is driven by this group's low income (Figure 4.13). Not surprisingly, saving rates are also lower among renters, while in a large number of countries owners with a mortgage exhibit the highest saving rates (Figure B.23). This finding confirms the evidence reviewed earlier that an outstanding

mortgage is not necessarily a strong predictor of income and asset risk of poverty[55]. Unsurprisingly, median saving rates by *education level* in Figure B.24 show higher savings among households headed by individuals with a tertiary level of education.

## Figure 4.13. Median saving rates by household type, working-age households, 2015 or closest available year




Note: Countries are ranked in ascending order of median saving rates for a single adult with children. Data for Japan, Korea and Norway are not available. Working-age households are defined as those with a head aged 16–64 years.

55 Using detailed information on regular saving flows allocated to each type of asset, as well as on discretionary savings and savings motives available in the Panel on Household Finances, Le Blanc and Schmidt (2017[35]) support the evidence that German households that own their main residence do save more than their renter counterparts. In particular, owner households do not substitute other savings with mortgage repayments but save on top.

Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_sr_02).

Beyond the level of spending in relation to income, it is also important to consider on what households spend their money. Essential spending, such as housing, utilities and food, is more difficult to adjust in the event of a temporary income shock, whereas it is easier to reduce spending on non-essential items. When households spend a large majority of their income on essentials, they have less room for changes in consumption to take the strain in the face of an income loss.

### *4.3.5. Essential expenditures represent a larger share of disposable income for households at the bottom of the distribution than for those at the top*

Figure 4.14 presents the percentage of disposable income spent on **essential and pre-committed expenditures,** both on average and across *income quintiles*[56]**.** Across the countries with available information, these expenditures make up almost 40% of household disposable income on average. This share is lowest in the United Kingdom and Luxembourg (below 30%) and highest in Greece, where households spend on average around 60% of their disposable income on essentials. Unsurprisingly, lower-income households in all countries spend a higher share of their income on essentials, which limits their capacity to respond to shocks by adjusting their spending. While essential expenses represent, in general, around one quarter of the disposable income of the richest 20% of households, this share goes up to almost 80% among the poorest 20%, and it is even higher in some countries (Figure 4.14). In Greece, Romania and Croatia, on average, essential expenditures exceed disposable income among low-income households. The steepest income gradients are observed in Canada, Germany and Lithuania, where the share of income devoted to essential expenditure is about four times as large for households in the bottom quintile as for those in the top quintile. This difference is mainly driven by large inequalities in housing costs (specifically, low housing costs towards the top of the distribution related to higher home-ownership rates).

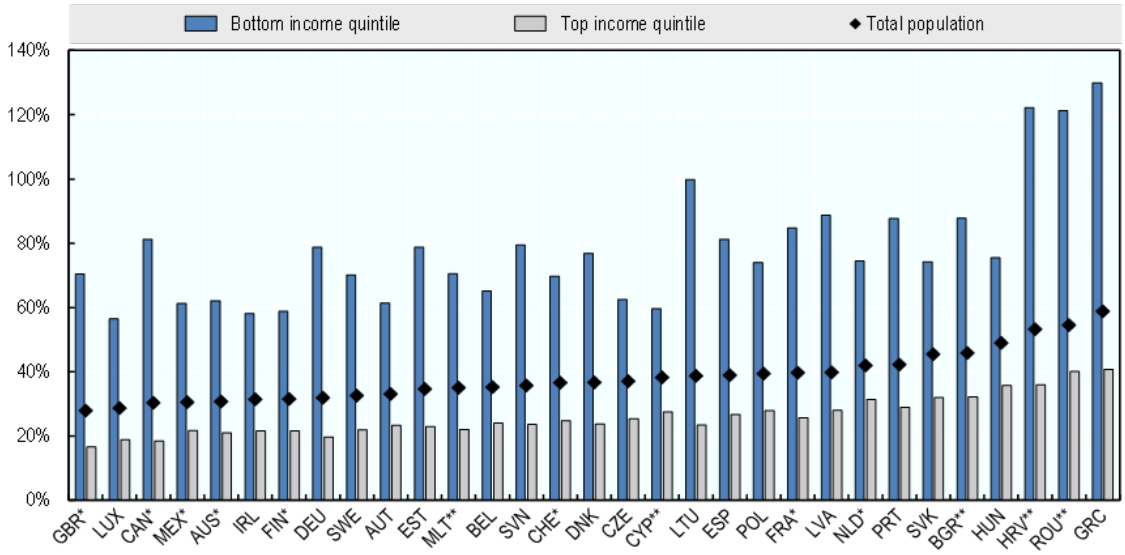
On average, essential expenditures represent a larger share of disposable income for households headed by younger (aged less than 35 years) and older (65 years and over) individuals. This reflects the bell-shaped development of income over a lifetime, rather than absolute differences in consumption levels. However, there are exceptions to the general pattern: in a number of Baltic or eastern EU Member States (for example, Latvia, Poland and Romania), the share of essential expenditures to income increases with *age*, while the reverse is true in Australia, France and Mexico (see Figure B.25)[57]. There are also wide differences across *household types*. Essential expenditures take up a larger share of the income of working-age single-person households and households composed of a single adult with children, while the presence of children does not systematically translate into higher shares for households with two or more adults (see Figure B.26).

---

56 Essential and pre-committed expenditures are defined here in a conservative way, by including only expenditure on housing (rent or mortgage), food and beverages (both at home and outside home), utilities (for example, water, electricity, gas and other (household) fuels) and other housing-related expenditure. The shares shown in Figure 4.14 are computed as the ratio of average essential and pre-committed expenditures to average household disposable income.

57 As already highlighted, the result for Australia could reflect the provision of lump-sum transfers as part of retirement schemes.

**Figure 4.14. Essential expenditure as a share of disposable income by income quintile, 2015 or closest available year**



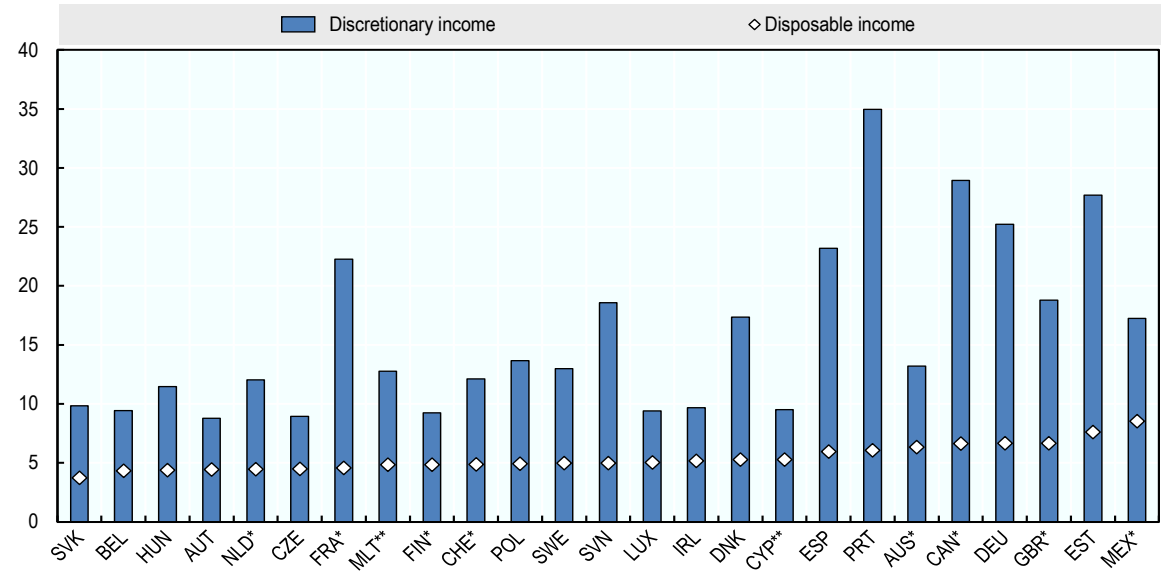

Note: Countries are ranked in ascending order of the share of income spent on essential expenditures for the total population. Values for Switzerland do not include housing-related consumption expenditure other than expenditure on utilities. In Eurostat's centralised exercise, essential expenditures do not include mortgages and information on rents is underestimated or not available for some of the countries. Data are not available for Israel, Japan, Korea, Norway and the United States.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

### *4.3.6. Inequality in discretionary income is higher than inequality in disposable income*

By combining spending and income, these new experimental data make it possible to compute additional welfare metrics, such as **discretionary income**, in other words, the amount of household disposable income that is left for saving, investing or spending on non-essential goods and services, such as luxury goods, leisure activities or vacations, once all essential or pre-committed expenditures have been made. Figure 4.15 illustrates how the regressive nature of essential expenditures translates into higher inequality levels in discretionary income than in disposable income. In general, across the countries with available information the S80/S20 ratio for discretionary income is close to 17 compared with a value of 5 when computed in terms of disposable income. While the experimental classification of these statistics means that cross-country comparisons should be treated with caution, this evidence suggests that inequalities in living standards are generally much higher when looking at discretionary income than at disposable income (or consumption expenditure)[58].

58 In both calculations, households are first ranked according to their household disposable income. The S80/S20 ratios for disposable income shown in Figure 4.15 should not be compared with those published by Eurostat and the OECD in their official reporting, not only due to differences in sources but also because the ratios analysed here refer to disposable rather than equivalised income.

## Figure 4.15. S80/S20 ratio in disposable and discretionary income, 2015 or closest available year




Note Countries are ranked in ascending order of the S80/S20 ratio in disposable income. Data points for Bulgaria, Croatia, Greece, Latvia, Lithuania and Romania are not shown due to the extremely low or negative values of discretionary income for the bottom 20% of households, which resulted in negative or extremely high levels of the S80/S20 ratio. Data for Israel, Japan, Korea, Norway and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

# 5 Conclusions and way forward

Statistics on household economic well-being tend to focus on just one of income, consumption or wealth, with each dimension being considered as a proxy of household economic well-being. However, analysing these three dimensions separately misses out on the important relationships between them. The work of the Eurostat-OECD expert group described in this paper is the first attempt to develop comparable experimental statistics on the joint distribution of ICW for the EU and beyond. Despite the experimental nature of the estimates and the methodological challenges associated with statistical matching, the analysis presented in this paper can help shed light on the extent of economic inequalities in ways that would not be possible with a univariate approach. For instance, an integrated approach can derive measures of savings, compute how much of a household's disposable income is taken up by essential expenditure and demonstrate how relying on wealth assets may help households sustain their living standards in the event of negative income shocks.

The added value of this exercise lies in combining information from different data sources thus enabling the production of two- and three-dimensional indicators. As such, matched ICW data are not meant to produce one-dimensional indicators since results would differ from estimates based on the primary data sources, and their quality would be worse. Moreover, the matched datasets should only be used for estimates on larger subgroups of the population rather than for small subgroups or individual households. This is different for joint ICW distributions that are based on an integrated (single) survey or record linking. Undoubtedly, an integrated survey provides the best basis for both uni- and multi-dimensional household economic analysis, as long as there is a sufficiently large sample size. As a result, whenever possible, an integrated survey or record linking should be considered.

The quality assessment of the matched ICW datasets derived as part of the work of the EG ICW showed good results for the statistical matching of income and consumption and, in most cases, acceptable results for the matching of income (or consumption) and wealth. However, the robustness of the results could be improved in two ways. First, improving the harmonisation of data collection practices and the definition of core socio-demographic and economic variables in the underlying household surveys could reduce the uncertainty associated with statistical matching. Second, improving the availability of auxiliary information would make it possible to test whether the conditional independence assumption is justified (in other words, the assumption that the entire correlation between the target variables goes through the selected matching variables). This would also improve matching between income and consumption or wealth data sources and help assess the robustness of the results. Auxiliary information could be obtained in two ways. The first is a modular approach (as France did to produce estimates as part of the EG ICW). A short module on consumption is introduced in the income survey, and a model is developed to impute the target variable (total consumption expenditure) using data from external and more detailed data sources. The second method is to introduce a limited number of simple additional questions such as 'the value of food consumed at home in a typical week' and 'the value of main residence' into household income surveys.

For EU and other European Economic Area countries, the new Framework Regulation on Social Statistics (IESS)[59], in force since 2021, provides a common framework for different data collections on individuals and households. This ensures that EU social statistics based on data collected from samples are produced in a consistent and coordinated way. This should improve the consistency of EU-SILC and HBS data in the future. Additionally, less frequent microdata collections will be used to complement the core social surveys. As such, an ad hoc module on over-indebtedness, consumption and wealth will be collected every six years. This might be used to model consumption expenditure in EU-SILC based on HBS, or to improve the statistical matching of EU-SILC, HBS and HFCS data and assess the validity and quality of the resulting joint distribution.

In a not-too-distant future, ideally, joint ICW distributions will also be the basis for distributing national account aggregates for income, consumption, savings, and financial and non-financial wealth over groups of households. However, a further alignment of micro- and macroeconomic concepts and definitions is necessary and the uncertainty of both distributional household accounts and ICW joint distributions should be further reduced before combining the two approaches.

[59] Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019 establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples.

# References


Attanasio, O. and L. Pistaferri (2016), "Consumption Inequality", *Journal of Economic Perspectives*, Vol. 30/2, pp. 3-28, https://doi.org/10.1257/JEP.30.2.3. [31]

Balestra, C. et al. (forthcoming), "Inequalities in household wealth – Trends, Drivers and Policy Implications"*, OECD Social, Employment and Migration Working Papers*, OECD Publishing, Paris. [33]

Balestra, C. and R. Tonkin (2018), "Inequalities in household wealth across OECD countries: Evidence from the OECD Wealth Distribution Database"*, OECD Statistics Working Papers*, No. 2018/01, OECD Publishing, Paris, https://doi.org/10.1787/7e1bf673-en. [18]

Bozio, A. et al. (2017), "Do the rich save more? Evidence from linked survey and administrative data", *Oxford Economic Papers*, Vol. 69/4, pp. 1101-1119, https://doi.org/10.1093/oep/gpx024. [29]

Brewer, M., B. Etheridge and C. O'Dea (2017), "Why are Households that Report the Lowest Incomes So Well-off?", *The Economic Journal*, Vol. 127/605, pp. F24-F49, https://doi.org/10.1111/ecoj.12334. [28]

Browning, M., T. Crossley and G. Weber (2003), "Asking Consumption Questions in General Purpose Surveys", *The Economic Journal*, Vol. 113/491, pp. F540-F567, https://doi.org/10.1046/j.0013-0133.2003.00168.x. [19]

Buhmann, B. et al. (1988), "Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) Database", *The Review of Income and Wealth*, Vol. 34/2, pp. 115-142, https://doi.org/10.1111/J.1475-4991.1988.TB00564.X. [36]

Carroll, C. et al. (2017), "The distribution of wealth and the marginal propensity to consume", *Quantitative Economics*, Vol. 8/3, pp. 977-1020, https://doi.org/10.3982/qe694. [25]

Coli, A. et al. (2022), *Distributional national account estimates for household income and consumption: methodological issues and experimental results*, Eurostat, https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/ks-tc-21-010 (accessed on 15 May 2022). [10]

Conti, P., D. Marella and A. Neri (2017), "Statistical matching and uncertainty analysis in combining household income and expenditure data", *Statistical Methods and Applications*, Vol. 26/3, pp. 485-505, https://doi.org/10.1007/S10260-016-0374-7. [24]

D'Orazio, M., M. Di Zio and M. Scanu (2006), *Statistical Matching*, Wiley, https://doi.org/10.1002/0470023554. [21]

European Central Bank (2016), "The Household Finance and Consumption Survey: methodological report for the second wave", *ECB Statistics Paper*, Vol. 17, https://doi.org/10.2866/260900. [32]

Eurostat (2020), *Interaction of household income, consumption and wealth - statistics on taxation*, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Interaction_of_household_income,_consumption_and_wealth_%E2%80%93_statistics_on_taxation (accessed on 8 September 2022). [8]

Eurostat (2013), *Statistical matching of EU-SILC and the Household Budget Survey to compare poverty estimates using income, expenditures and material deprivation*, Eurostat, https://ec.europa.eu/eurostat/fr/web/products-statistical-working-papers/-/KS-RA-13-007. [11]

Fisher, J. et al. (2021), "Inequality in 3-D: Income, Consumption, and Wealth", *Review of Income and Wealth*, Vol. 68/1, pp. 16-42, https://doi.org/10.1111/roiw.12509. [1]

Fisher, J. et al. (2018), "Inequality in 3-D: Income, Consumption, and Wealth", *Finance and Economics Discussion Series*, Vol. 2018/001, https://doi.org/10.17016/feds.2018.001. [15]

Fisher, D. et al. (2009), "Identifying the poorest older Americans", *The Journals of Gerontology. Series B, Psychological sciences and social sciences*, Vol. 64/6, pp. 758-766, https://doi.org/10.1093/GERONB/GBP022. [27]

Garbinti, B. and P. Lamarche (2014), "Les hauts revenus épargnent-ils davantage ?", *Economie et Statistique* 472-473, pp. 49-64, https://doi.org/10.3406/estat.2014.10488. [20]

Garbinti, B. and P. Lamarche (2014), "Les hauts revenus épargnent-ils davantage ?", *Economie et statistique*, Vol. 472/1, pp. 49-64, https://doi.org/10.3406/estat.2014.10488. [30]

Hacker, J. (2018), "Economic security", in *For Good Measure: Advancing Research on Well-being Metrics Beyond GDP*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264307278-10-en. [34]

Kuypers, S., F. Figari and G. Verbist (2021), "Redistribution from a joint income-wealth perspective: Results from 16 European OECD countries"*, OECD Social, Employment and Migration Working Papers*, No. 257, OECD Publishing, Paris, https://doi.org/10.1787/22103c5e-en. [7]

Lamarche, P., F. Oehler and I. Rioboo (2020), "European household's income, consumption and wealth", *Statistical Journal of the IAOS*, Vol. 36/4, pp. 1175-1188, https://doi.org/10.3233/sji-190528. [4]

Le Blanc, J., D. Bundesbank and T. Schmidt (2017), *Do homeowners save more?-Evidence from the Panel on Household Finances (PHF)*, http://www.bundesbank.de/phf (accessed on 13 May 2021). [35]

Leulescu, A. and M. Agafitei (2013), *Statistical matching: a model based approach for data integration*, Eurostat. [12]

OECD (2013), *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*, OECD Publishing, Paris, https://doi.org/10.1787/9789264194830-en. [3]

ONS (2020), *Household income, spending and wealth, Great Britain*, https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/householdincomespendingandwealthgreatbritain/april2016tomarch2018 (accessed on 23 July 2021). [16]

Rässler, S. (2004), "Data Fusion: Identification Problems, Validity and Multiple Imputation", *Austrian Journal of Statistics*, Vol. 33, pp. 153-171. [23]

Renssen, H. (1998), "Use of statistical matching techniques in calibration estimation", *Survey Methodology*, Vol. 24/2, pp. 171-183, https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1998002/article/4354-eng.pdf?st=cjljzpwd. [22]

Serafino, P. and R. Tonkin (2017), *Statistical matching of European Union statistics on income and living conditions (EU-SILC) and the household budget survey 2017 edition*, Eurostat, https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-TC-16-026 (accessed on 9 July 2021). [13]

Serafino, P. and R. Tonkin (2017), *Statistical matching of European Union statistics on income and living conditions (EU-SILC) and the household budget survey 2017 edition*, Eurostat, Luxembourg, https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-TC-16-026 (accessed on 19 July 2021). [26]

Smeeding, T. (2016), "Gates, Gaps, and Intergenerational Mobility: The Importance of an Even Start", in *The Dynamics of Opportunity in America*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-25991-8_8. [6]

Social Metrics Commission (2018), *A new measure of poverty for the UK*, https://socialmetricscommission.org.uk/MEASURING-POVERTY-FULL_REPORT.pdf (accessed on 1 March 2021). [5]

Stiglitz, J., A. Sen and J. Fitoussi (2009), *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Commission on the Measurement of Economic Performance and Social Progress, https://ec.europa.eu/eurostat/documents/8131721/8131772/Stiglitz-Sen-Fitoussi-Commission-report.pdf. [2]

Törmälehto, V., O. Kannas and M. Säylä (2013), *Integrated measurement of household-level income, wealth and non-monetary well-being in Finland*, Working Papers 1, Statistics Finland, https://www.stat.fi/tup/julkaisut/tiedostot/julkaisuluettelo/ywrp1_201300_2013_10518_net.pdf (accessed on 19 July 2021). [14]

United Nations Economic Commission for Europe (2011), *Handbook on Household Income Statistics Second Edition*, United Nations, https://unece.org/statistics/publications/canberra-group-handbook-household-income-statistics-2nd-edition (accessed on 16 May 2021). [17]

Zwijnenburg, J. et al. (2021), "Distribution of household income, consumption and saving in line with national accounts: Methodology and results from the 2020 collection round", *OECD Statistics Working Papers*, No. 2021/01, OECD Publishing, Paris, https://doi.org/10.1787/615c9eec-en. [9]

# Annex A. Additional methodological information

## Table A.1. Countries covered and reference years

| | | Income | Consumption | Wealth |
|---|---|---|---|---|
| Eurostat's centralised exercise | Austria | 2015 | 2015 | 2014 |
| | Belgium | 2014 | 2014 | 2014 |
| | Bulgaria** | 2015 | 2015 | : |
| | Croatia** | 2014 | 2014 | : |
| | Cyprus** | 2015 | 2015 | 2014 |
| | Czech Republic | 2015 | 2015 | : |
| | Denmark | 2015 | 2015 | : |
| | Estonia | 2015 | 2015 | 2013 |
| | Finland | 2016 | 2016 | 2017 |
| | France | 2017 | 2017 | 2017 |
| | Germany | 2013 | 2013 | 2014 |
| | Greece | 2015 | 2015 | 2014 |
| | Hungary | 2015 | 2015 | 2014 |
| | Ireland | 2015 | 2015 | 2013 |
| | Latvia | 2015 | 2015 | 2014 |
| | Lithuania | 2016 | 2016 | : |
| | Luxembourg | 2015 | 2015 | 2014 |
| | Malta** | 2015 | 2015 | 2013 |
| | Netherlands | 2015 | 2015 | 2013 |
| | Poland | 2015 | 2015 | 2014 |
| | Portugal | 2015 | 2015 | 2013 |
| | Romania** | 2015 | 2015 | : |
| | Slovak Republic | 2015 | 2015 | 2014 |
| | Slovenia | 2015 | 2015 | 2014 |
| | Spain | 2015 | 2015 | 2014 |
| | Sweden | 2012 | 2012 | : |
| National exercises | Austria | 2015 | 2015 | 2014 |
| | Canada | 2015 | 2012–2013–2014–2015 | 2016 |
| | Finland | 2016 | 2016 | 2016 |
| | Korea | 2014 | 2015 | 2015 |
| | Switzerland | 2014 | Pooled 2012–2014 | 2014 |
| | United Kingdom | 2015 | 2015 | 2015 |
| | United States | 2016 | 2016 | 2016 |

Note: Data for the United Kingdom are limited to Great Britain. ** denotes non-OECD countries
Source: Eurostat and EG ICW members.

## Table A.2. Hellinger distances between common variables in EU-SILC and HBS in Eurostat's centralised exercise

| | Household size | Household type | Degree of urbanisation | Age of reference person (RP) | Aggregate level of education (RP) | Aggregate activity status (RP) | Occupation status (RP) | Tenure status | Tenure and rent | Aggregate source of income | Income ventiles |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Austria** | 0.002 | 0.006 | 0.028 | 0.013 | 0.026 | 0.052 | 0.578 | 0.006 | 0.019 | 0.013 | 0.001 |
| **Belgium** | 0.008 | 0.020 | 0.235 | 0.041 | 0.032 | 0.047 | 0.248 | 0.027 | 0.048 | 0.059 | 0.001 |
| **Bulgaria**** | 0.049 | 0.063 | 0.036 | 0.062 | 0.070 | 0.085 | 0.259 | 0.029 | 0.029 | 0.046 | 0.002 |
| **Croatia**** | 0.028 | 0.048 | 0.030 | 0.029 | 0.093 | 0.070 | 0.181 | 0.004 | 0.004 | 0.014 | 0.002 |
| **Cyprus**** | 0.058 | 0.065 | 0.011 | 0.016 | 0.011 | 0.071 | 0.278 | 0.063 | 0.063 | 0.008 | 0.003 |
| **Czech Republic** | 0.050 | 0.120 | 0.010 | 0.042 | 0.060 | 0.087 | 0.606 | 0.261 | 0.261 | 0.011 | 0.001 |
| **Denmark** | 0.046 | 0.062 | 0.021 | 0.020 | 0.067 | 0.099 | 0.282 | 0.043 | 0.084 | 0.014 | 0.003 |
| **Estonia** | 0.013 | 0.024 | 0.157 | 0.023 | 0.142 | 0.079 | 0.299 | 0.004 | 0.004 | 0.052 | 0.003 |
| **Finland** | 0.010 | 0.020 | 0.013 | 0.008 | 0.022 | 0.054 | 0.286 | 0.020 | 0.020 | 0.012 | 0.002 |
| **France** | 0.005 | 0.021 | 0.008 | 0.036 | 0.021 | 0.111 | 0.234 | 0.025 | 0.033 | 0.040 | 0.002 |
| **Germany** | 0.004 | 0.015 | 0.130 | 0.043 | 0.376 | 0.064 | 0.588 | 0.044 | 0.048 | 0.007 | 0.000 |
| **Greece** | 0.000 | 0.017 | 0.034 | 0.043 | 0.096 | 0.062 | 0.186 | 0.041 | 0.041 | 0.015 | 0.050 |
| **Hungary** | 0.003 | 0.009 | 0.002 | 0.024 | 0.004 | 0.019 | 0.217 | 0.023 | 0.035 | 0.055 | 0.001 |
| **Ireland** | 0.005 | 0.047 | 0.067 | 0.064 | 0.074 | 0.069 | 0.243 | 0.030 | 0.030 | 0.023 | 0.002 |
| **Latvia** | 0.018 | 0.033 | 0.173 | 0.012 | 0.079 | 0.099 | 0.243 | 0.092 | 0.135 | 0.011 | 0.002 |
| **Lithuania** | 0.010 | 0.025 | 0.018 | 0.016 | 0.054 | 0.096 | 0.276 | 0.005 | 0.006 | 0.051 | 0.004 |
| **Luxembourg** | 0.046 | 0.057 | 0.234 | 0.039 | 0.164 | 0.116 | 0.270 | 0.012 | 0.046 | 1.000 | 0.002 |
| **Malta**** | 0.020 | 0.027 | 0.038 | 0.010 | 0.030 | 0.045 | 0.282 | 0.149 | 0.148 | 0.002 | 0.002 |
| **Netherlands** | 0.008 | 0.014 | 0.009 | 0.009 | 0.912 | 0.166 | 0.590 | 0.050 | 0.052 | 0.022 | 0.001 |
| **Poland** | 0.002 | 0.067 | 0.006 | 0.035 | 0.187 | 0.199 | 0.225 | 0.014 | 0.080 | 0.018 | 0.001 |
| **Portugal** | 0.001 | 0.017 | 0.009 | 0.034 | 0.010 | 0.024 | 0.220 | 0.012 | 0.017 | 0.036 | 0.001 |
| **Romania**** | 0.062 | 0.066 | 0.011 | 0.057 | 0.113 | 0.085 | 0.203 | 0.041 | 0.041 | 0.017 | 0.020 |
| **Slovak Republic** | 0.103 | 0.110 | 0.101 | 0.059 | 0.020 | 0.069 | 0.062 | 0.025 | 0.025 | 0.018 | 0.002 |
| **Slovenia** | 0.048 | 0.067 | 0.018 | 0.028 | 0.056 | 0.041 | 0.245 | 0.565 | 0.565 | 0.005 | 0.002 |
| **Spain** | 0.001 | 0.012 | 0.015 | 0.017 | 0.049 | 0.081 | 0.238 | 0.108 | 0.111 | 0.103 | 0.001 |
| **Sweden** | 0.037 | 0.132 | 0.190 | 0.052 | 0.064 | 0.958 | 0.582 | 0.139 | 0.124 | 0.113 | 0.003 |

Note: Hellinger distances above the threshold of 0.05 (used as a rule of thumb) are shaded. Source: Eurostat calculations based on EU-SILC and HBS microdata.

## Table A.3. Hellinger distances between common variables in donor and recipient datasets in country-based exercises

| | Common variable | Hellinger distance income-consumption datasets | Hellinger distance income-wealth datasets |
|---|---|---|---|
| **Austria** | NUTS 2 region | 0.000 | 0.006 |
| | Level of education of reference person | 0.012 | 0.031 |
| | Dwelling year of construction | 0.016 | 0.078 |
| | Household has a private car | : | 0.000 |
| | Tenure status | 0.044 | 0.044 |
| | Sex of reference person | 0.000 | 0.000 |
| | Size of municipality | 0.029 | : |
| | Number of children in the household | 0.018 | 0.045 |
| | Number of persons in the household | 0.028 | 0.063 |
| | Size of dwelling | 0.017 | 0.025 |
| | Age of reference person | 0.022 | 0.050 |
| | Employee income within the household | 0.031 | 0.150 |
| | Household income self-reported | 0.044 | 0.060 |
| | Household income estimated from administrative data | 0.019 | : |
| | Pension income within the household | 0.027 | 0.114 |
| | Taxable income | 0.026 | : |
| **Canada** | Household disposable income quintile | 0.000 | : |
| | Household type | 0.022 | : |
| | Housing tenure | 0.024 | : |
| | Age of household reference person | 0.014 | : |
| | Sex of household reference person | 0.013 | : |
| | Region (Atlantic provinces, Quebec, Ontario, Prairie provinces, British Columbia) | 0.013 | : |
| | Household size | 0.022 | : |
| | Number of children in household | 0.018 | : |
| | Urbanisation (urban and rural areas) | 0.027 | : |
| **Finland** | Ownership of a car | 0.012 | : |
| | Household has debt | 0.015 | : |
| | Income decile | 0.002 | : |
| | Education level of reference person | 0.022 | : |
| | Socio-economic group of reference person | 0.015 | : |
| | Region | 0.003 | : |
| **South Korea** | Household income quintile | 0.018 | : |
| | Age of household reference person | 0.028 | : |
| | Household size | 0.076 | : |
| | Marital status of reference person | 0.012 | : |
| | Employment status of reference person | 0.067 | : |
| | Housing | 0.051 | : |
| **Switzerland** | Activity status of reference person | 0.075 | : |
| | Age of reference person | 0.029 | : |
| | Sex of reference person | 0.034 | : |
| | Income decile | 0.008 | : |
| | Greater region | 0.014 | : |
| | Linguistic region | 0.007 | : |
| | Household size | 0.053 | : |
| | Tenure status (1st/2nd/3rd/4th quartile of paid rents/no rents paid) | 0.025 | : |

| | Common variable | Hellinger distance income-consumption datasets | Hellinger distance income-wealth datasets |
|---|---|---|---|
| **United Kingdom** | Wages | 0.001 | : |
| | Self-employment income | 0.015 | : |
| | Pension income | 0.014 | : |
| | Total income | 0.000 | : |
| | Benefit income | 0.013 | : |
| | Property income | 0.001 | : |
| | Age (banded) | 0.046 | : |
| | North East (region) | 0.008 | : |
| | North West (region) | 0.003 | : |
| | Yorkshire and the Humber (region) | 0.001 | : |
| | East Midlands (region) | 0.004 | : |
| | West Midlands (region) | 0.004 | : |
| | East of England (region) | 0.005 | : |
| | London (region) | 0.007 | : |
| | South East (region) | 0.011 | : |
| | South West (region) | 0.007 | : |
| | Wales (region) | 0.001 | : |
| | Scotland (region) | 0.001 | : |
| | Northern Ireland (region) | 0.010 | : |
| | Council tax band | 0.022 | : |
| | Mixed (ethnicity) | 0.014 | : |
| | Asian (ethnicity) | 0.001 | : |
| | Black(ethnicity) | 0.008 | : |
| | Chinese (ethnicity) | 0.001 | : |
| | Other (ethnicity) | 0.003 | : |
| | Single (marital status) | 0.045 | : |
| | Married (marital status) | 0.022 | : |
| | Civil partnership (marital status) | 0.008 | : |
| | Separated (marital status) | 0.004 | : |
| | Divorced (marital status) | 0.028 | : |
| | Widowed (marital status) | 0.002 | : |
| | Civil partner separated (marital status) | 0.002 | : |
| | Gender | 0.049 | : |
| | Degree (education) | 0.033 | : |
| **United States** | Race | 0.12 | : |
| | Education | 0.20 | : |
| | Rank in the income distribution | 0.07 | : |
| | Negative income | 0.31 | : |
| | Government transfer income | 0.02 | : |
| | Wage or salary income | 0.03 | : |
| | Positive capital income | 0.09 | : |
| | Negative capital income | 0.18 | : |

Source: National or country-based calculations based on income, consumption and wealth original data sources.

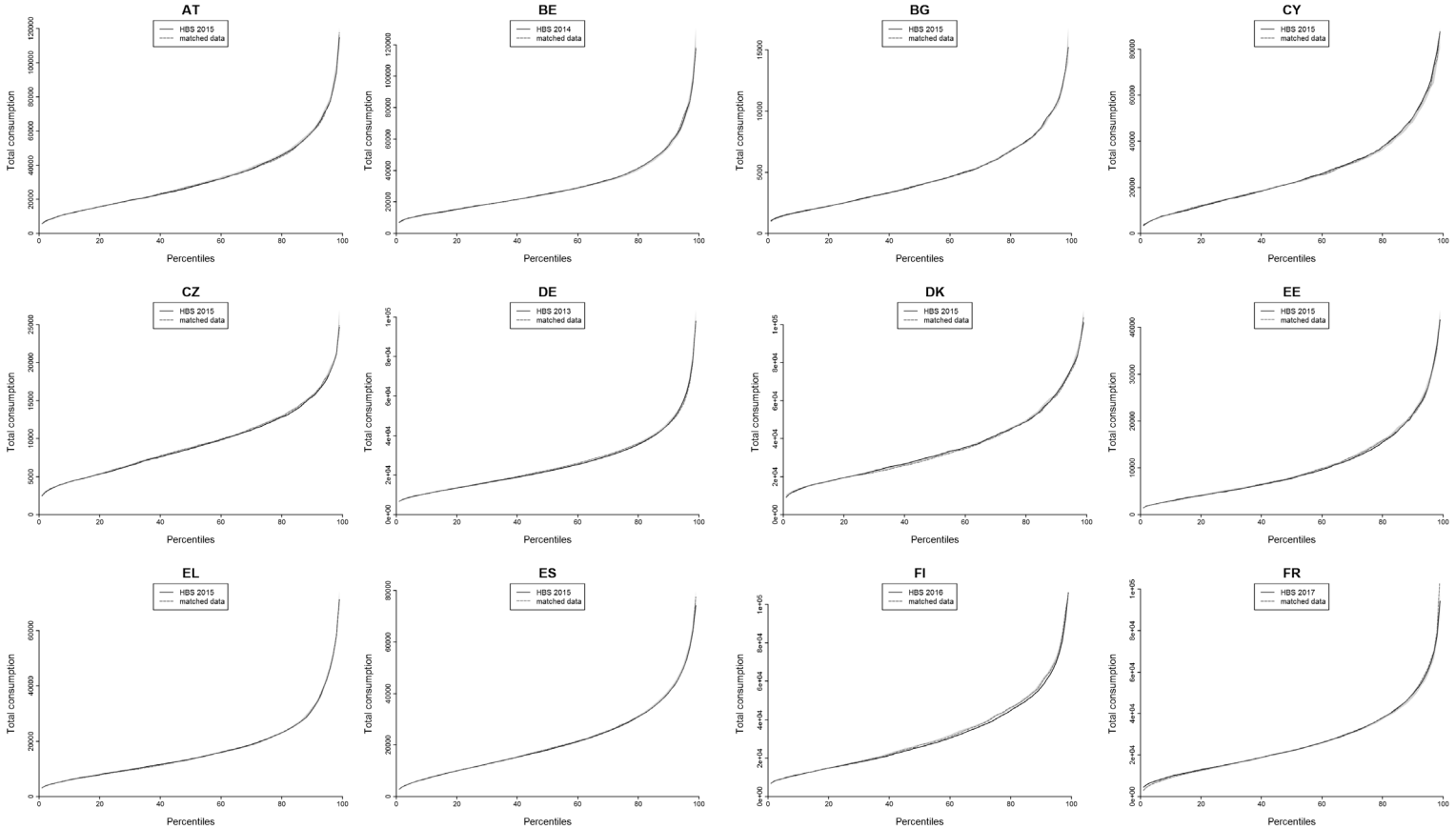
## Table A.4. Final list of matching variables and adjusted R-squared values as obtained with a regression involving all matching variables

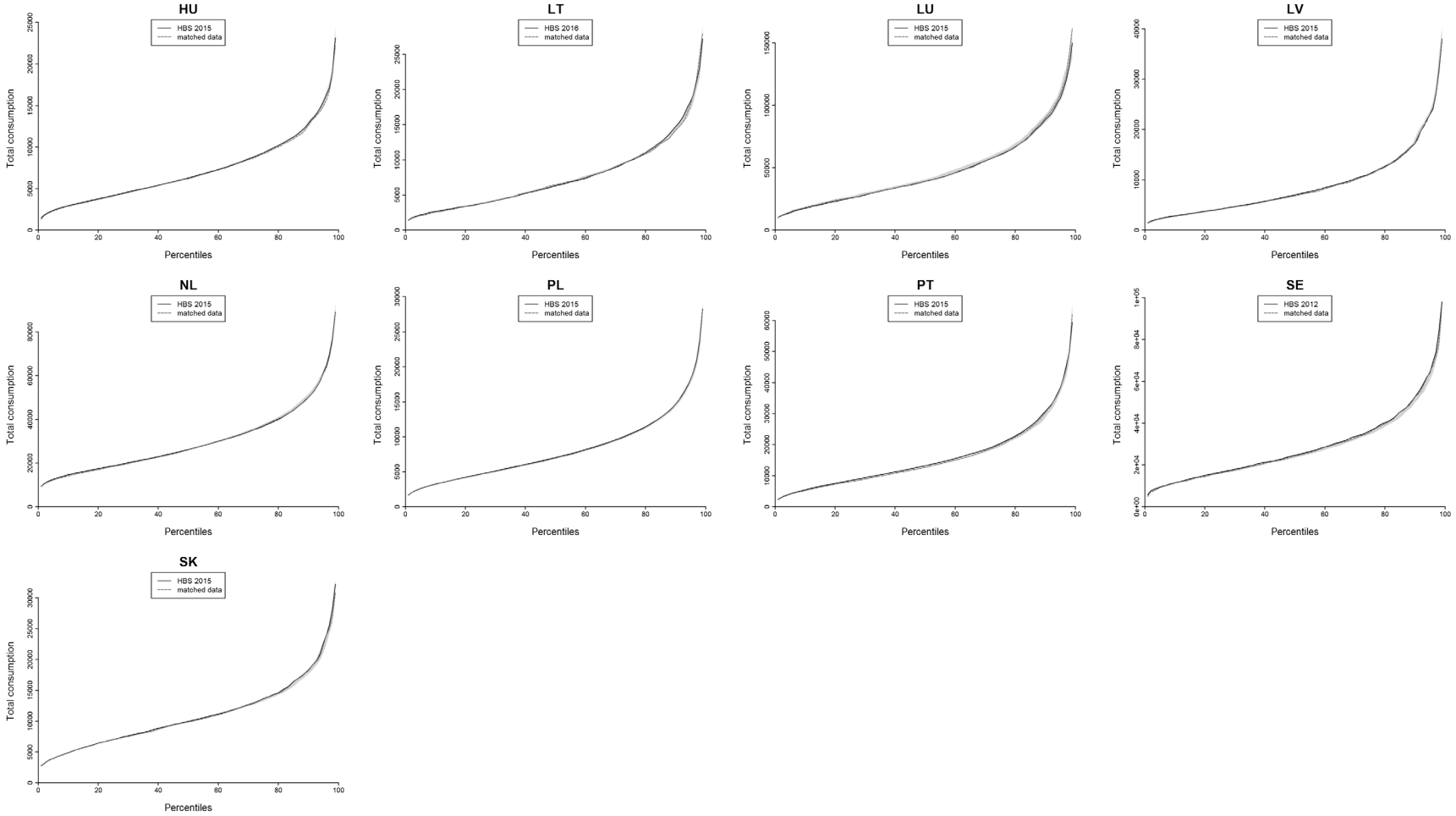
| | | Matching variables | R-squared | |
|---|---|---|---|---|
| | | | consumption | income |
| Eurostat's centralised exercise | **Austria** | household type, level of education of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.30 | 0.78 |
| | **Belgium** | household type, level of education of household reference person, activity status of household reference person, income ventile | 0.26 | 0.72 |
| | **Bulgaria**** | household type, age class of household reference person, degree of urbanisation, level of education of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.67 | 0.83 |
| | **Croatia**** | household type, activity status of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.57 | 0.95 |
| | **Cyprus**** | household type, age class of household reference person, level of education of household reference person, tenure, income ventile | 0.59 | 0.63 |
| | **Czech Republic** | household size, age class of household reference person, level of education of household reference person, main source of income (aggregate), income ventile | 0.63 | 0.79 |
| | **Germany** | household type, age class of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.41 | 0.70 |
| | **Denmark** | household type, level of education of household reference person, tenure, income ventile | 0.39 | 0.83 |
| | **Estonia** | household type, activity status of household reference person, tenure or rent quintile, income ventile | 0.44 | 0.96 |
| | **Finland** | household type, age class of household reference person, activity status of household reference person, tenure or rent quintile, level of education of household reference person, income ventile | 0.48 | 0.79 |
| | **France** | household type, level of education of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.37 | 0.75 |
| | **Greece** | household type, tenure or rent quintile, main source of income (aggregate), age class of household reference person, degree of urbanisation, activity status of household reference person, income ventile | 0.44 | 0.58 |
| | **Hungary** | household type, level of education of household reference person, tenure or rent quintile, degree of urbanisation, main source of income (aggregate), income ventile | 0.59 | 0.83 |
| | **Ireland** | household type, age class of household reference person, activity status of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.50 | 0.75 |
| | **Lithuania** | household type, main source of income (aggregate), income ventile | 0.40 | 0.87 |
| | **Luxembourg** | household type, tenure or rent quintile, income ventile | 0.36 | 0.85 |
| | **Latvia** | household type, age class of household reference person, income ventile | 0.53 | 0.90 |
| | **Malta**** | household type, age class of household reference person, level of education of household reference person, activity status of household reference person, degree of urbanisation, income ventile | 0.27 | 0.89 |
| | **Netherlands** | household type, age class of household reference person, main source of income (aggregate), income ventile | 0.30 | 0.66 |
| | **Poland** | household type, tenure, age class of household reference person, main source of income (aggregate), income ventile | 0.40 | 0.94 |
| | **Portugal** | household type, level of education of household reference person, tenure or rent quintile, activity status of household reference person, main source of income (aggregate), income ventile | 0.44 | 0.79 |

| | | Matching variables | R-squared | |
|---|---|---|---|---|
| | | | consumption | income |
| | **Romania**** | household type, age class of household reference person, tenure or rent quintile, main source of income (aggregate), income ventile | 0.58 | 0.93 |
| | **Slovak Republic** | level of education of household reference person, main source of income (aggregate), income ventile | 0.47 | 0.90 |
| | **Slovenia** | household type, age class of household reference person, degree of urbanisation, level of education of household reference person, activity status of household reference person, main source of income (aggregate), income ventile | 0.45 | 0.77 |
| | **Spain** | household type, age class of household reference person, level of education of household reference person, income ventile | 0.44 | 0.87 |
| | **Sweden** | household size, level of education of household reference person, income ventile | 0.27 | 0.76 |
| **National exercises** | **Austria** | income decile, NUTS level 2, level of education of household reference person, construction year, dwelling size, tenure, household size, number of children | 0.76 | : |
| | **Canada** | household disposable income quintile, household type, housing tenure, age of household reference person, sex of household reference person, region, household size, number of children in household, urbanisation | 0.48 | : |
| | **Finland** | car ownership, household debt, income decile, level of education of household reference person, socio-economic group of household reference person, region | 0.52 | : |
| | **Korea** | income quintile, age of household reference person, household size, marital status of household reference person, activity status of household reference person, tenure | 0.37 | : |
| | **Switzerland** | activity status of household reference person, age of household reference person, sex of household reference person, household size, tenure or rent quintile, income decile, administrative region, linguistic region | 0.35 | : |
| | **United Kingdom** | total income, property income, wages, self-employment income, benefits, age class of household reference person, council tax amount, pension, gender, ethnicity, marital status, region | 0.39 | : |
| | **United States** | race, education, rank in the income distribution, whether the household reported negative income, whether the household received government transfer income, whether the household received wage or salary income, whether the household reported positive capital income such as interest and dividends, whether the household reported negative capital income | : | : |

Source: Eurostat and national calculations based on income, consumption and wealth original data sources.

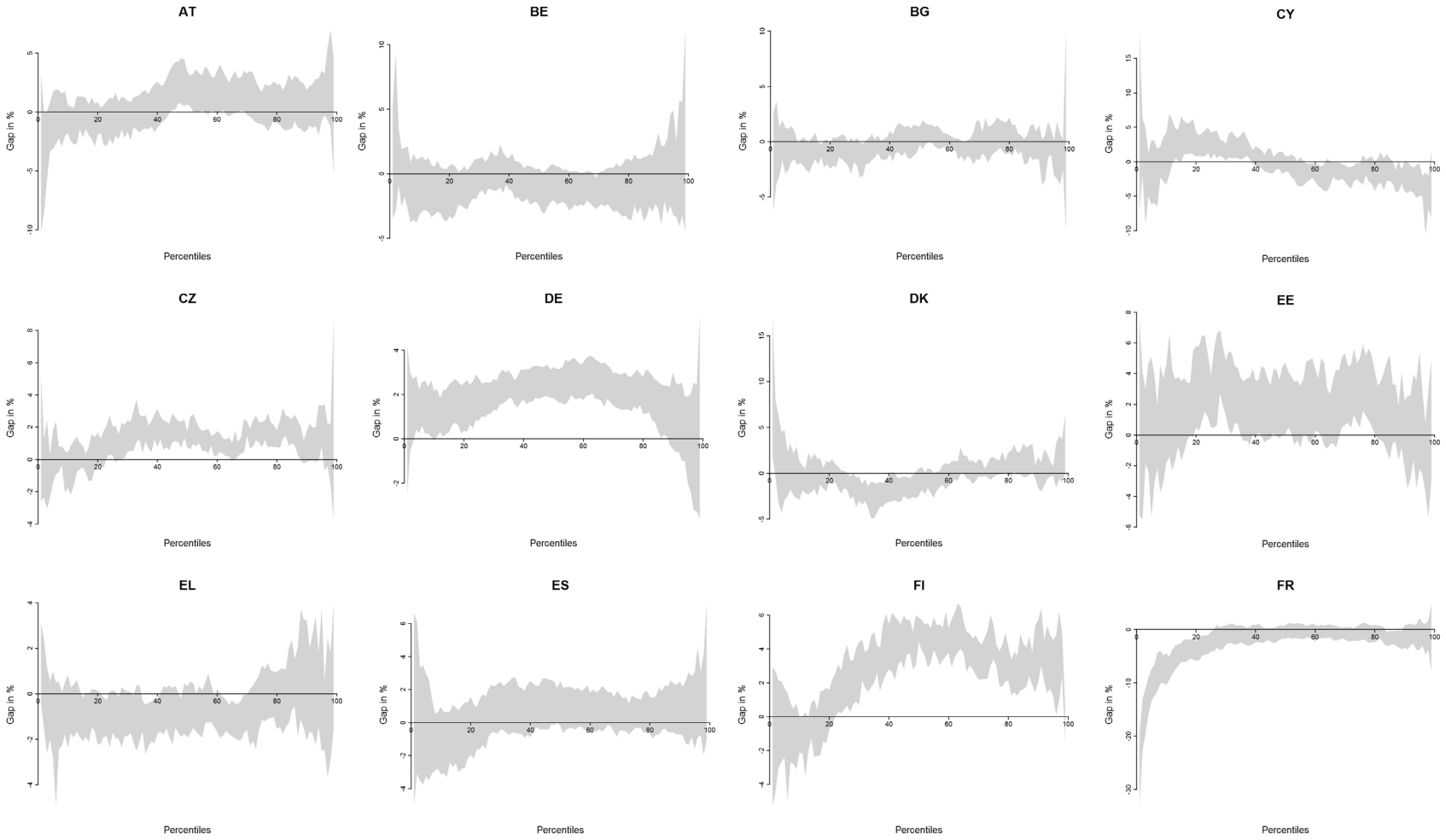
## Figure A.1. Consumption percentiles observed in original HBS data and in the matched IC dataset for countries in Eurostat's centralised exercise

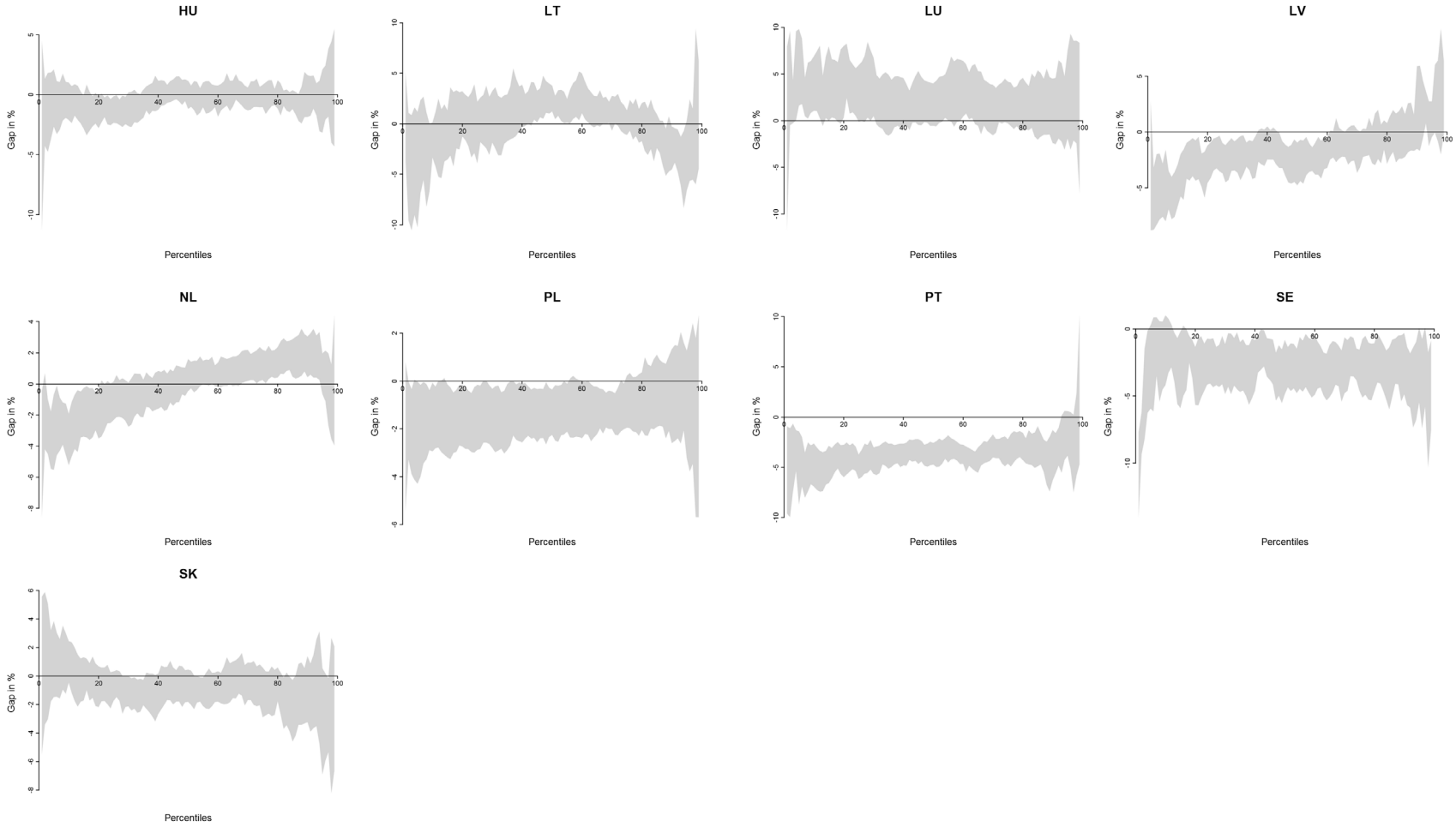





Note: Scales vary across plots.
Source: Eurostat calculations based on joint income, consumption and wealth micro data and original data sources.

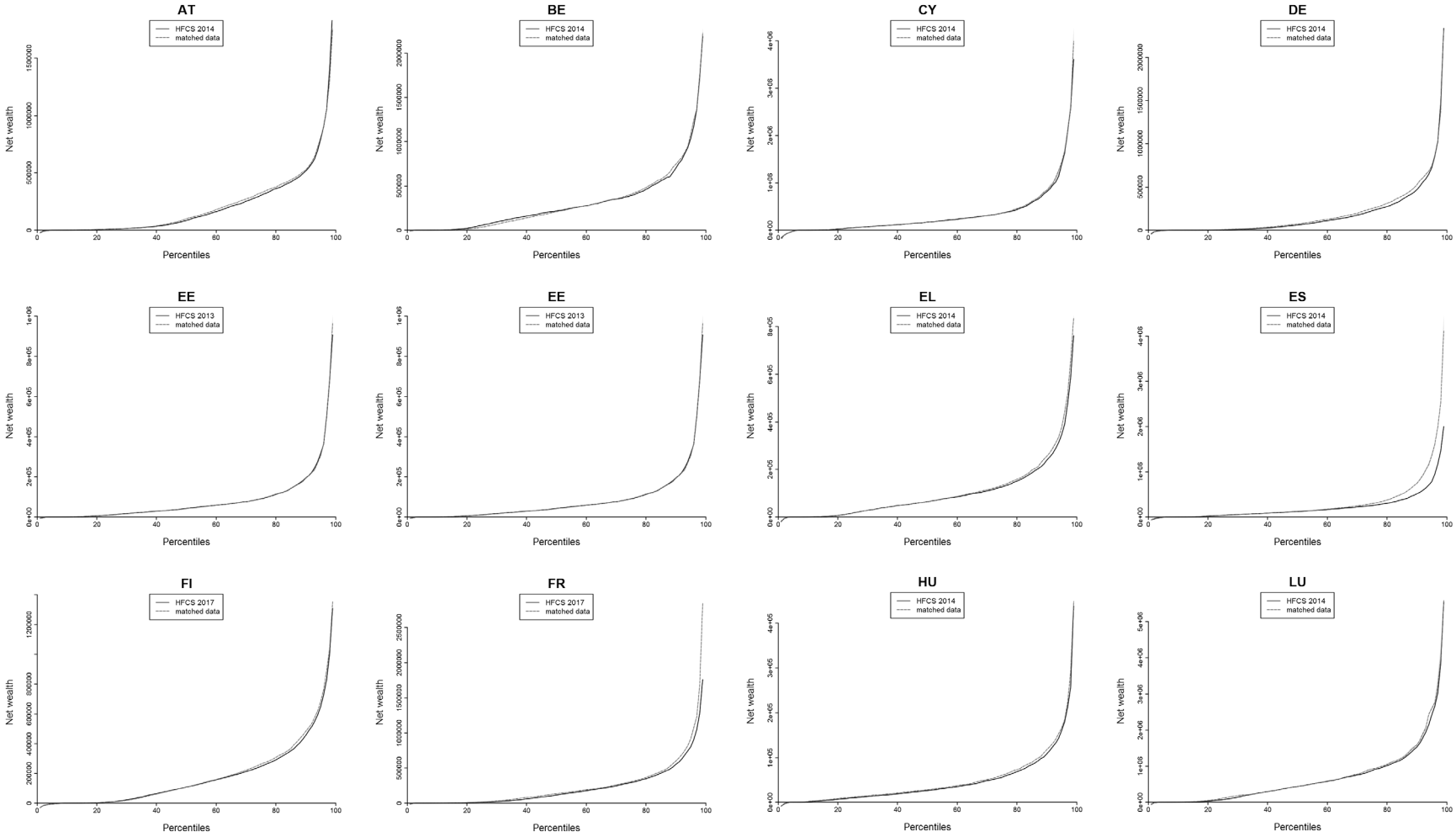
## Figure A.2. Relative gap sizes between consumption percentiles observed in original HBS data and in the matched IC dataset for countries in Eurostat's centralised exercise

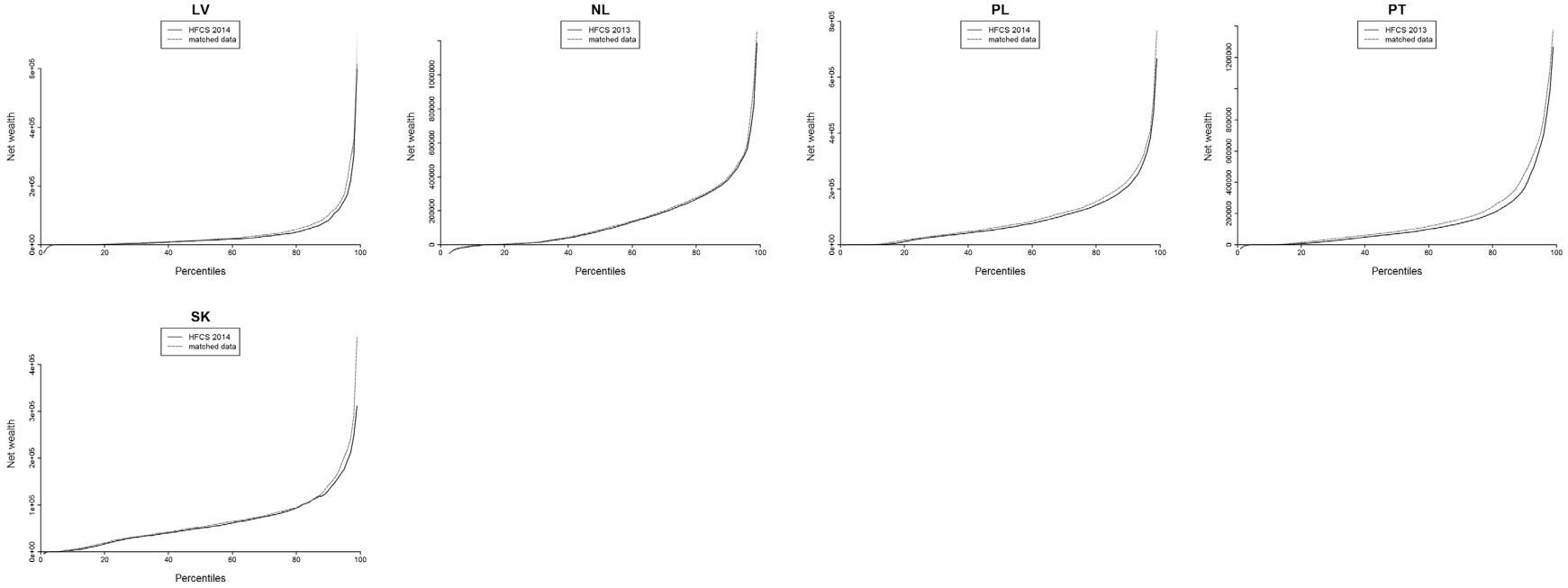





Note: Scales vary across plots.
Source: Eurostat calculations based on joint income, consumption and wealth micro data and original data sources.

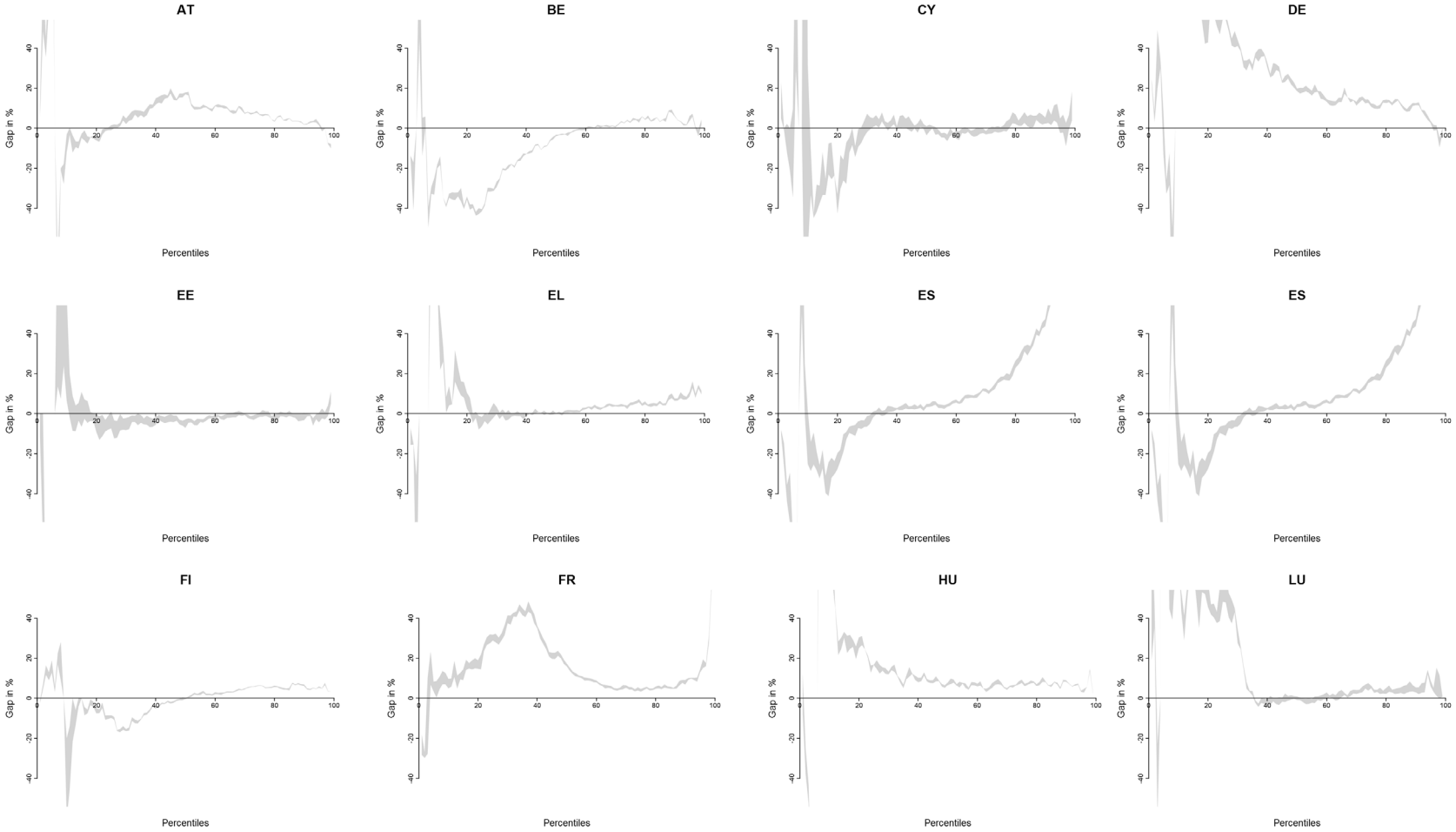
## Figure A.3. Net wealth percentiles observed in original HFCS data and in the matched ICW dataset for countries in Eurostat's centralised exercise

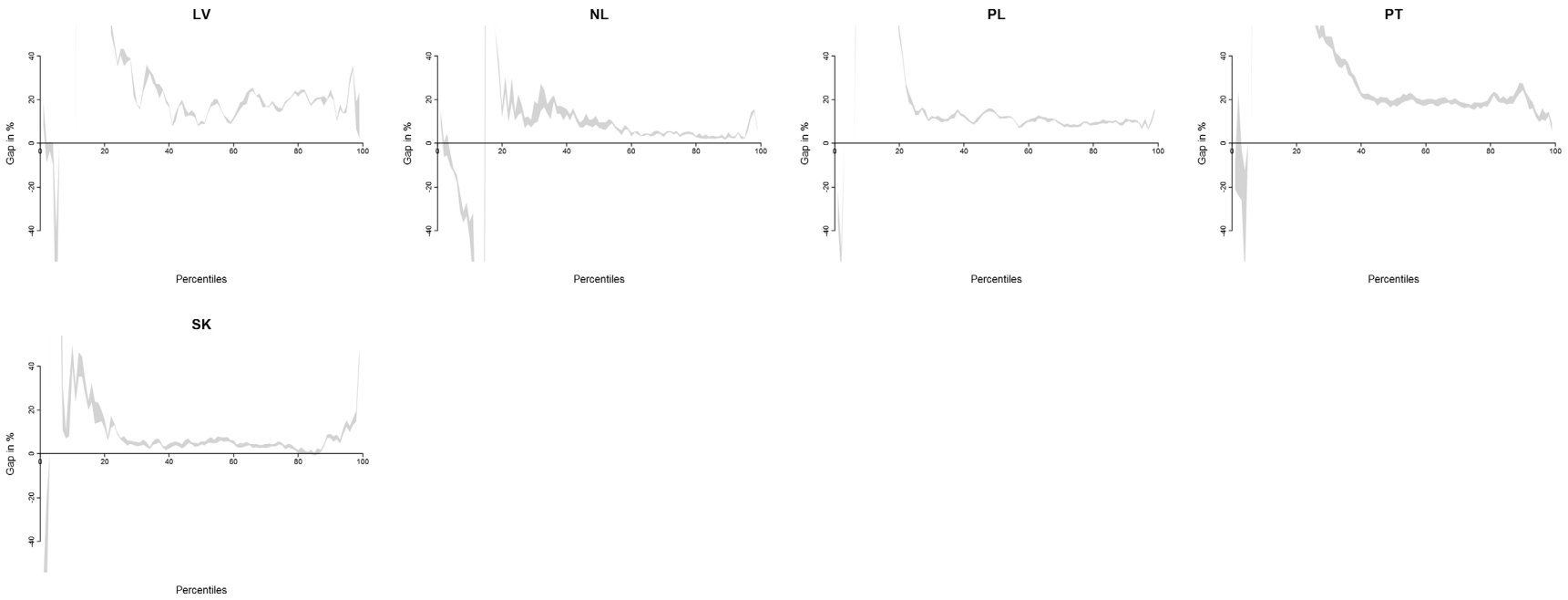





Note: Scales vary across plots.
Source: Eurostat calculations based on joint income, consumption and wealth micro data and original data sources.

## Figure A.4. Relative gap sizes between net wealth percentiles observed in original HFCS data and in the matched ICW dataset for countries in Eurostat's centralised exercise

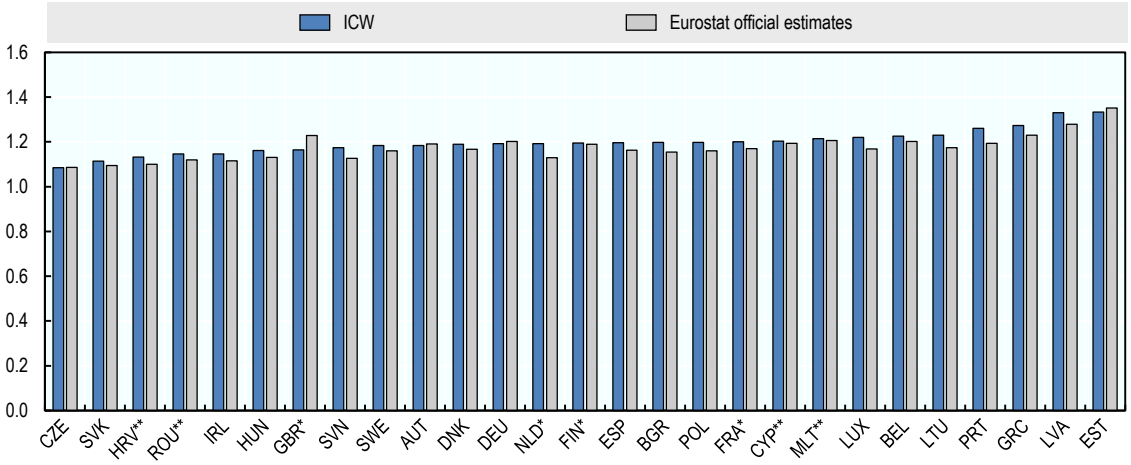





Note: The grey area designates the 95% confidence interval computed through the 30 repetitions of the random hot deck procedure resulting into 30 net wealth estimates.
*Source:* Eurostat calculations based on joint income, consumption and wealth micro data and original data sources.

## Figure A.5. Consumption inequality – experimental versus official statistics, 2015 or closest available year

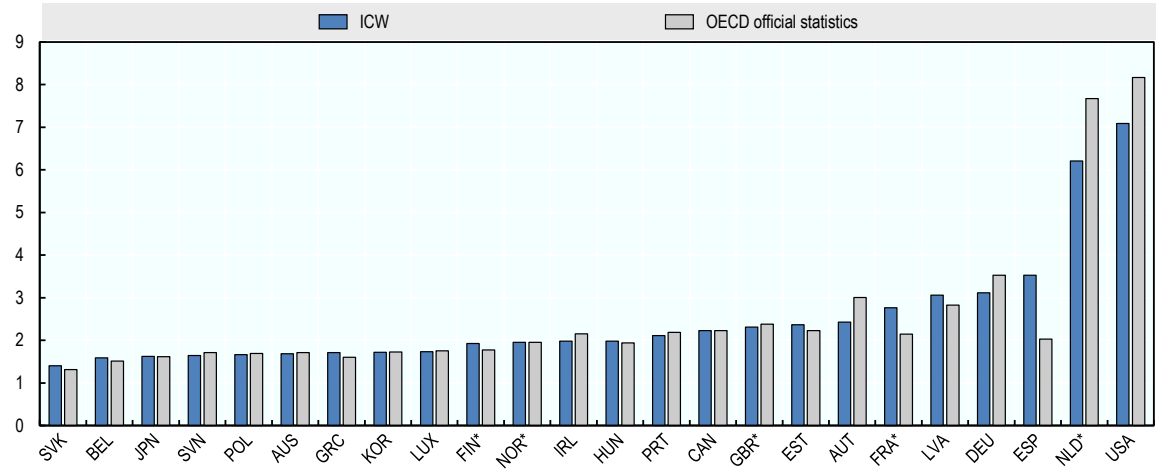
(mean-to-median ratio)




Note: Countries are ranked in ascending order of mean-to-median consumption ratio in the ICW database.
Source: Eurostat-OECD Secretariat calculations based on national estimates, Eurostat's experimental statistics on income, consumption and wealth and Eurostat's public dissemination database (online data code: hbs_exp_t111).

## Figure A.6. Wealth inequality – experimental versus official statistics, 2015 or closest available year

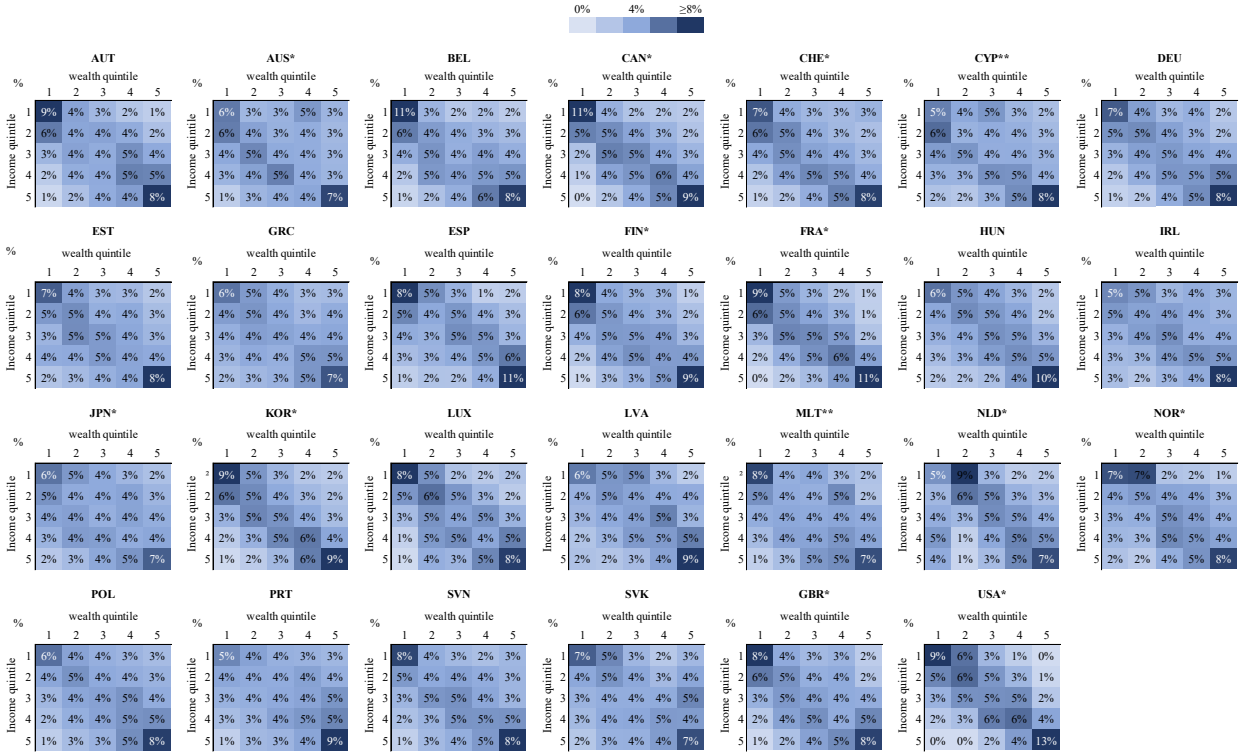
(mean-to-median wealth ratio)




Note: Countries are ranked in ascending order of the mean-to-median wealth ratio in the ICW database.
Source: Eurostat-OECD Secretariat calculations based on national estimates, Eurostat's experimental statistics on income, consumption and wealth and OECD's Wealth Distribution Database (oe.cd/wealth).

# Annex B. Additional charts and tables

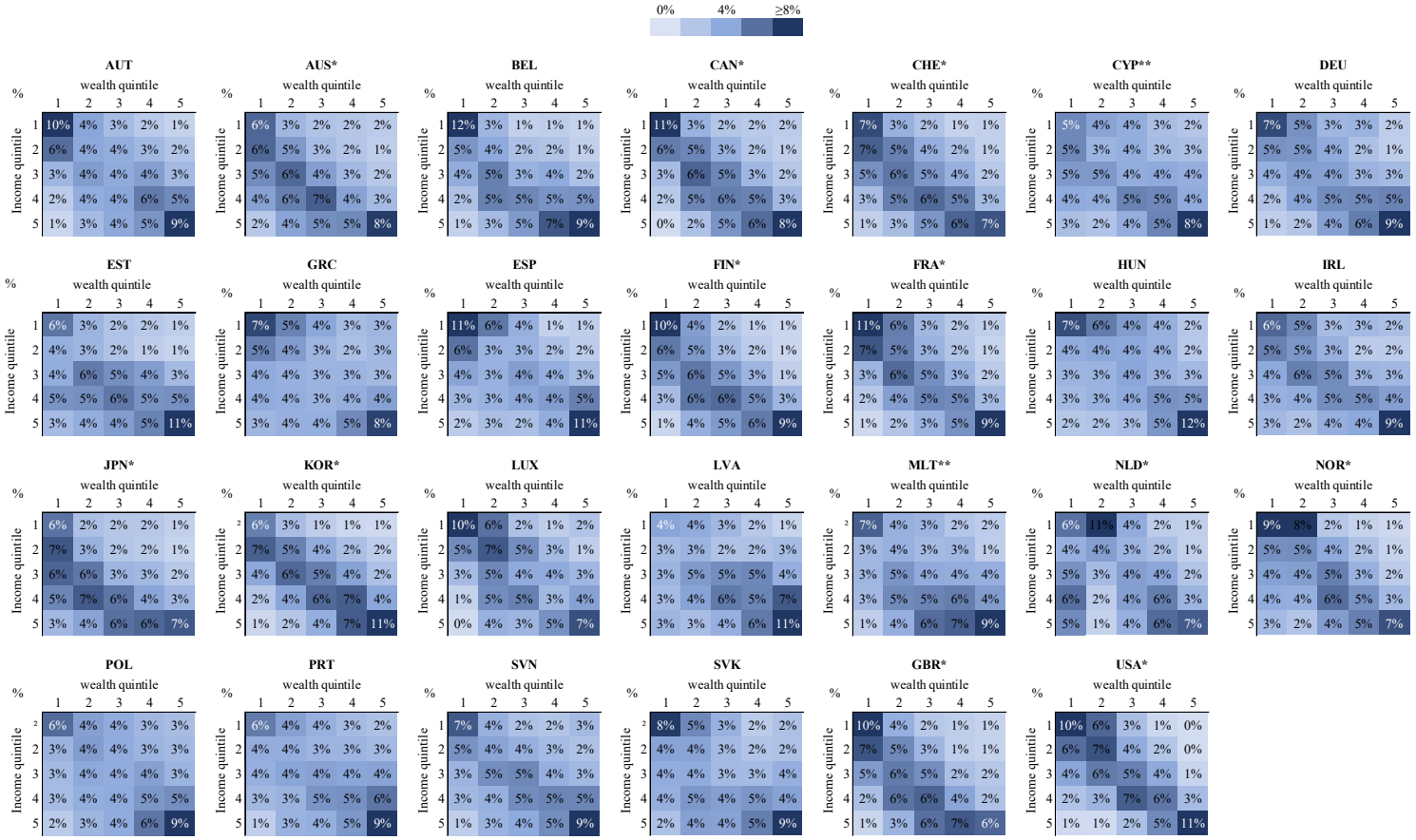
## Figure B.1. Share of households along the joint distribution of household disposable income and net wealth, total population




Note: The value shown is the share of all households in the respective quintiles of the distribution of disposable household income and household net wealth, with 1 denoting the bottom quintile and 5 denoting the top quintile. For instance, in Austria 9% of all households find themselves in the bottom quintile of both the disposable income and the net wealth distribution. Cells with a value larger than 4% have a greater population share than they would if households' position in the wealth and income distributions were completely unrelated; by contrast, cells with a value lower than 4% contain fewer households than expected from a random distribution. * denotes national or country-based estimates, ** denotes non-OECD countries.

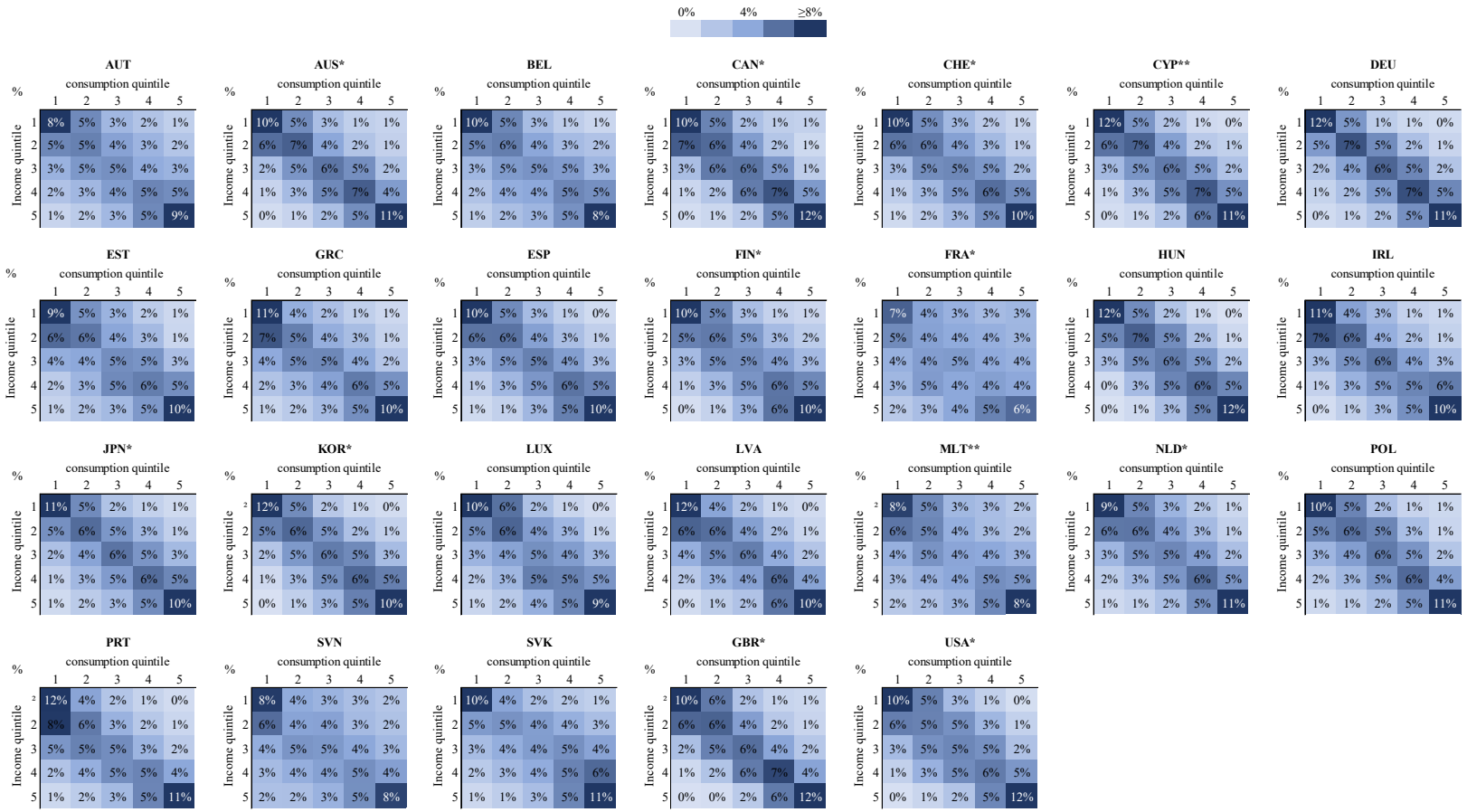
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

## Figure B.2. Share of households along the joint distribution of household disposable income and net wealth, working-age households



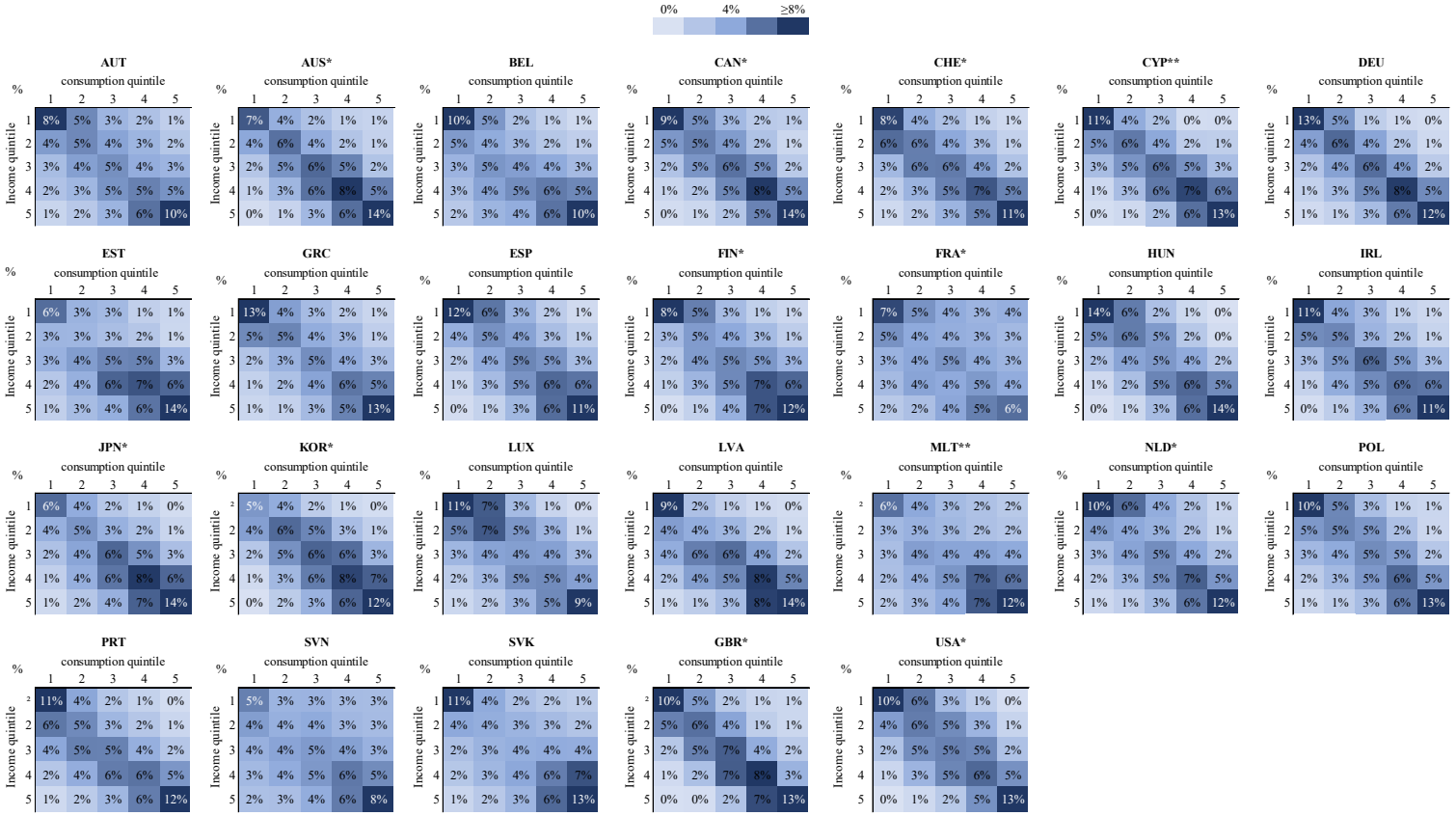

Note: The value shown is the share of working-age households in the respective quintiles of the distribution of disposable household income and household net wealth, with 1 denoting the bottom quintile and 5 denoting the top quintile. For instance, in Austria 10% of working-age households find themselves in the bottom quintile of both the disposable income and the net wealth distribution. Cells with a value larger than 4% have a greater population share than they would if households' position in the wealth and income distributions were completely unrelated; cells with a value lower than 4% contain fewer households than expected from a random distribution. Working-age households are defined as those with a head aged 16–64 years.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.3. Share of households along the joint distribution of household disposable income and consumption expenditure, total population




Note: The value shown is the share of households in the respective quintiles of the distribution of disposable household income and consumption, with 1 denoting the bottom quintile and 5 denoting the top quintile. Cells with a value larger than 4% have a greater population share than they would if households' position in the income and consumption distributions were completely unrelated; cells with a value lower than 4% contain fewer households than expected from a random distribution. Data for Norway are not shown due to possible data issues.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

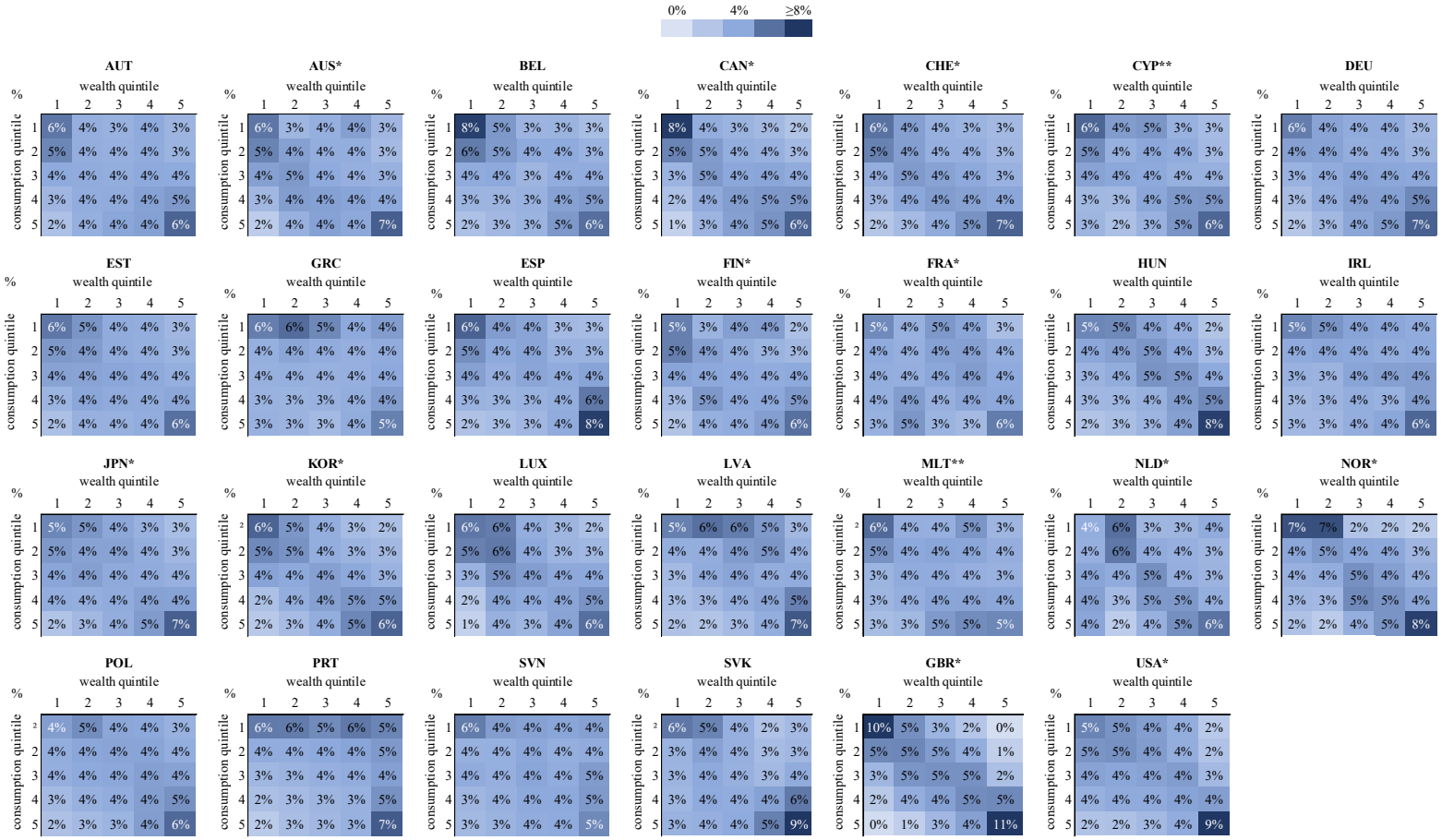
## Figure B.4. Share of households along the joint distribution of household disposable income and consumption expenditure, working-age households




Note: The value shown is the share of working-age households in the respective quintiles of the distribution of household disposable income and household consumption expenditure, with 1 denoting the bottom quintile and 5 denoting the top quintile. Cells with a value larger than 4% have a greater population share than they would if households' position in the income and consumption distributions were completely unrelated; cells with a value lower than 4% contain fewer households than expected from a random distribution. Data for Norway are not shown due to possible data issues. Working-age households are defined as those with a head aged 16-64.

Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

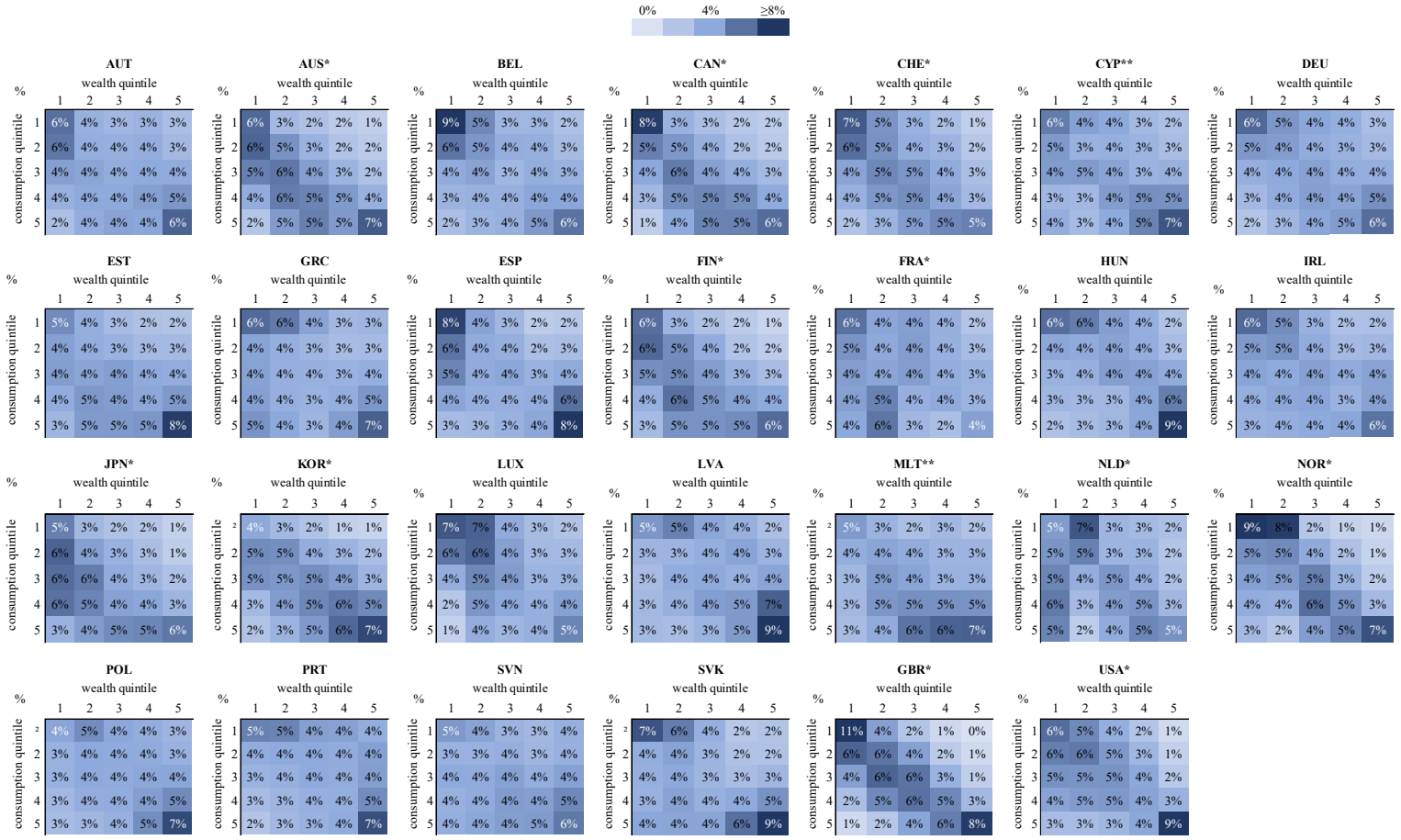
## Figure B.5. Share of households along the joint distribution of household consumption expenditure and net wealth, total population




Note: The value shown is the share of households in the respective quintiles of the distribution of household net wealth and consumption expenditure, with 1 denoting the bottom quintile and 5 denoting the top quintile. Cells with a value larger than 4% have a greater population share than they would if households' position in the wealth and consumption distributions were completely unrelated; cells with a value lower than 4% contain fewer households than expected from a random distribution.

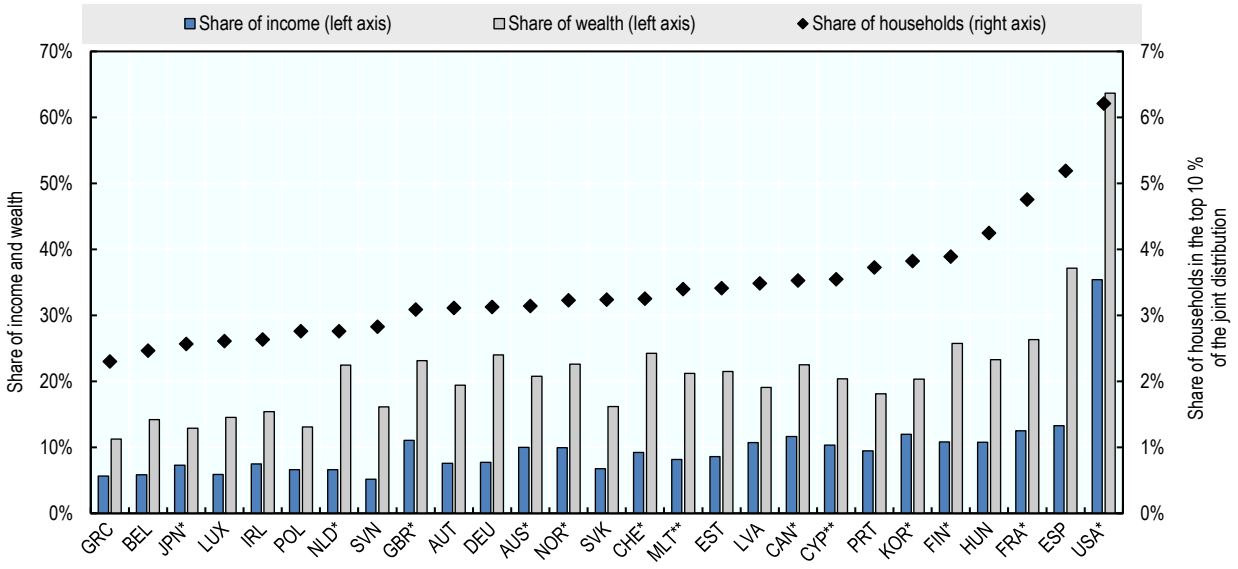
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

## Figure B.6. Share of households along the joint distribution of household consumption expenditure and net wealth, working-age households



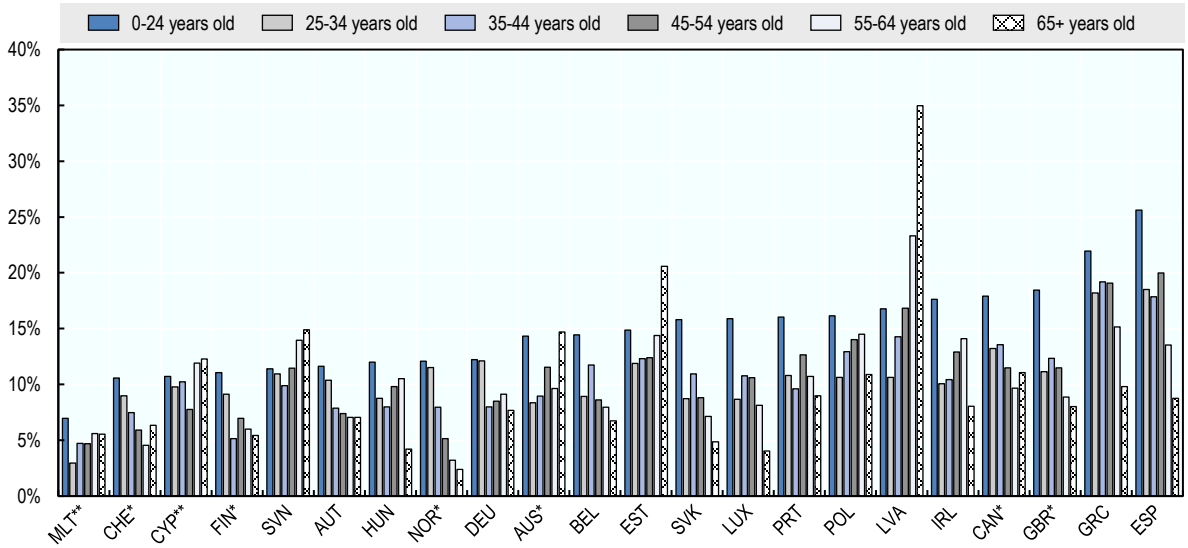

Note: The value shown is the share of working-age households in the respective quintiles of the distribution of household net wealth and consumption expenditure, with 1 denoting the bottom quintile and 5 denoting the top quintile. Cells with a value larger than 4% have a greater population share than they would if households' position in the wealth and consumption distributions were completely unrelated; cells with a value lower than 4% contain fewer households than expected from a random distribution. Working-age households are defined as those with a head aged 16–64 years.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.7. Share of resources held by the top 10% of the joint distribution of income and wealth, 2015 or closest available year



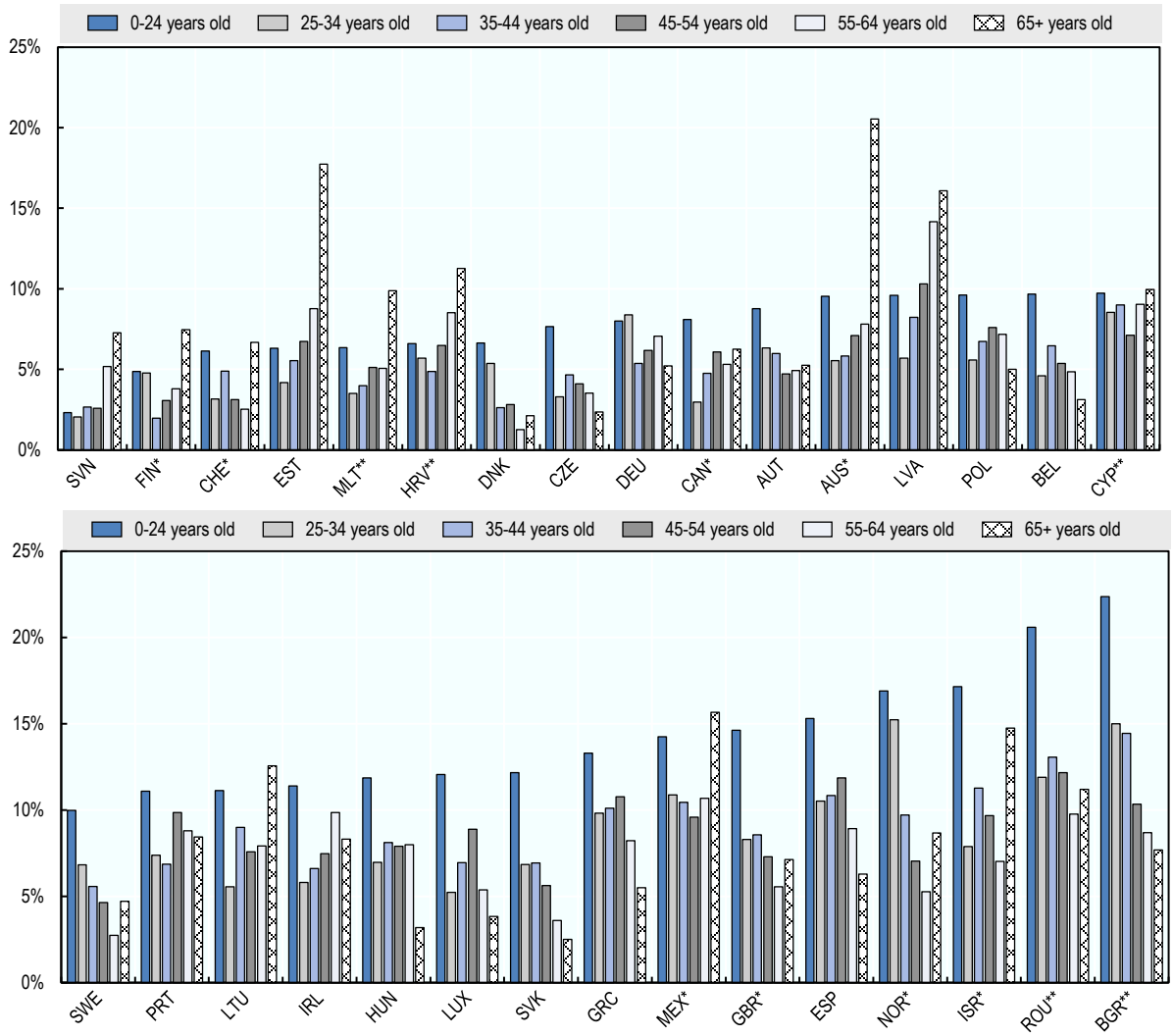

Note: Countries are ranked in ascending order of the share of households belonging to the top 10% of the joint distribution of income and wealth (right vertical axis).
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_res_01).

## Figure B.8. Share of individuals who are both income and asset poor, by age, 2015 or closest available year



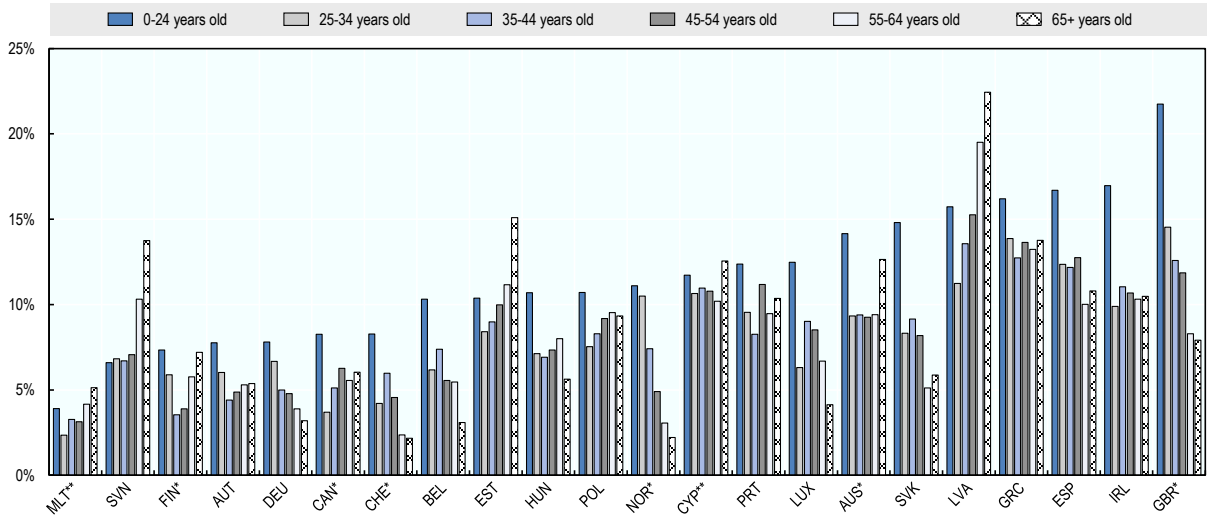

Note: Countries are ranked in ascending order of income and asset risk-of-poverty rates among the youngest age group. Age groups 0–17 and 18–24 years have been combined and the weighted average (0–24 years old) is shown. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the age of the head of household rather than the individual.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_10).

## Figure B.9. Share of individuals who are both income and consumption poor, by age, 2015 or closest available year



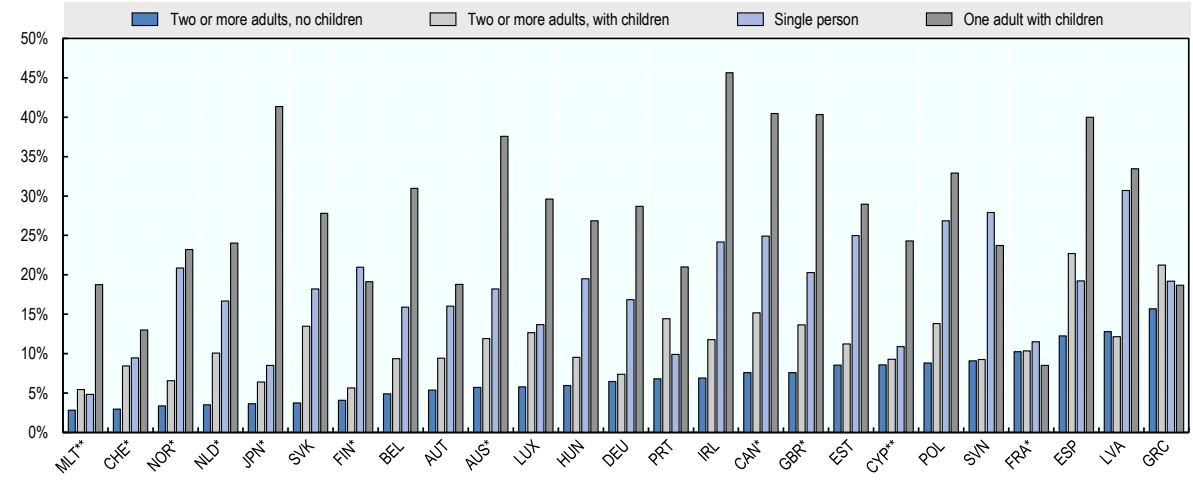

Note: Countries are ranked in ascending order of income and consumption risk-of-poverty rates among the youngest age group. Age groups 0–17 and 18–24 years have been combined and the weighted average (0–24 years old) is shown. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the age of the head of household rather than the individual.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_10).

## Figure B.10. Share of individuals who are both consumption and asset poor, by age, 2015 or closest available year



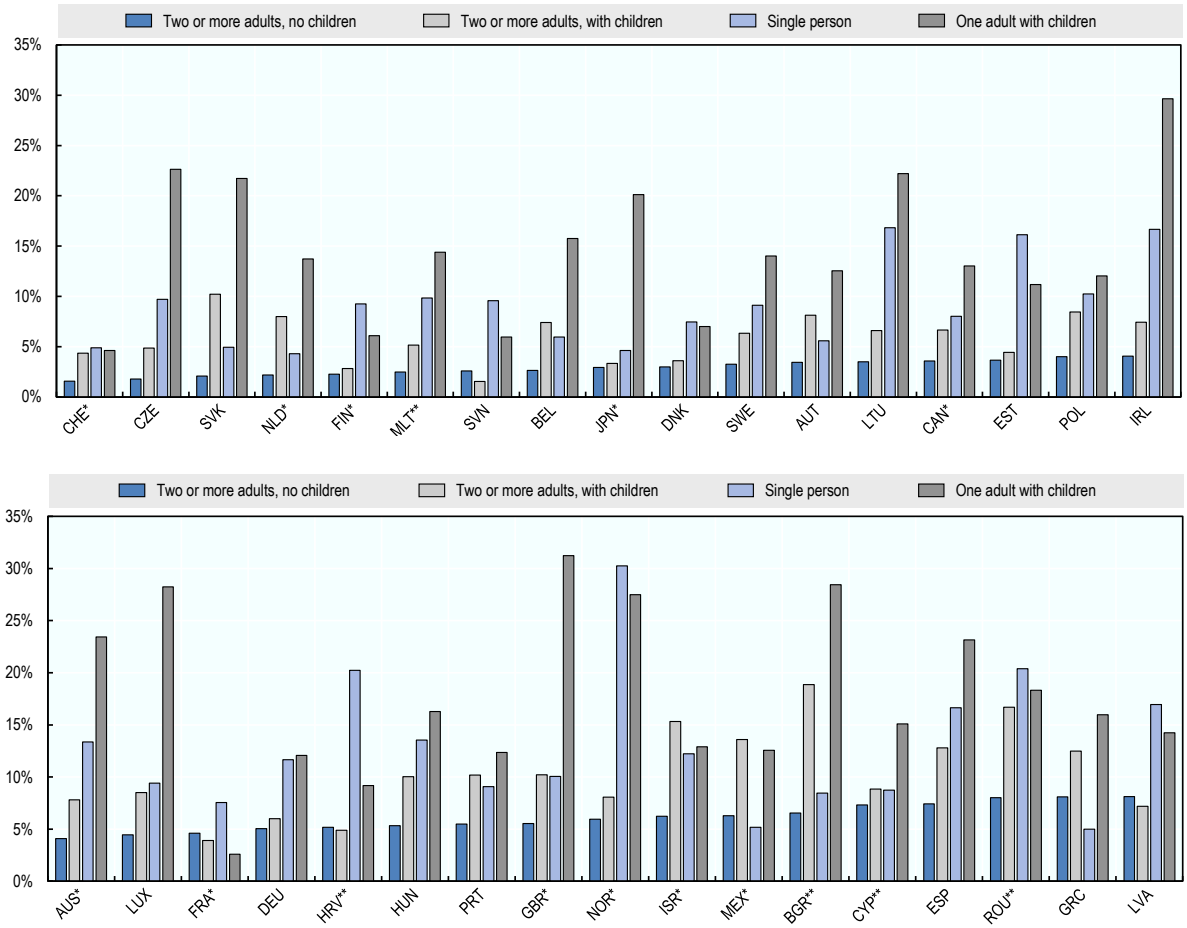

Note: Countries are ranked in ascending order of consumption and asset risk-of-poverty rates among the youngest age group. Age groups 0–17 and 18–24 years have been combined and the weighted average (0–24 years old) is shown. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the age of the head of household rather than the individual. * denotes country-based estimates; ** denotes non-OECD countries.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code:icw_pov_10).

## Figure B.11. Share of individuals who are both income and asset poor, by household type, working-age households, 2015 or closest available year




Note: Countries are ranked in ascending order of income and asset risk-of-poverty rates among working-age people living in households of two or more adults without children. Working-age households are defined as those with a head aged 16–64 years. Risk-of-poverty rates for older people are omitted as they are shown for the age group 65 years and over in Figure B.8. Data for Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_11).

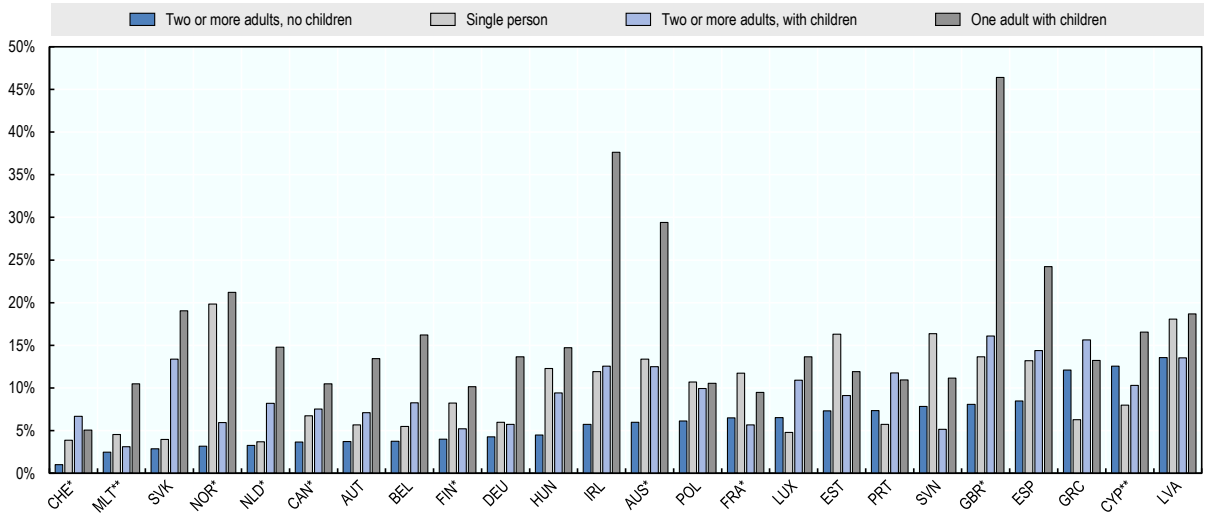
## Figure B.12. Share of individuals who are both income and consumption poor, by household type, working-age households, 2015 or closest available year




Note: Countries are ranked in ascending order of income and consumption risk-of-poverty rates among the working-age people living in households of two or more adults without children. Working-age households are defined as those with a head aged 16–64 years. Risk-of-poverty rates for older people are omitted as they are shown for the age group 65 years and over in Figure B.9. Data for Korea and the United States are not available.

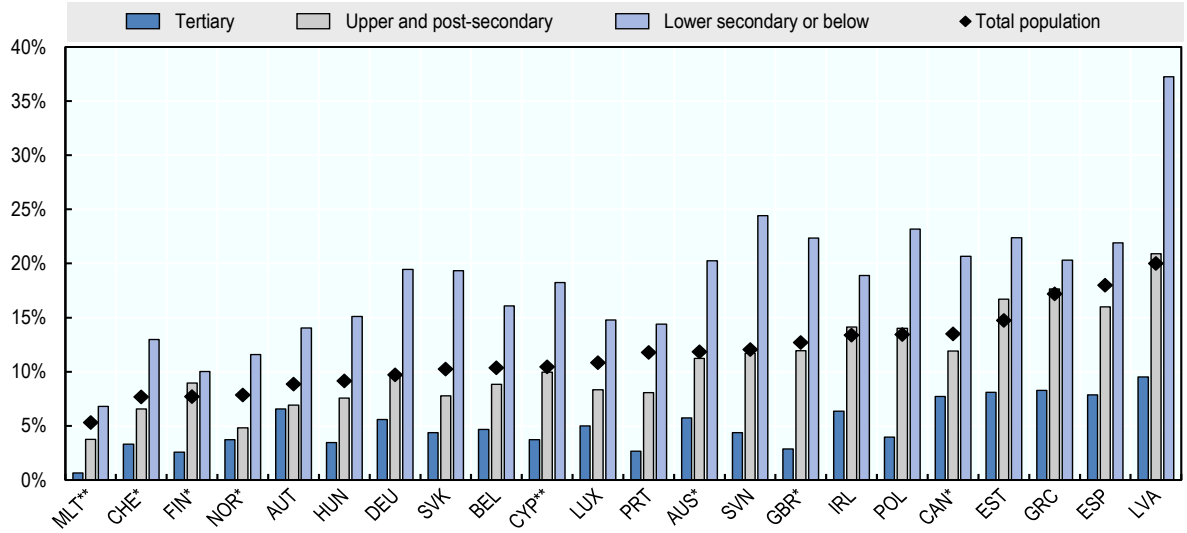
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_11).

## Figure B.13. Share of individuals who are both consumption and asset poor, by household type, working-age households, 2015 or closest available year



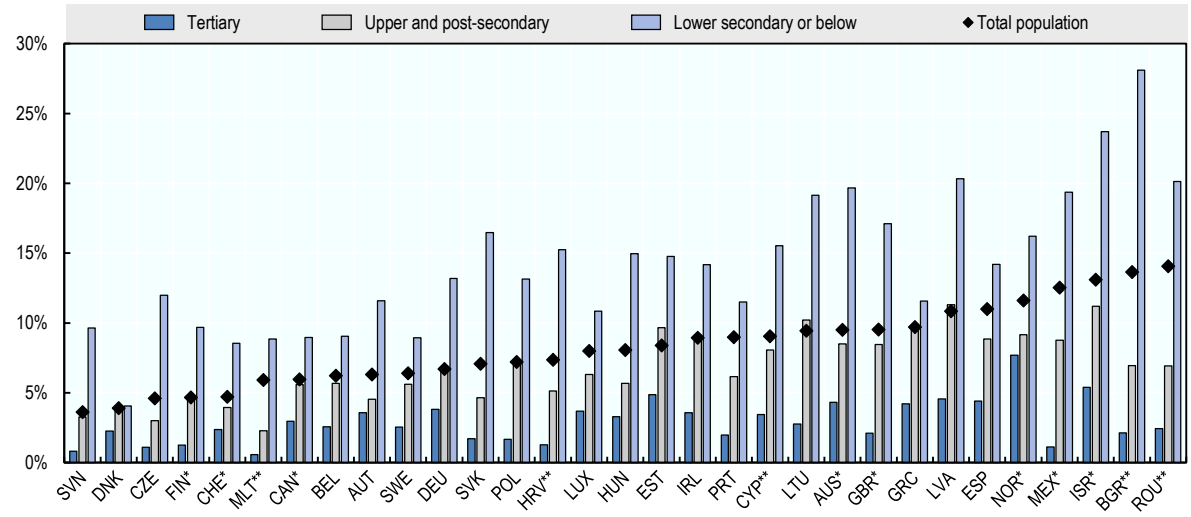

Note: Countries are ranked in ascending order of consumption and asset risk-of-poverty rates among the working-age people living in households of two or more adults without children. Working-age households are defined as those with a head aged 16–64 years. Risk-of-poverty rates for older people are omitted as they are shown for the age group 65 years and over in Figure B.10. Data for Japan, Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_11).

## Figure B.14. Share of individuals who are both income and asset poor, by education level, 2015 or closest available year



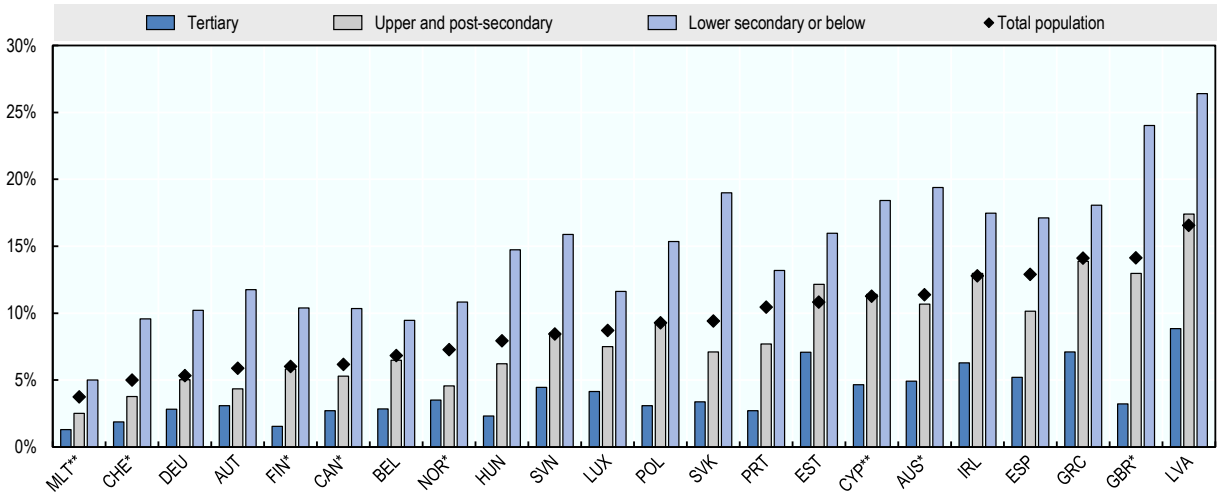

Note: Countries are ranked in ascending order of income and asset risk-of-poverty rates for the total population. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the education level of the head of household rather than the individual.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.15. Share of individuals who are both income and consumption poor, by education level, 2015 or closest available year



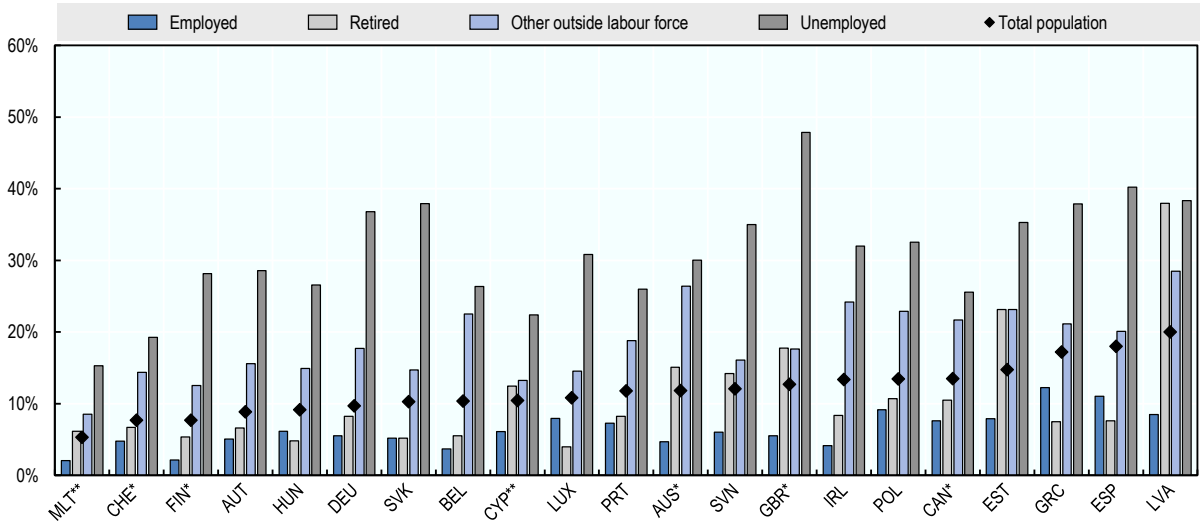

Note: Countries are ranked in ascending order of income and consumption risk-of-poverty rates for the total population. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the education level of the head of household rather than the individual.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.16: Share of individuals who are both consumption and asset poor, by education level, 2015 or closest available year



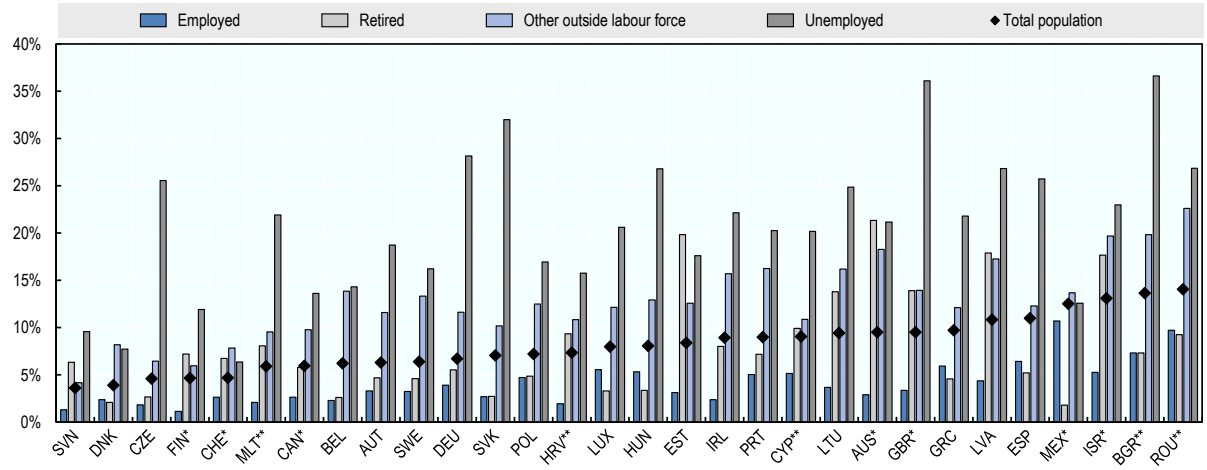

Note: Countries are ranked in ascending order of consumption and asset risk-of-poverty rates for the total population. Data for Japan, Korea and the United States are not available. Data for France and the Netherlands are not shown as they refer to the education level of the head of household rather than the individual.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.17. Share of individuals who are both income and asset poor, by activity status, 2015 or closest available year



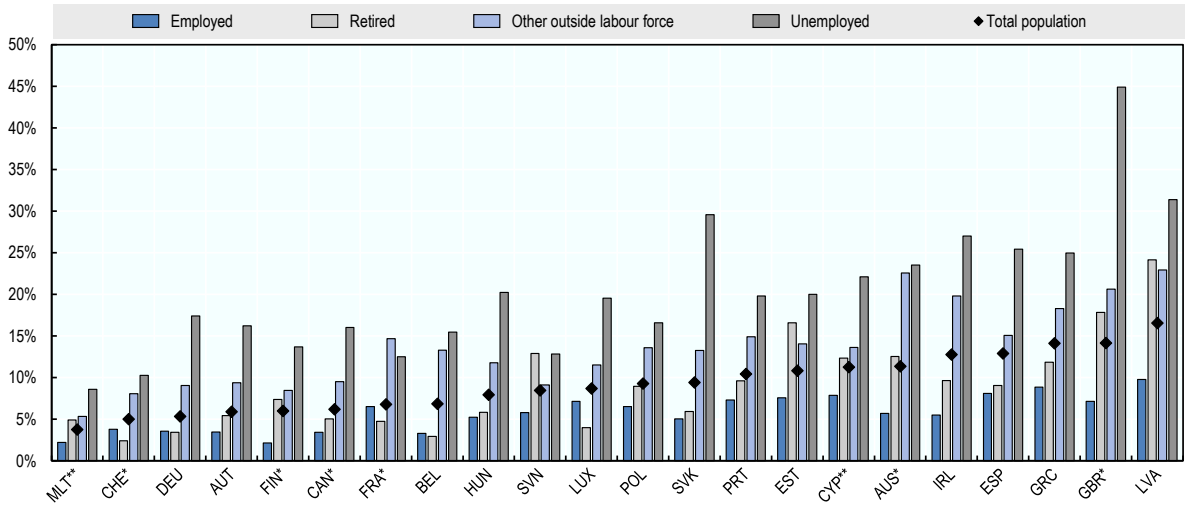

Note: countries are ranked in ascending order of income and asset risk-of-poverty rates for the total population. Data for Japan, Korea, Norway and the United States are not available. Data for France and the Netherlands are not shown as they refer to the activity status of the head of household.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_12).

## Figure B.18. Share of individuals who are both income and consumption poor, by activity status, 2015 or closest available year



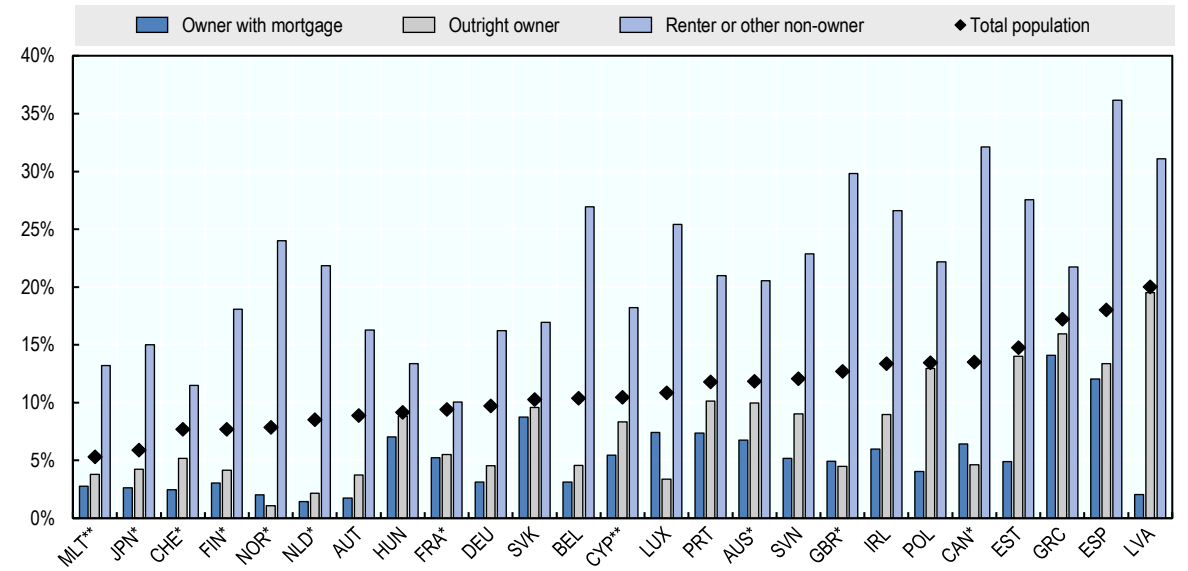

Note: Countries are ranked in ascending order of income and consumption risk-of-poverty rates for the total population. Data for Japan, Korea, Norway and the United States are not available. Data for France and the Netherlands are not shown as they refer to the activity status of the head of household.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_12).

## Figure B.19. Share of individuals who are both consumption and asset poor, by activity status, 2015 or closest available year



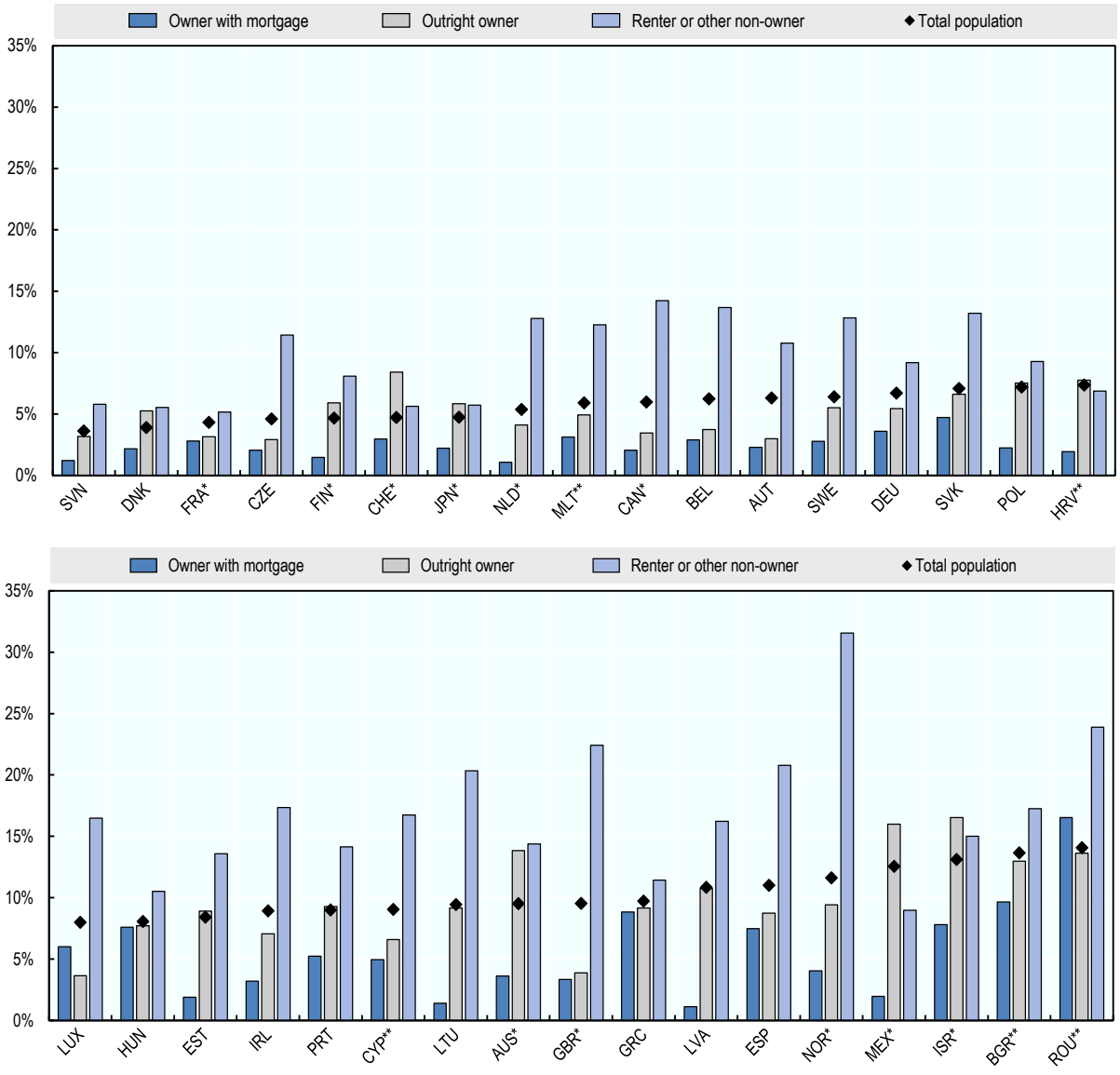

Note: Countries are ranked in ascending order of consumption and asset risk-of-poverty rates for the total population. Data for Japan, Korea, Norway and the United States are not available. Data for France and the Netherlands are not shown as they refer to the activity status of the head of household.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_pov_12).

## Figure B.20. Share of individuals who are both income and asset poor, by tenure status, 2015 or closest available year



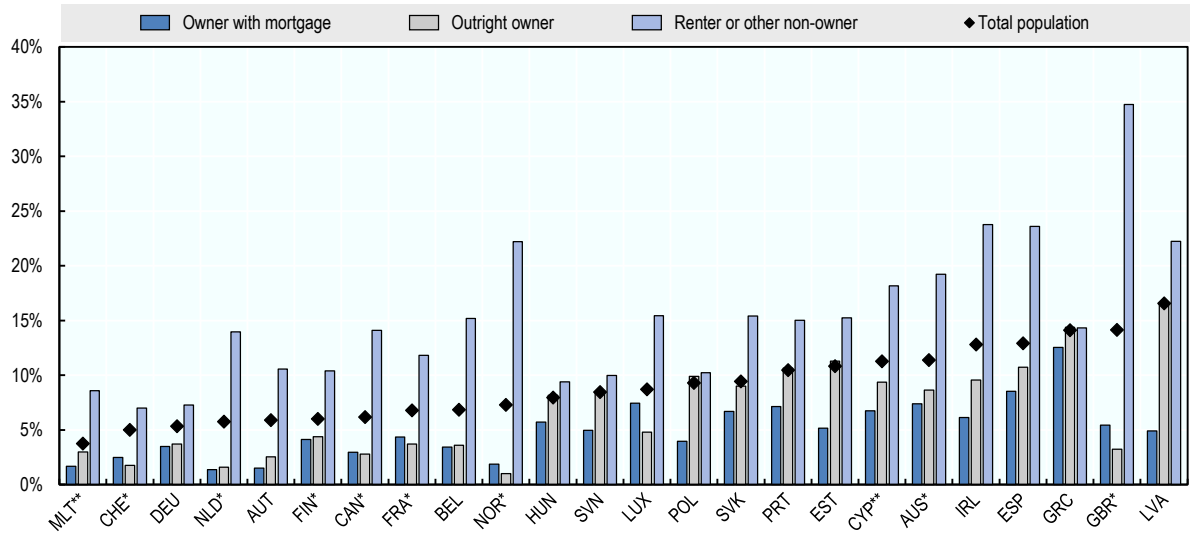

Note: Countries are ranked in ascending order of income and asset risk-of-poverty rates for the total population. Data for Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.21. Share of individuals who are both income and consumption poor, by tenure status, 2015 or closest available year



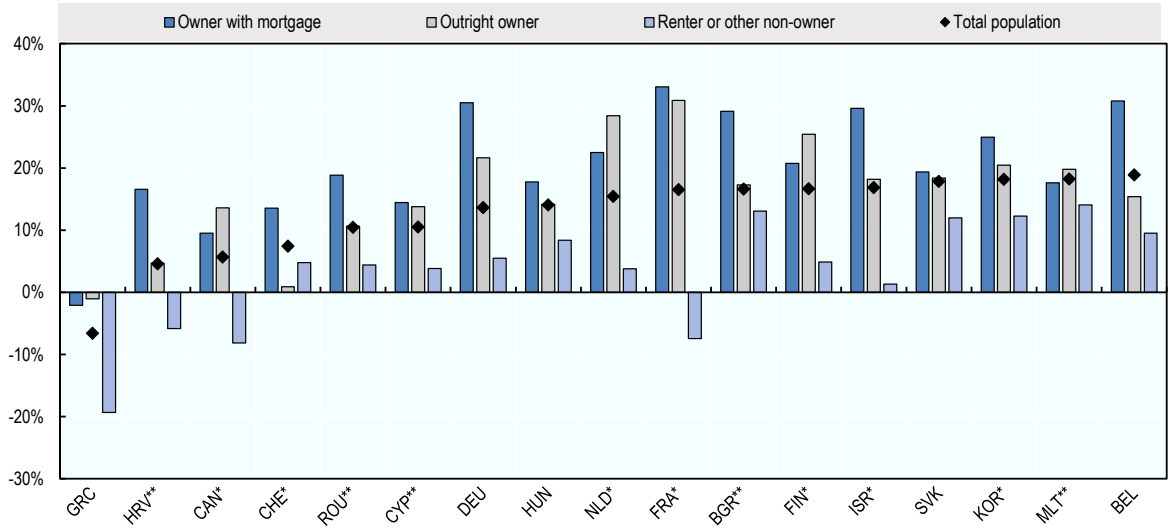

Note: Countries are ranked in ascending order of income and consumption risk-of-poverty rates for the total population. Data for Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

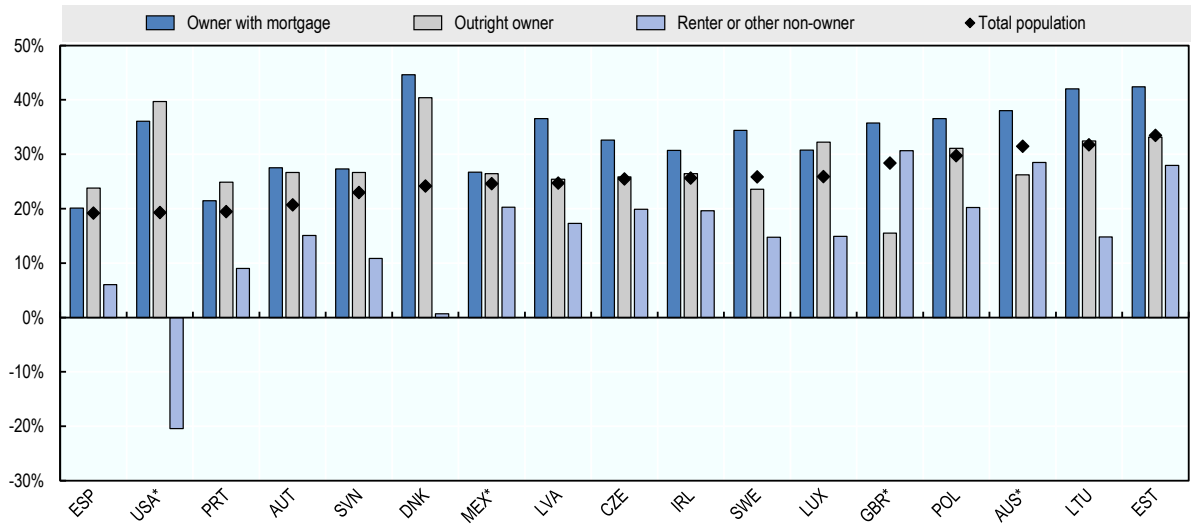
## Figure B.22. Share of individuals who are both consumption and asset poor, by tenure status, 2015 or closest available year




Note: Countries are ranked in ascending order of consumption and asset risk-of-poverty rates for the total population. Data for Japan, Korea and the United States are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

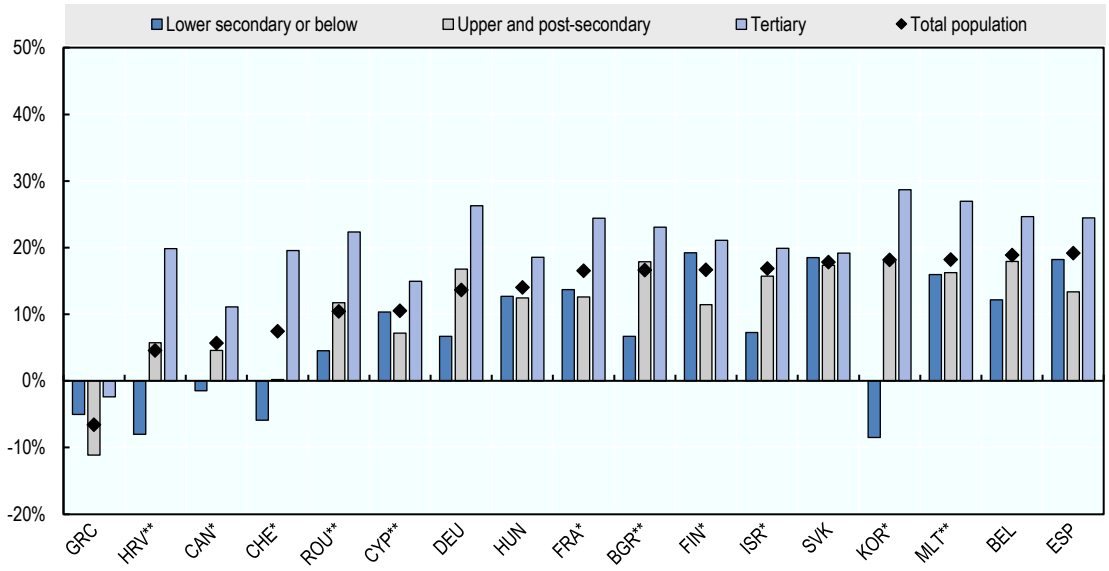
## Figure B.23. Median saving rates, by tenure status, 2015 or closest available year

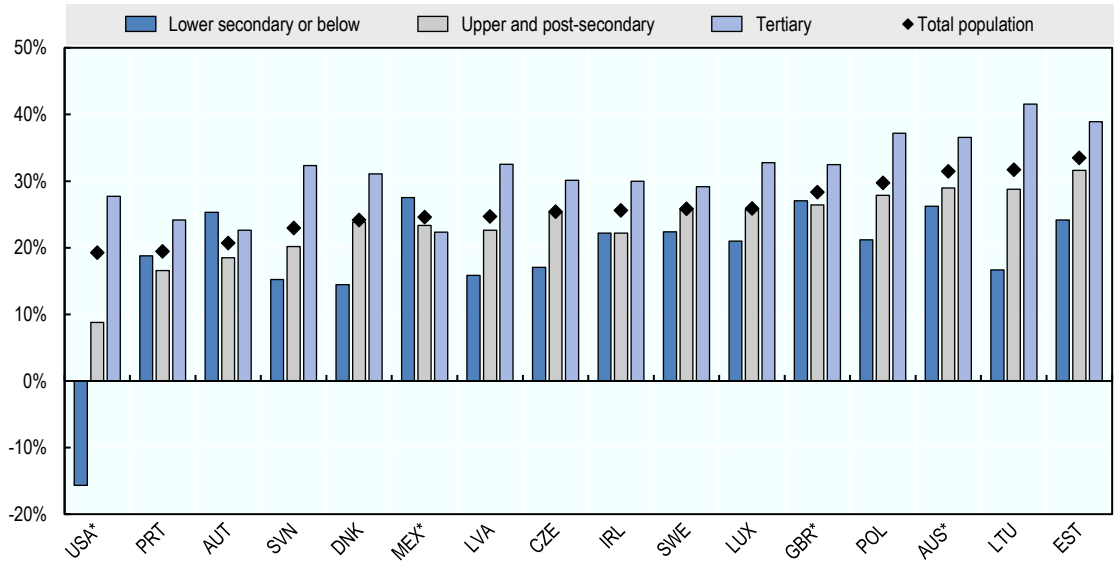





Note: Countries are ranked in ascending order of median saving rates for the total population. Data for Norway are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

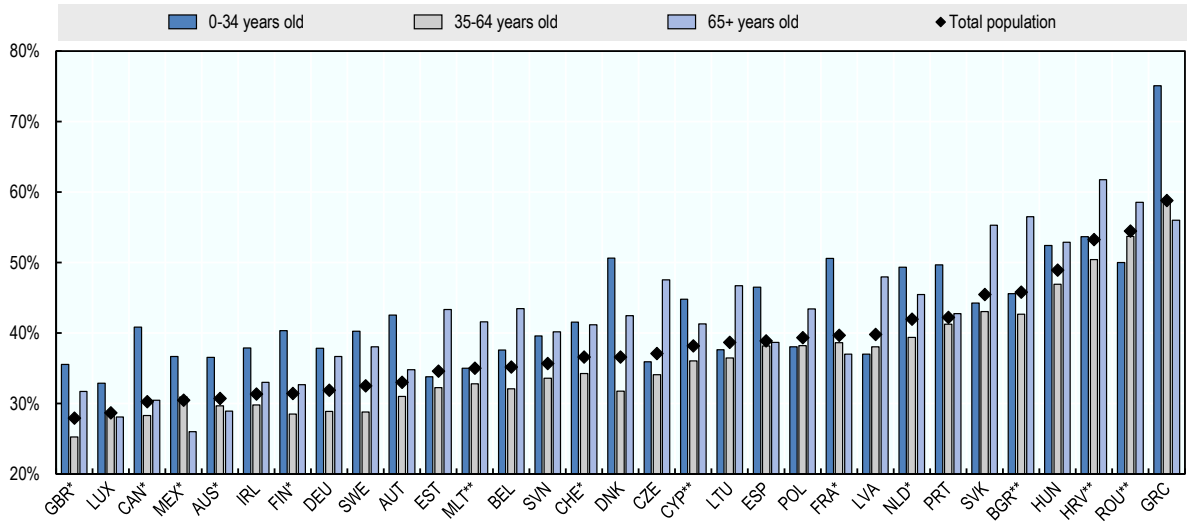
## Figure B.24. Median saving rates, by education level of the household head, 2015 or closest available year





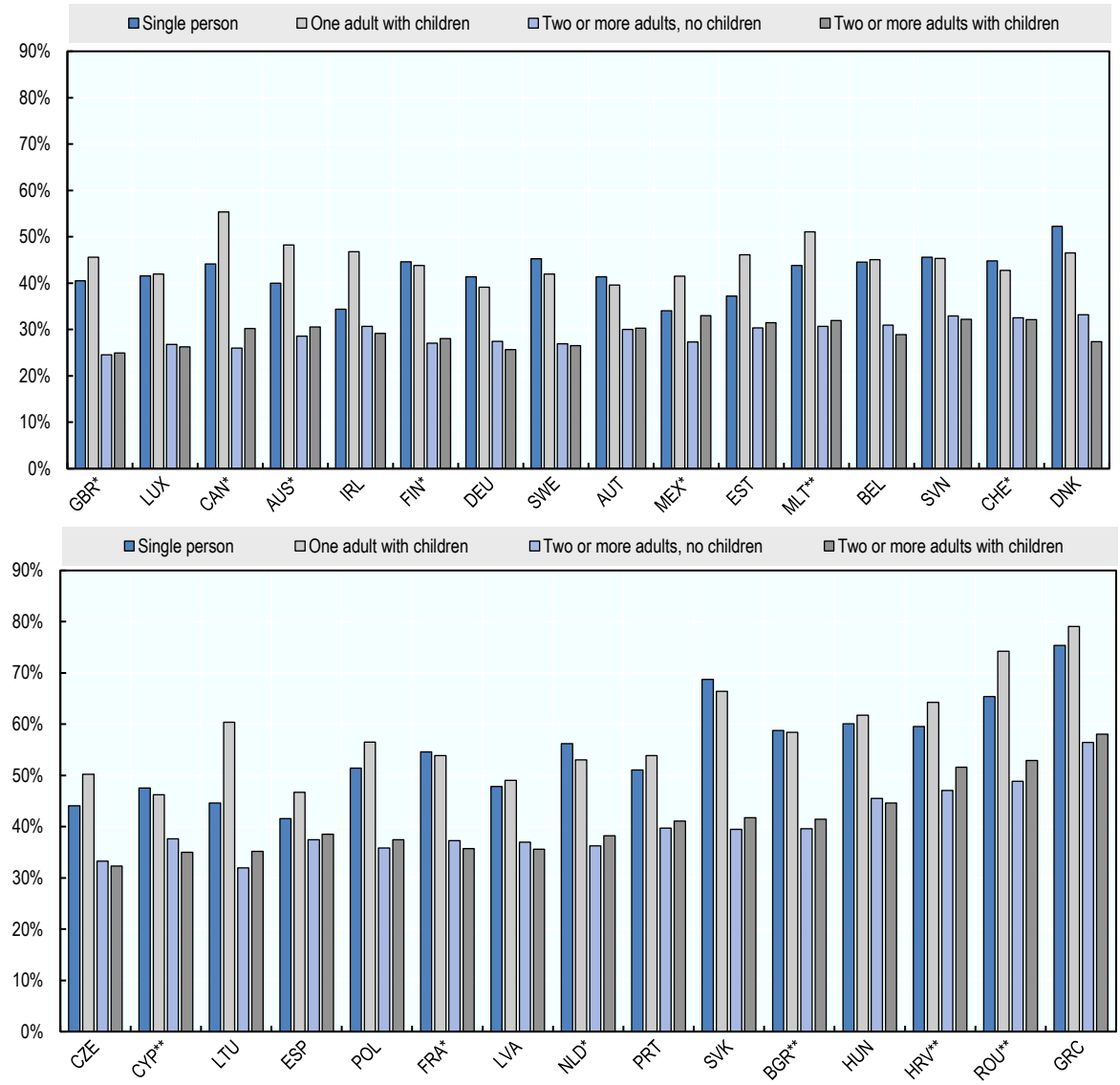

Note: Countries are ranked in ascending order of median saving rates for the total population. Data for Norway and the Netherlands are not available.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth (online data code: icw_sr_04).

## Figure B.25. Essential expenditures as share of disposable income, by age of the household head, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of income spent on essential expenditures for the total population. Values for Switzerland do not include housing-related consumption expenditure other than expenditure on utilities. In Eurostat's centralised exercise, essential expenditures do not include mortgages and information on rents is underestimated or unavailable for some of the countries. Data are not available for Israel, Japan, Korea, Norway and the United States.
Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.

## Figure B.26. Essential expenditures as share of disposable income, by household type, working-age households, 2015 or closest available year




Note: Countries are ranked in ascending order of the share of income spent on essential expenditures for the total population. Values for Switzerland do not include housing-related consumption expenditure other than expenditure on utilities. In Eurostat's centralised exercise, essential expenditures do not include mortgages and information on rents is underestimated or not available for some of the countries covered in the centralised exercise. Data are not available for Israel, Japan, Korea, Norway and the United States. Working-age households are defined as those with a head aged 16–64 years.

Source: Eurostat-OECD Secretariat calculations based on national estimates and Eurostat's experimental statistics on income, consumption and wealth.